

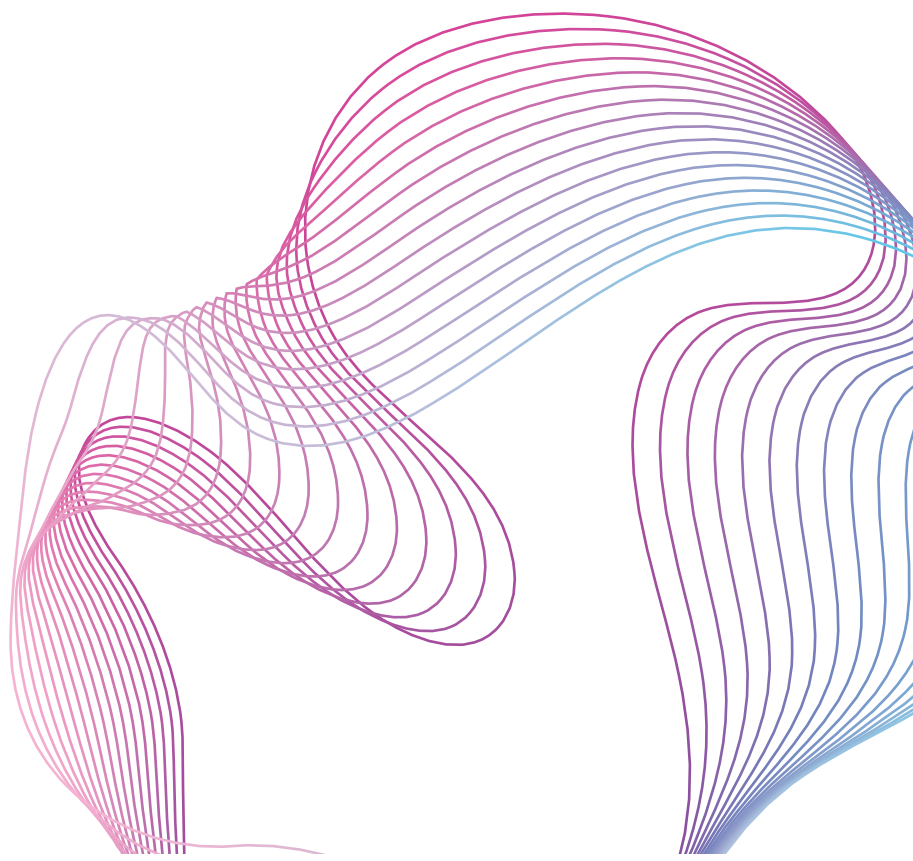
ИТМО

АЛЬМАНАХ НАУЧНЫХ РАБОТ

**молодых ученых
Университета ИТМО**

Санкт-Петербург,
2-5 февраля 2022 г.

Том 2



МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ
НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ УНИВЕРСИТЕТ ИТМО

**АЛЬМАНАХ НАУЧНЫХ РАБОТ
МОЛОДЫХ УЧЕНЫХ
Университета ИТМО**

Том 2

ИТМО

Санкт-Петербург
2022

УДК 004, 681,

ББК 22, 32, 85

Альманах научных работ молодых ученых Университета ИТМО. Том 2. – СПб.: Университет ИТМО, 2022. – 424 с.

Издание содержит результаты научных работ молодых ученых, доложенные на Пятьдесят первой (LI) научной и учебно-методической конференции Университета ИТМО, проходившей с 2 по 5 февраля 2022 г., по тематикам: системы управления и робототехника, безопасность информационных технологий, программная инженерия и компьютерная техника; фотоника; информационные технологии и программирование; инфокоммуникационные технологии; дизайн и урбанистика.

ISBN 978-5-7577-0669-6

ISBN 978-5-7577-0671-9 (Том 2)

The logo for ITMO University, consisting of the letters 'ITMO' in a bold, black, sans-serif font. The letter 'I' has a small dot above it, and the 'O' is a solid circle.

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 и искусственный интеллект, 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). Представлен в мировом ТОП-200 по телекоммуникационным технологиям (Telecommunication engineering), а также в ТОП-300 по нанонаукам и нанотехнологиям (Nanoscience & Nanotechnology) ARWU. Входит в ТОП-200 по инженерным наукам (Engineering and Technology), в ТОП-300 по физике и астрономии (Physics & Astronomy), наукам о материалах (Materials Sciences), а также по машиностроению, аэрокосмической и промышленной инженерии (Mechanical, Aeronautical & Manufacturing Engineering) рейтинга QS. Лидер проекта «Приоритет - 2030».

© Университет ИТМО, 2022

© Авторы, 2022

РЕДАКЦИОННАЯ КОЛЛЕГИЯ

по направлению системы управления и робототехника, безопасность информационных технологий, программная инженерия и компьютерная техника

Председатель редколлегии:

*профессор, доктор технических наук **Бобцов Алексей Алексеевич***

Члены редколлегии:

*профессор, доктор технических наук **Пыркин Антон Александрович***

*кандидат технических наук **Кустарев Павел Валерьевич***

*доцент, кандидат технических наук **Заколдаев Данил Анатольевич***

*доцент, кандидат технических наук **Андреев Юрий Сергеевич***

по направлению фотоника

Председатель редколлегии:

*профессор, доктор физико-математических наук **Бугров Владислав Евгеньевич***

Члены редколлегии:

*доцент, кандидат технических наук **Бахолдин Алексей Валентинович***

*профессор, доктор технических наук **Вейко Вадим Павлович***

*доцент, кандидат технических наук **Вознесенская Анна Олеговна***

*профессор, доктор технических наук **Кортаев Валерий Викторович***

*профессор, доктор технических наук **Успенская Майя Валерьевна***

по направлению информационные технологии и программирование; инфокоммуникационные технологии; дизайн и урбанистика

Председатель редколлегии:

*доктор технических наук **Бухановский Александр Валерьевич***

Члены редколлегии:

*кандидат технических наук **Ульянцев Владимир Игоревич***

*доцент, доктор технических наук **Басов Олег Олегович***

*доцент, доктор технических наук **Духанов Алексей Валентинович***

*кандидат технических наук **Капитонов Александр Александрович***

*кандидат технических наук **Митягин Сергей Александрович***

ВВЕДЕНИЕ

Издание содержит результаты научных работ молодых ученых, доложенные на Пятьдесят первой (LI) научной и учебно-методической конференции Университета ИТМО, проходившей 2 – 5 февраля 2022 г., по тематикам: системы управления и робототехника, безопасность информационных технологий, программная инженерия и компьютерная техника; фотоника; 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 № 1251 от 16 октября 2017 года "Об утверждении Правил предоставления субсидии из федерального бюджета на оказание государственной поддержки центров Национальной технологической инициативы на базе образовательных организаций высшего образования и научных организаций и Положения о проведении конкурсного отбора на предоставление грантов на государственную поддержку центров Национальной технологической инициативы на базе образовательных организаций высшего образования и научных организаций", 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, Государственной программы Российской Федерации «Научно-технологическое развитие Российской Федерации», включающей в себя федеральные проекты национального проекта "Наука и университеты" и Национальной программы "Цифровая экономика Российской Федерации", в рамках федерального проекта "Развитие масштабных научных и научно-технологических проектов по приоритетным исследовательским направлениям" национального проекта "Наука и университеты" в университете выполняются работы по Федеральной научно-технической программе развития синхротронных и нейтронных исследований и исследовательской инфраструктуры на 2019 – 2027 годы (ПП 16 марта 2020 г. № 287) и по инициативным научно-исследовательским проектам, выполняемыми преподавателями, научными сотрудниками, молодыми учеными, 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.

КОМПЬЮТЕРНЫЕ ТЕХНОЛОГИИ И УПРАВЛЕНИЕ

УДК 004.932

СЕМАНТИЧЕСКАЯ СЕГМЕНТАЦИЯ КАК ЧАСТНЫЙ СЛУЧАЙ ЗАДАЧИ КЛАССИФИКАЦИИ

Антонов Р.М.¹ (студент)

Научный руководитель – канд. техн. наук Щемелинин В.Л.²

¹*Университет ИТМО*

²*ООО «Сибур Диджитал»*

e-mail: ruslanantonof@yandex.ru

В данной работе проведено исследование метода семантической сегментации применительно к задаче классификации. Проанализированы достоинства и недостатки метода, исследованы особенности его практического применения. Опираясь на полученные знания, была решена реальная практическая задача по контролю качества продукции при производстве термоэластопластов.

Ключевые слова: классификация, сегментация, семантическая сегментация, машинное обучение, компьютерное зрение.

Классификация – раздел машинного обучения, посвященный решению задачи соотнесения некоторого множества объектов к заранее известным классам, для которых существует конечное множество объектов, определяющих классы. Это множество называется обучающей выборкой. Задача классификации относится к разделу задач машинного обучения с учителем.

Существует 4 метода решения задачи классификации:

1. Распознавание объектов.
2. Детектирование объектов.
3. Семантическая сегментация.
4. Сегментация экземпляров.

Задача распознавания объектов состоит в том, чтобы обнаружить на переданном на вход изображении некий объект, определить его класс и вернуть координаты прямоугольника, содержащего этот объект. Задача детектирования объектов схожа с предыдущей, однако, изображение может содержать множество объектов, относящихся к разным классам. Так же, как и в предыдущей задаче, на выходе ожидается предсказанный класс и координаты прямоугольника для каждого из объектов на изображении.

Задача семантической сегментации состоит в разделении изображения на отдельные группы пикселей, соответствующих некоторому объекту, и дальнейшему определению класса объекта для каждой группы. Данный метод применяется тогда, когда необходимо не просто определить координаты прямоугольника, содержащего объект, но и точно определить его контуры. Благодаря такому принципу работы, обучение систем, построенных с использованием семантической сегментации, требует обучающей выборки меньших размеров, так как, по сути, каждый пиксель изображения является маленьким обучающим примером. Однако, это же является и недостатком этого метода, так как предсказание класса для каждого пикселя может занимать значительное количество времени. Так же особенность метода заключается в том, что он выделит два стоящих рядом объекта одного класса как один объект. Задачу дифференцированного сегментирования изображения решают методы сегментации экземпляров, являющиеся по сути развитием идеи семантической сегментации.

Одной из наиболее известных реализаций метода семантической сегментации является U-Net. U-Net – это сверточная нейронная сеть, разработанная в 2015 году О. Роннебергером, Ф. Фишером и Т. Броксом для решения задачи сегментации клеток на микроскопических снимках [1].

Архитектура данной сети, показана на рисунке 1, состоит из двух частей: сужающейся (кодировщика) и расширяющейся (декодировщика). Кодировщик представляет собой сверточную нейронную сеть, состоящую из блоков, содержащих два сверточных слоя 3×3 , с функцией активации ReLU и пакетной нормализации, применяемой к каждому из них, а также слоя понижения дискретизации с окном 2×2 . Декодировщик имеет такое же количество блоков, что и кодировщик. Каждый блок декодировщика состоит из слоя повышения дискретизации с окном 2×2 , слияния с соответствующим набором признаков из кодировщика, двух сверточных слоев с фильтрами 3×3 и функцией активации ReLU, примененной к каждому из них. Между кодером и декодером существуют сквозные соединения (skip-connections), с помощью которых карты признаков из кодировщика передаются в декодировщик, что позволяет сохранять пространственную информацию (контекст). Наличие таких соединений также ускоряет работу сети и помогает снизить необходимый для обучения размер обучающей выборки. Последний слой использует сигмоидную функцию активации признаков для классификации на уровне пикселей.

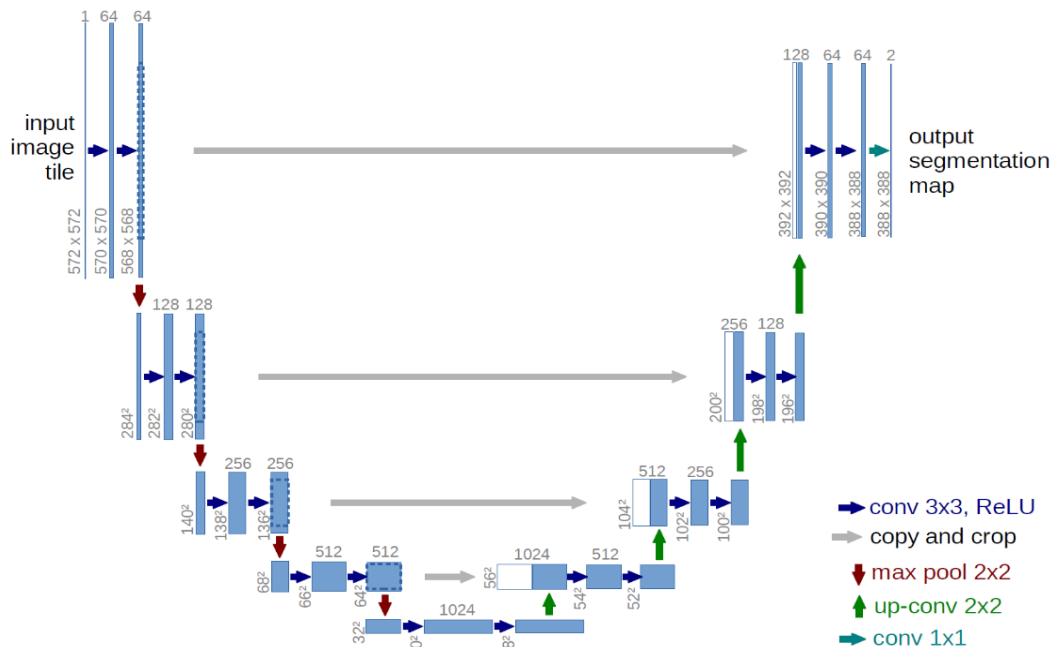


Рис. 1. Архитектура U-Net

Как уже было сказано, изначально U-Net создавался для решения задач биомедицинского характера. Однако, вскоре выяснилось, что архитектура позволяет решать многие другие задачи, в том числе сегментацию спутниковых снимков, видеопотока в реальном времени.



Рис. 2. Внешний вид готового термоэластопласта

Архитектура нашла применение в задачах промышленности, дефектологии, контроля качества [2, 3]. В рамках проблемы контроля качества существует задача детектирования и классификации агломератов при производстве термоэластопластов, показанных на рисунке 2. Сложность задачи состоит в том, что агломераты вещества могут иметь различную сложную форму, а также не иметь четко выраженных краев. При этом, так как задача существует в рамках реального производства, и изображение передается с производственных установок, возникают искажения, связанные с изменяемым освещением и особенностями технического процесса.

Учитывая специфику задачи, был сделан вывод, что наиболее оптимальным методом ее решения является семантическая сегментация. За основу была взята архитектура U-Net, которая была модифицирована таким образом, чтобы одновременно с сегментацией проводить многоклассовую классификацию. Для ее обучения был собран датасет изображений агломератов как нормального, так и слишком большого размера, содержащий несколько сотен изображений. Затем была проведена ручная разметка данных. Был проведен ряд опытов по подбору оптимальных параметров обучения сети. На выход нейросеть передает изображение, объекты на котором размечены и классифицированы по признаку принадлежности к одному из классов. В результате была обучена нейросеть, способная с высокой точностью проводить сегментацию и классификацию агломератов вещества. Примеры работы обученной нейросети представлены на рисунках 3 и 4.



Рис. 3. Классификация агломератов нормального размера



Рис. 4. Классификация агломератов слишком большого размера

Таким образом, можно сделать вывод, что метод семантической сегментации (и в частности, архитектура U-Net) хорошо проявляет себя в задачах сегментирования и классификации объектов сложной формы. Благодаря особенностям обучения метод может быть применен к объектам произвольной формы. Также метод хорошо проявляет себя при работе с объектами, обладающими нечеткими краями. Семантическая сегментация показывает достаточно высокую устойчивость к внешним искажениям на изображении. Также значительным преимуществом метода является то, что для обучения сети достаточно небольшого количества обучающих данных, что особенно актуально в ситуациях, когда собрать большой датасет для обучения модели является проблематичным.

Литература

1. Ronneberger O., Fischer P., Brox T. U-Net: Convolutional networks for biomedical image segmentation // International Conference on Medical Image Computing and Computer-Assisted Intervention - MICCAI 2015. Lecture Notes in Computer Science. – 2015. – V.9351. – Pp. 234–241.
2. Neven R., Goedemé T.A Multi-Branch U-Net for Steel Surface Defect Type and Severity Segmentation. Metals 2021. – V.11. – Pp. 870.
3. Гаскаров Р.Д., Бирюков А.М., Никонов А.Ф., Агниашвили Д.В., Хайрисламов Д.А. Определение дефектов на стальных листах с использованием сверточных нейронных сетей. Электронные библиотеки. – 2020. – Т.23. – №6. – С. 1155– 1171.

УДК 004.056

ИССЛЕДОВАНИЕ ЭФФЕКТИВНОСТИ ВНЕДРЕНИЯ HONEYPOT-РЕШЕНИЙ В ОБЛАСТИ ИОТ

Барина Я.В.¹ (студент)

Научный руководитель – канд. техн. наук, доцент Меншиков А.А.¹

¹*Университет ИТМО*

e-mail: byvvv@yandex.ru, menshikov@itmo.ru

В работе проанализированы существующие honeypot-решения для устройств промышленного Интернета вещей (IIoT), рассмотрена эффективность данных решений и их преимущества перед другими средствами защиты информации. Полученные результаты используются для составления требований к разрабатываемой системе.

Ключевые слова: система honeypot, IIoT устройства, промышленный Интернет вещей, информационная безопасность, киберугрозы.

В настоящее время концепция IIoT устройств стала неотъемлемой частью повседневной жизни, предоставляя пользователям таких устройств уникальные возможности, начиная с автоматизации простейших бытовых вопросов и заканчивая медицинскими устройствами жизнеобеспечения. IIoT устройства способны обрабатывать и передавать большие массивы данных, а также отправлять информацию о состоянии в режиме реального времени. Однако, IIoT устройства не лишены проблем с безопасностью и конфиденциальностью – с ростом их количества, растет количество и масштаб киберпреступлений на эти устройства.

В результате стремительного развития технологического сектора возрастает количество устройств промышленного контроля в критически важной инфраструктуре, что приводит к увеличению числа потенциальных кибератак. IIoT (Industrial Internet of Things) относятся к промышленному подмножеству устройств Интернет вещей и представляют собой автоматизацию бизнес-процессов предприятий.

Существующие средства защиты умных устройств, такие как Межсетевой экран (firewall) и системы обнаружения вторжений (IDS, Intrusion Detection Systems) являются достаточно эффективными для обнаружения некоторых типов вредоносной активности и для выявления аномального поведения в сети предприятия. Но этим инструментам зачастую не хватает возможности обнаружить новые угрозы, такие как уязвимости нулевого дня, основанные на атаках нулевого дня. Также эти инструменты не могут собрать больше информации о злонамеренных действиях злоумышленника, его полезной нагрузке (payload), доступных эксплоитах (exploit) и навыках ввиду отсутствия сигнатур атак нулевого дня в соответствующих базах [1].

Одним из способов прогнозирования и предотвращения атак на устройства промышленной инфраструктуры может выступать технология ловушек (honeypot). Технология позволяет провести анализ поведения киберпреступника в сети и определить, каким образом могут быть нанесены удары по существующим объектам безопасности. Одним из преимуществ данных систем является быстрое обнаружение инцидентов, потому что любое действие против такой системы или взаимодействие с ней является аномалией. Собранные данные о сетевом трафике впоследствии могут быть изучены для разработки современных средств противодействия.

В зависимости от цели, honeypot-решения можно разделить на два класса: производственные и исследовательские. Производственные honeypot нацелены на снижение рисков в производственной среде и используются для решения основных задач обеспечения информационной безопасности. Исследовательские honeypot применяются для сбора и анализа инцидентов, с целью разработки современных средств защиты от атак. Как правило, исследовательские решения не уменьшают риски для предприятия, но полученные с их помощью данные могут найти свое применение на реальных промышленных системах, например для доработки систем обнаружения, предупреждения и протоколирования [2].

В таблице приведен краткий обзор некоторых наиболее популярных фреймворков honeypot для устройств промышленного Интернета вещей. Анализ решений для устройств IIoT позволил классифицировать honeypots по разным основаниям. Существуют различные реализации honeypot/honeynet, отличия которых между собой заключаются в уровне взаимодействия, масштабируемости, эмулируемых сервисах. Некоторые из данных решений являются бесплатными (Open Source).

Таблица. Классификация honeypot/honeynet для устройств IIoT

Решение	Год	Уровень взаимодействия	Масштабируемость	Open Source	Эмулируемые сервисы
Conpot	2013	низкий	✓	✓	IEC 60870-5-104, BACnet, EtherNet/IP, Guardian AST, Kamstrup, Modbus, S7comm, HTTP, FTP, SNMP, IPMI, TFTP
DiPot	2018	низкий	✓		HTTP, Modbus, Kamstrup, SNMP, IMPI, BAC-net, Guardian AST, S7comm
SIPHON	2017	высокий	✓	✓	SSH, HTTP
Piggin	2016	высокий			SSH, HTTP, RDP
ShaPe	2015	низкий	✓	✓	IEC 61850 MMS, HTTP, FTP, SMB
Murillo	2018	низкий	✓	✓	EtherNet/IP
Haney	2019	высокий	✓		Modbus/TCP, SSH, Telnet, SNMP, HTTP
HoneyIo3	2018	смешанный	✓		IPMI, S7comm, Kamstrup, SNMP, HTTP(S), Ntopng, SSH
Du and Wang	2020	смешанный	✓		N/A

С точки зрения назначения honeypot и honeynet решений, рассмотренных в таблице, большинство фреймворков нацелены на исследование инцидентов. Единственный рассмотренный honeypot, преследующий производственные цели – Piggin et al. И данное решение имеет закрытый исходный код.

Исследование работы и эффективности рассмотренных решений проводилось инструментами Nmap, Wireshark, Snort IDS, Shodan, TCPDUMP, syslog и другими.

Фреймворки поддерживают широкий спектр протоколов, которые специфичны для промышленных сред. Наиболее популярные из них - Modbus, S7comm, EtherNet/IP. В исследованиях также распространены непромышленные протоколы - HTTP, SNMP. HTTP используется в качестве интерфейса HMI промышленных систем, а также для удаленной настройки промышленных устройств. SNMP применяется в целях мониторинга и управления промышленных объектов [3].

В общем виде honeypot система представляет собой исполнение промышленной сети. Главным условием является изолирование сегмента сети с промышленными устройствами и сегмента с установленной ловушкой, во избежание реального получения доступа к объектам в процессе исследовательской деятельности. Каждый экземпляр системы должен иметь строго определенную функцию в сети, однако, требования к их реализации остаются одинаковыми [4]. Предлагаемая в работе архитектура honeypot-системы предполагает наличие следующих ключевых особенностей:

1. Для исследовательских honeypot-решений важно удерживать злоумышленника в системе в течение длительного времени, что означает, что система должна быть максимально реалистичной (vetting system).
2. Система не должна содержать официально опубликованных уязвимостей CVE, чтобы атакующие с большей вероятностью стали применять неопубликованные атаки нулевого дня на систему honeypot.
3. Каждый экземпляр системы honeypot обеспечивает защиту от неправомерного доступа и не раскрывает реальную сеть организации (граница honeypot должна быть прочной).
4. Honeypot-система обеспечивает протоколирование происходящих действий злоумышленника сразу после его проникновения внутрь IT-инфраструктуры предприятия. Каждое его действие фиксируется, включая использование вредоносного и иного программного обеспечения, для возможности их дальнейшего анализа.

Концепция honeypot-решений является важным и эффективным инструментом для обеспечения безопасности устройств промышленных Интернет вещей. Honeypot-система может снизить количество ложных срабатываний, выдаваемых другими средствами защиты информации, внедренными в промышленную систему, такими как IDS/IPS-системы, firewall.

Далее, на основе сформулированных требований, строится модель honeypot системы, направленная на решение следующих задач:

- выявление ранее неизвестных уязвимостей в промышленной сети;
- выявление ранее неизвестного программного обеспечения;
- пошаговое изучение действий киберпреступников в системе;
- сбор сведений об используемой инфраструктуре злоумышленника.

Литература

1. Архипова А.Б., Каревский Д.Р. Honeypot как инструмент создания эффективной защищенной системы // Безопасность цифровых технологий. – 2021. – №2(101). – С. 122–135.
2. Honeypots in Network Security. [Электронный ресурс]. – Режим доступа: <https://www.diva-portal.org/smash/get/diva2:327476/fulltext01> (дата обращения: 01.03.2022).
3. Javier F., Ahmet A., Berk C., A. Uluagac «A Survey of Honeypots and Honeynets for Internet of Things, Industrial Internet of Things, and Cyber-Physical Systems» (PDF) // arXiv – CS – Cryptography and Security (IF), arxiv-2108.02287.
4. Поляков В.В., Лапин С.А. Архитектура системы Honeypot для изучения целевых атак // XIV Международная научно-техническая конференция АПЭП – 2018. – Т.3. – С. 173–175.

УДК 67.05

РАЗРАБОТКА МЕТОДИКИ АКТИВНОГО КОНТРОЛЯ РЕЖУЩЕГО ИНСТРУМЕНТА НА СТАНКАХ С ЧПУ В УСЛОВИЯХ БЕЗЛЮДНОГО ПРОИЗВОДСТВА

Басова Т.В.¹ (аспирант)

Научный руководитель – канд. техн. наук, доцент Андреев Ю.С.¹

¹*Университет ИТМО*

e-mail: tvbasova@itmo.ru, ysandreev@corp.ifmo.ru

В работе рассмотрены задачи контроля режущего инструмента в условиях безлюдного производства, приведены основные особенности существующих методик активного контроля, а также основы разрабатываемой методики активного контроля режущего инструмента в условиях безлюдного производства, имеющей некоторые преимущества по сравнению с существующими производственными методиками. Новая методика основана на авторских алгоритмах функционирования станка с числовым программным управлением (ЧПУ), датчиков наладки режущего инструмента и постпроцессора станка с ЧПУ. Внедрение представленной методики автоматизирует процесс измерения инструмента, способствует поддержанию автономной работы оборудования при изменении технологических характеристик обработки без дополнительной установки дорогостоящих датчиков на оборудование.

Ключевые слова: износ режущего инструмента, активный контроль, станки с ЧПУ, датчики наладки инструмента, безлюдное производство.

Износ РИ является одним из значимых факторов, влияющих на качество изготовления деталей на станках с ЧПУ [1]. Несвоевременное выявление сломанного или износившегося режущего инструмента (РИ) может привести к производственному браку, порче оборудования, простоем производства и прочим негативным последствиям. Кроме того, износ РИ приводит к снижению точности механической обработки, увеличению погрешностей формы и размера обрабатываемых поверхностей деталей, что может в дальнейшем вызывать сложности при сборке деталей, повышенное трение сопрягающихся поверхностей и вибрации, и, следовательно, сокращение срока эксплуатации механизмов и машин. В связи с этим, в процессе изготовления деталей на станках с ЧПУ вопросы повышения качества РИ и его контроля являются важным, особенно в связи с тенденцией к минимизации человеческого вмешательства в технологические и производственные процессы.

Повышение производительности механической обработки на станках с ЧПУ при изготовлении деталей машиностроения и приборостроения возможно обеспечить путем перехода предприятий на безлюдное производство, функционирующее на максимально допустимых режимах резания, с минимальными простоями. Явной проблемой как автоматизированных, так и неавтоматизированных производств является отсутствие надежных производственных методик активного контроля и измерения РИ, обеспечивающих своевременное выявление превышения допустимого износа РИ. Результатами этого являются:

- занижение режимов резания в 2, а в некоторых случаях и в 5 раз от рекомендуемых производителями РИ, что значительно снижает скорость изготовления детали;
 - простой, вызванный отказом РИ, и, следовательно, низкая эффективность производства;
 - высокая вероятность технологического брака;
 - необходимость частых доработок деталей или неисправимый брак;
 - увеличение себестоимости изготовления деталей.
-

На сегодняшний день разработка методик активного контроля РИ с целью обеспечения бесперебойной работы безлюдного производства является актуальной задачей. Под активным контролем понимается контроль в процессе обработки, по результатам которого возможна корректировка каких-либо размеров детали [2]. Существующие методики активного контроля основаны на методах:

1. Акустической эмиссии.
2. Виброакустической диагностики.
3. Контроля температуры в зоне резания.
4. Анализа силовых характеристик.

Необходимо отметить, что в большинстве случаев, методики активного контроля РИ основываются на одновременном применении нескольких датчиков и методах обработки данных [3], что значительно увеличивает стоимость системы мониторинга, а также требует дальнейшей трудоемкой обработки данных и разработки сложных математических моделей [4]. Методы активного контроля РИ, основанные, например, на акустической эмиссии и виброакустической диагностике могут быть малонадежными из-за влияния других факторов механической обработки на результаты измерений. Многие системы мониторинга не могут приемлемо адаптироваться к изменению технологических характеристик из-за использования недостаточно точных алгоритмов функционирования. Указанные недостатки делают актуальным разработку простой и экономичной системы мониторинга состояния РИ с минимальным количеством устанавливаемых датчиков и высокой эффективностью контроля.

В данной работе предлагается решение вышеперечисленных недостатков посредством разработки новой авторской методики активного контроля РИ, основанной на интеграции общеприменимых датчиков наладки РИ. Указанная интеграция настраивается при помощи совместной работы макросов станков с ЧПУ и разработанного постпроцессора станка. Активный контроль РИ в этой методике обеспечивается за счет контроля износа РИ посредством измерения его геометрических размеров в начале и в конце технологических переходов при помощи бесконтактных датчиков наладки инструмента для выявления поломки, а также корректировки в процессе обработки координат движения РИ на станке с ЧПУ на величину выявленного износа.

Датчики наладки инструмента применяются для измерения и определения нулевой точки РИ на станках с ЧПУ.



Рис. 1. Блок-схема последовательности изготовления детали при функционировании по стандартной производственной методике контроля РИ на станках с ЧПУ

Управление датчиками наладки инструмента осуществляется, в основном, посредством подачи через систему управления ЧПУ программы измерения инструмента – цикл измерения инструмента. Параметры цикла измерения отвечают за точки измерения РИ на датчике наладки, а также влияют на характеристики процесса измерения.

Параметры цикла измерения должны быть выбраны в зависимости от типа РИ его конструкторско-технологических характеристик; в случае если подстановки некорректных значений в цикл инструмента, не подходящих данному РИ, высок риск производственного брака за счет смещения нулевой точки инструмента на станке с ЧПУ, а также сбой работы самого датчика наладки инструмента. В основном, в существующих производственных методиках данные цикла вносятся неавтоматизированным методом, что повышает риск ошибки измерения по причине человеческого фактора [5], что особенно критично для РИ со сложной геометрической формой.

Последовательность изготовления детали при функционировании станка по стандартной производственной методике контроля РИ, основанной на применении датчиков наладки РИ представлен на рисунке 1.

Основными недостатками функционирования стандартной методики контроля РИ являются:

- отсутствие активного контроля РИ;
- потребность ручного ввода данных в систему управления ЧПУ;
- невозможность измерения и контроля РИ со сложной геометрической формой;
- отсутствие автоматической адаптации к изменению технологических характеристик обработки.



Рис. 2. Блок-схема функционирования разрабатываемой методики активного контроля РИ

Вышеописанная методика позволит выполнять активный контроль РИ посредством автоматизации заполнения параметров цикла измерения РИ, а также частично обеспечить точность изготовления элементов детали, например, резьбовых соединений. Повышение уровня автоматизации способствует внедрению безлюдного производства. Разрабатываемая методика может быть применена не только в автоматизированных производствах, но также адаптирована для неавтоматизированных станков с ЧПУ за счет применения стандартного станочного оборудования.

Литература

1. Темпель Ю.А., Темпель О.А. Особенности и факторы, оказывающие влияние на точность механической обработки на станках с ЧПУ // Новые технологии-нефтегазовому региону. – 2016. – С. 265–270.
2. Калмыков В.В., Юсупова К.Н. Сравнительный анализ методов активного контроля состояния режущего инструмента // Научно-технологические инновации в приборостроении и развитии инновационной деятельности в вузе. – 2013. – С. 44–51.
3. Mohanraj T. et al. Tool condition monitoring techniques in milling process—A review // Journal of Materials Research and Technology. – 2020. – Vol.9. – №1. – С. 1032–1042.
4. Сырицкий А.Б., Болдасов Д.Д. Фазохронометрическая система мониторинга износа режущего инструмента // Металлообработка. – 2015. – №5(89). С. 2–10.
5. Junior M.V. et al. The role of tool presetting in milling stability uncertainty // Procedia Manufacturing. – 2018. – Т.26. – Рр. 164-172.
6. Басова Т.В. Разработка методики активного контроля резьбовых фрез на станке с ЧПУ // Альманах научных работ молодых ученых Университета ИТМО. – 2021. – Т.2. – С. 45–47.

УДК 681.7.069.24

РАЗРАБОТКА СИСТЕМЫ ТЕРМОСТАБИЛИЗАЦИИ ДЛЯ УСТАНОВКИ ПО КОНТРОЛЮ ПАРАМЕТРОВ ГАЗОВЫХ ЯЧЕЕК КВАНТОВЫХ ДАТЧИКОВ

Власова М.А.^{1,2} (студент), Шевченко А.Н.², Дементьев М.М.²

Научный руководитель – канд. техн. наук Безмен Г.В.²

¹*Университет ИТМО*

²*АО «Концерн «ЦНИИ «Электроприбор»*

e-mail: mary_vlasova446@mail.ru, standw@yandex.ru, mysha2009@yandex.ru, gbezmen@eprib.ru

Работа выполнена в рамках темы НИР №620164 «Методы искусственного интеллекта для киберфизических систем».

Разработана система термостабилизации для установки по контролю параметров газовых ячеек квантовых датчиков. Представлено решение задачи оптимального расположения датчика температуры в установке. Произведён синтез закона управления температурой вблизи ячейки. С его использованием проведено экспериментальное определение параметров четырёх газовых ячеек по ширине линии оптического резонанса, оценено качество их заполнения.

Ключевые слова: квантовый датчик, газовая ячейка, термостабилизация, оптический резонанс, математическая модель, температура.

Существуют приборы, принцип работы которых основан на взаимодействии света и рабочего вещества, заключённого в объёме прозрачной ячейки [1, 2]. К таким устройствам относятся квантовые магнитометры, гироскопы, стандарты частоты. Для проверки качества ячеек квантовых приборов используются установки для контроля их параметров, а именно прозрачности, поляризационных свойств и ширины оптического резонанса (добротности).

Ширина оптического резонанса определяется материалом стенок ячейки, давлением буферного газа, а также наличием загрязнений внутри объёма ячейки [3]. Для проверки на наличие нарушений в технологическом процессе заполнения ячеек необходимо по ширине линии оптического резонанса определить давление буферного газа и количество загрязнений в ячейке. Если давление отличается от номинального, ширина линии резонанса будет либо слишком широкой, либо слишком узкой. Загрязнения только увеличивают линию оптического резонанса. Кроме ширины оптического резонанса, его важным параметром является амплитуда. Она определяется температурой рабочего вещества газовой ячейки. При низкой температуре отношение сигнал/шум низкое, и линия резонанса слабо различима на фоне шумов. При высокой температуре происходит искажение линии резонанса из-за поглощения щелочным металлом всего света от детектирования. Таким образом, существует оптимальная температура, при которой и производится определение ширины линии резонанса без искажения его формы. В предлагаемой установке температура подобрана экспериментально с учетом имеющегося лазерного излучателя - использован вертикально-излучающий лазер для стандарта частоты с мощностью оптического излучения 0,2 мВт. Для такой мощности оптимальной оказалась температура 72°C.

Таким образом, целью работы является разработка системы стабилизации температуры ячейки для установки по контролю её параметров. Такая система должна обеспечивать

постоянство амплитуды оптического резонанса, по ширине которого можно определять давление буферного газа, а также наличие загрязнений в ячейке.

Состав установки. Установка для контроля параметров газовых ячеек устроена следующим образом. Через ячейку, находящуюся в термостабилизированном контуре, проходит свет от лазерного излучателя. Фотодетектор регистрирует интенсивность прошедшего через ячейку света. Нагрев ячейки производится нагнетаемым вентилятором воздухом, проходящим через нагревательный элемент. При разработке контуров термостабилизации необходимо решить задачу оптимального расположения датчика температуры [4]. В нашей работе рассмотрено два варианта.

Первый вариант системы управления изображён на рисунке 1. В нём термодатчик размещается непосредственно у нагревательного элемента. Преимуществами такой системы являются простота настройки (используется релейный закон управления) и высокая динамика (то есть система почти мгновенно реагирует на резкие скачки температуры забираемого вентилятором воздуха). Недостатками такой системы являются требование установки датчика с большим температурным диапазоном (это обусловлено достаточно высокими температурами у нагревательного элемента), а также сильная зависимость установившейся температуры у ячейки от температуры окружающей среды.

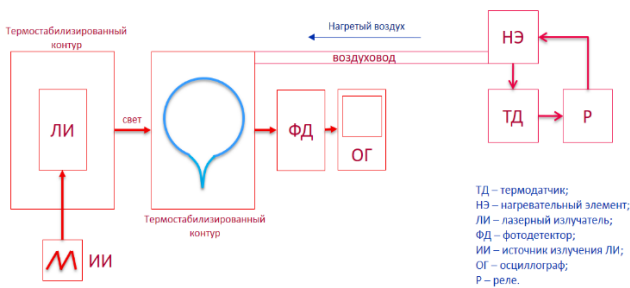


Рис. 1. Размещение термодатчика непосредственно у нагревательного элемента

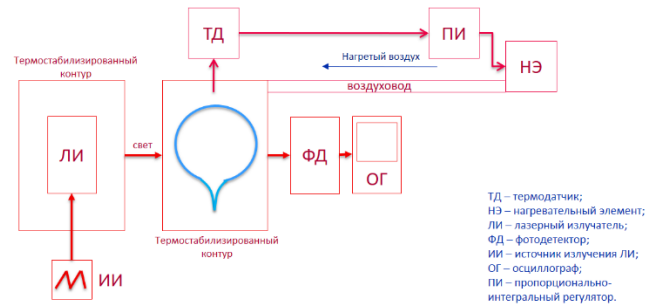


Рис. 2. Размещение термодатчика около ячейки

Второй вариант системы термостабилизации представлен на рисунке 2. В нём термодатчик размещается около ячейки. Преимущество такой системы – поддержка постоянной температуры непосредственно около ячейки. Недостатки – низкая динамика системы, а также сложный закон управления (требуется подбор коэффициентов закона регулирования, ввиду большого расстояния между нагревателем и ячейкой).

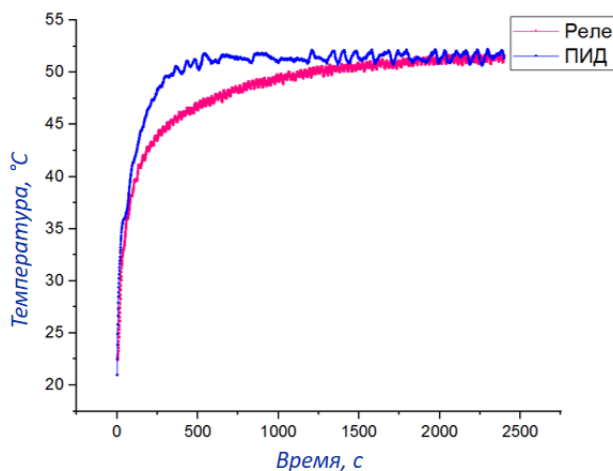


Рис. 3. Графики выхода на температурный режим

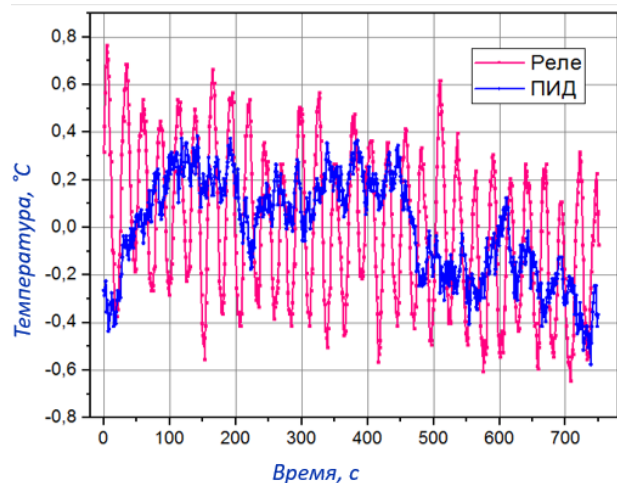


Рис. 4. Графики ошибок схем термостабилизации

Коэффициенты пропорционально-интегрального регулятора (ПИ-регулятора) для такой схемы подбирались посредством ПО Matlab Simulink. Для этого были записаны входные и выходные температуры воздухопровода. Далее, по этим данным была построена передаточная

функция и, с помощью тюнера пропорционально-интегрального регулятора, были подобраны коэффициенты такие, что система обладает устойчивостью и наименьшей постоянной времени.

Выбор системы термостабилизации. На рисунке 3 представлено сравнение релейной схемы и схемы с ПИ-регулятором при выходе на температурный режим. По графикам можно сделать вывод, что система термостабилизации с ПИ-регулятором имеет меньшее время выхода на режим, а значит, позволяет быстрее приступить к проведению измерений

На рисунке 4 представлены графики ошибки стабилизации температуры предложенных схем. По ним можно сделать вывод, что система термостабилизации с ПИ-регулятором имеет не худшие шумовые характеристики в сравнении с релейной схемой.

Получение линии резонанса. Таким образом, для получения оптимальной линии оптического резонанса была выбрана система термостабилизации с использованием ПИ-регулятора. По ширине этой линии были определены параметры четырёх газовых ячеек, произведена оценка добротности контура и давления буферного газа. Для эксперимента были использованы сферические ячейки с диаметром 5 мм, содержащие 6 мг цезия и 100 торр азота. По результатам проверки было определено, что все 4 экспериментальные ячейки удовлетворяют требованиям отсутствия примесей. В ячейке с максимальной добротностью минимальное давление буферного газа обусловлено его вытеканием при отпайке ячейки от установки. Ячеек с повышенным значением азота, по которому можно идентифицировать примеси и загрязнения, обнаружено не было.

С помощью аппроксимации Лоренцевой кривой, представленной на рисунке 5, удалось рассчитать ширину резонанса $\Gamma_{рез.}$, а далее вычислить расчетное парциальное давление азота $P_{рез.}$ и добротность контура $Q_{рез.}$ для каждой из ячеек.

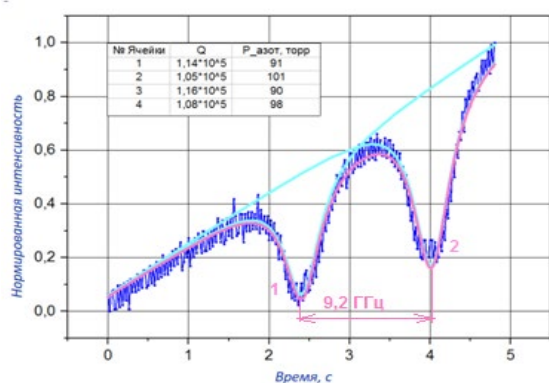


Рис. 5. Аппроксимация Лоренцевой кривой

Добротность контура $Q_{рез.}$ вычислялась по следующей формуле:

$$Q_{рез.} = \frac{\omega_{рез.}}{\Gamma_{рез.}},$$

где $\omega_{рез.}$ — резонансная частота цезия на линии D1.

В результате были обнаружены небольшие расхождения с заявленными 100 торр внутри ячейки, но скорее всего это было вызвано наличием дополнительных примесей в газовом составе ячеек или некачественной герметизацией ячеек.

В рамках работы определена оптимальная температура для измерения параметров газовой ячейки для получения наилучшей формы оптического резонанса. Реализована система стабилизации температуры установки для контроля параметров газовой ячейки, проведено сравнение двух вариантов закона управления данной системой, определён оптимальный закон. Проведено экспериментальное определение параметров четырёх ячеек по ширине линии оптического резонанса на установке с использованием стабилизации температуры с помощью ПИ-регулятора, оценено качество заполнения экспериментальных ячеек.

Литература

1. Peshekhonov V.G. Gyroscopic navigation systems: Current status and prospects. Gyroscopy and navigation 2(3). – 2011. – Pp. 111-118. DOI:10.1134/S2075108711030096.

2. Вершовский А.К., Шевченко А.Н. Ядерный магнитный гироскоп: принцип действия, история, перспективы // Материалы XVII конференции молодых ученых. Навигация и управление движением. – 2015. – С. 19–28.
3. Шевченко А.Н., Захарова Е.А. Исследование зависимости качества резонанса от распределения цезия в ячейке квантового датчика вращения // Научно-технический вестник информационных технологий, механики и оптики. – 2019. – Т.19. – №4. DOI: 10.17586/2226-1494-2019-19-4-567-573.
4. Рубцов В.П., Горячих Е.В., Щербаков А.В. Исследование влияния неоднородности электрической печи сопротивления как объекта управления // Электротехника. – 2015. – №7. – С. 41а-45.
5. Greg A. Pitz, Douglas E. Wertepny and Glen P. Perram - Pressure broadening and shift of the cesium D1 transition by the noble gases and N₂, H₂, HD, D₂, CH₄, C₂H₆, CF₄, and 3He. Phys. Rev A 80, 062718 (2009). DOI:10.1103/PHYSREVA.80.062718.

УДК 004.056.55

КРИПТОГРАФИЧЕСКИЙ МЕТОД ОБЪЕДИНЕНИЯ СООБЩЕНИЙ НЕСКОЛЬКИХ ПОЛЬЗОВАТЕЛЕЙ В ОДНОМ ШИФРТЕКСТЕ

Голованов А.А.¹ (студент)Научный руководитель – д. т. н., профессор Беззатеев С.В.¹¹Университет ИТМО

e-mail: agolovanov2403@gmail.com

Работа выполнена в рамках темы НИР №620164 «Методы искусственного интеллекта для киберфизических систем».

Работа частично финансировалась Федеральной программой академического лидерства Приоритет 2030.

Работа посвящена задаче проектирования криптографической схемы, позволяющей в одном шифртексте объединить несколько открытых текстов. В такой схеме из одного шифртекста при дешифровании можно получить различные открытые тексты в зависимости от используемого ключа. В ходе работы созданы два прототипа описанной криптографической схемы: на основе китайской теоремы об остатках и на основе ортогональных линейных кодов. Проведён предварительный анализ их особенностей.

Ключевые слова: криптографическая схема, шифрование, теория чисел, теория помехоустойчивых кодов.

Стоит задача по разработке криптографической схемы, которая позволяет объединить зашифровать несколько открытых текстов в одном шифртексте. Так, в зависимости от используемого ключа из одного и того же шифртекста операция дешифрования может получить разные открытые тексты. Подобная схема может использоваться в таких системах, как аутентификация типа «вход по принуждению», организация доступа в секретные архивы, создание карт лояльности нескольких компаний.

Представим общий вид криптографической схемы в формальном виде. Схема определена над пространствами:

K – пространство ключей;

M – пространство открытых текстов;

C – пространство шифртекстов.

Криптографическая схема состоит из трёх функций:

– функция генерации ключей:

$$KEYGEN : N \rightarrow K^*; n \mapsto (k_1, \dots, k_n), \quad (1)$$

где N – множество натуральных чисел,

n – количество различных ключей криптографической схемы,

– функция шифрования:

$$ENC : M^n \times K^n \rightarrow C; (m_1, \dots, m_n, k_1, \dots, k_n) \mapsto s, \quad (2)$$

– функция дешифрования:

$$DEC : C \times K \rightarrow M; (s, k_i) \mapsto m_i. \quad (3)$$

Описанная схема концептуально напоминает схему разделения секрета, но в некотором смысле является противоположностью (рис. 1). В схеме разделения секрета некоторая часть информации (секрет) в ходе процедуры разделения преобразуется в новые n частей информации (доли секрета), чтобы затем по ним можно было восстановить изначальную часть в ходе процедуры восстановления секрета.

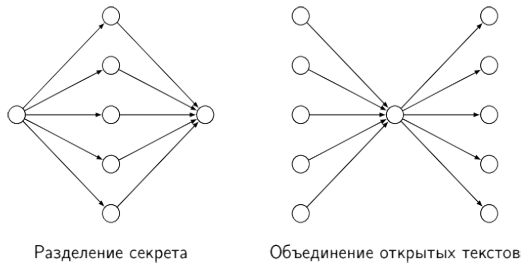


Рис. 1. Сравнение схемы разделения секрета и схемы объединения открытых текстов в одном шифртексте

Схема на основе китайской теоремы об остатках

Первая криптографическая схема получена в ходе переосмысления схемы разделения секрета М. Миньотта, основанной на китайской теореме об остатках [1]. Китайская теорема об остатках при заданных попарно взаимно простых натуральных числах $\{k_1, k_2, \dots, k_n\}$ устанавливает биекцию $s \mapsto (m_1, m_2, \dots, m_n)$, такую что:

$$\{m_1 = s \bmod k_1 \ m_2 = s \bmod k_2 \ \dots \ m_n = s \bmod k_n \Leftrightarrow s = (\sum_{i=1}^n m_i \cdot y_i \cdot z_i) \bmod z, \quad (4)$$

где $z = \prod_{i=1}^n k_i$, $z_i = z/k_i$, $y_i = z_i^{-1} \pmod{k_i}$, $0 \leq m_i < k_i$, $0 \leq s < z$ – целые числа.

В схеме разделения секрета Миньотта секретом выступает s , а долями секрета – m_i , а также наложено дополнительное ограничение на ключи k_i . В схеме объединения открытых текстов в одном шифртексте, разработанной в ходе данной работы, ограничений на ключи k_i не накладываемся, s является шифртекстом для открытых текстов m_i .

Таким образом, пространства, на которых определена схема:

$$K = N, M = N, C = \{0 \leq s < z\}, \text{ где } Z \text{ – множество целых чисел.}$$

Функции криптографической схемы, аналогично (1), (2), (3) определены следующим образом:

$$KEYGEN(n) = \{i \neq j \Rightarrow (k_i, k_j) = 1\}, \quad (5)$$

$$s = ENC(m_1, \dots, m_n, k_1, \dots, k_n) = (\sum_{i=1}^n m_i \cdot z_i \cdot y_i) \bmod z, \quad (6)$$

$$m_i = DEC(s, k_i) = s \bmod k_i. \quad (7)$$

Также предложено расширение данной схемы – для добавления вероятностной составляющей в функции шифрования (6) вводится дополнительный открытый текст, представляющий из себя случайное число:

$$s = ENC(m_1, \dots, m_n, m_{n+1}, k_1, \dots, k_n, k_{n+1}), \quad (8)$$

где m_{n+1} – случайное число. Для этого необходим дополнительный ключ k_{n+1} .

Особенности схемы на основе китайской теоремы об остатках

Пусть для генерации ключей условились выбирать только простые числа. Имеет место теорема о распределении простых чисел [2]:

$$\pi(N) \sim N / \ln \ln N, \quad (9)$$

где $\pi(N)$ – количество простых чисел, не превосходящих N . Эта теорема, по аналогии с RSA, говорит о том, что нахождение простых чисел для использования их в качестве ключей – простая задача, в то время как подбор ключей – сложная задача.

Также можно заметить, что знание одного из ключей в $\{k_1, \dots, k_n\}$ не даёт дополнительной информации об остальных ключах. Планируется уточнение сложности взлома данной схемы в следующей работе.

Размер шифртекста в данной криптографической схеме равен

$$|s| = n \cdot |k|, \quad (10)$$

где $|s|$ – размер числа s в битах,
 $|k|$ – средний размер ключа.

Вариант криптографической схемы, в которой шифрование проводится по формуле (6), является семантически не стойким: набор равных открытых текстов каждый раз будет давать равные шифртексты при одинаковых ключах:

$$(m_1, \dots, m_n) = (m'_1, \dots, m'_n) \Rightarrow s = s'.$$

Однако, если используется формула (8), то семантическая стойкость обеспечивается.

Схема на основе ортогональных линейных кодов

Вторая схема, разработанная в ходе данной работы, основана на свойствах ортогональных линейных кодов.

Пусть $\{G_1, G_2, \dots, G_n\}$ – порождающие матрицы n линейных кодов, попарно ортогональных:

$$i \neq j \Rightarrow G_i \cdot G_j^T = O, \quad (11)$$

где O – нулевая матрица. Для таких матриц проявляется следующее свойство. Пусть матрица D_i определена как:

$$D_i = G_i^T \cdot (G_i G_i^T)^{-1}. \quad (12)$$

Тогда, благодаря свойству (11), умножение на порождающую матрицу с таким же индексом даст единичную матрицу I :

$$G_j D_j = G_j G_j^T \cdot (G_j G_j^T)^{-1} = I, \quad (13)$$

а умножение на порождающую матрицу с отличающимся индексом даст нулевую матрицу:

$$i \neq j \Rightarrow G_i D_j = G_i G_j^T \cdot (G_j G_j^T)^{-1} = O \cdot (G_j G_j^T)^{-1} = O. \quad (14)$$

Это свойство использовано для создания криптографической схемы. Зададим схему на следующих пространствах:

$$M = F^k, C = F^{2k}, K = F^{k \times 2k},$$

где F – некоторое поле,

k – размерность используемых кодов.

Таким образом, порождающие матрицы данных ортогональных кодов являются ключами криптографической схемы ($k_i = G_i$), а информационные слова соответствующих кодов – открытыми текстами. Шифртекст является суммой кодовых слов, полученных из сообщений:

$$s = ENC(m_1, \dots, m_n, G_1, \dots, G_n) = \sum_{i=1}^n c_i, \text{ где } c_i = m_i G_i.$$

Функция дешифрования определена следующим образом:

$$m_j = DEC(s, G_j) = s \cdot D_j.$$

Доказательство корректности, используя свойства (11) и (12):

$$s \cdot D_j = m_j \cdot G_j G_j^T \cdot (G_j G_j^T)^{-1} + \sum_{i=1, i \neq j}^n m_i \cdot G_i G_j^T \cdot (G_j G_j^T)^{-1} = m_j \cdot I + \sum_{i=1, i \neq j}^n m_i \cdot O = m_j.$$

Обратите внимание, что матрица D_j однозначно определяется из G_j по формуле (12).

Особенности схемы на основе ортогональных линейных кодов

Сложность подбора ключа ещё не исследовано для данной схемы. Также не исследовано можно ли, зная один из ключей, получить дополнительную информацию о других ключах.

В отличие от предыдущей схемы (10), размер шифртекста в данной схеме не зависит от n :

$$|s| = 2 \cdot |m|. \quad (15)$$

В дальнейшей работе планируется анализ классов кодов и поиск подходящих классов для данной схемы. В частности, интересуют такие особенности классов кодов:

- высокая сложность взлома схемы;
- низкая сложность алгоритма *KEYGEN*;
- обратимость матрицы $G_j G_j^T$.

Исследована задача проектирования криптографической схемы, позволяющей объединить несколько открытых текстов в одном шифртексте. Спроектированы два варианта такой схемы, доказана их корректность. В дальнейшей работе планируется исследование безопасности данных схем.

Литература

1. Mignotte M. How to Share a Secret? // EUROCRYPT '82, LNCS 149. – Heidelberg: Springer-Verlag Berlin. – 1983. - С. 371-375.
2. Crandall R., Pomerance C. How primes are distributed // Prime Numbers. A Computational Perspective. – New York: Springer Science+Business Media. – 2000. – С. 10-14.

УДК 681.51

ОБЗОР МЕТОДОВ ПОСТРОЕНИЯ ТЕМПЕРАТУРНОЙ МОДЕЛИ ВОЛОКОННО-ОПТИЧЕСКОГО ГИРОСКОПА

Гонтарь Д.А.¹ (студент), Егоров Д.А.²

Научный руководитель – канд. техн. наук, доцент Драницына Е.В.^{1,2}

¹Университет ИТМО

²АО ЦНИИ «Электроприбор»

e-mail: denisgontar99@gmail.com

Работа выполнена в рамках темы НИР №620164 «Методы искусственного интеллекта для киберфизических систем».

В данной работе произведен сравнительный обзор методов машинного обучения с целью выбора наиболее эффективного для компенсации температурного дрейфа волоконно-оптического гироскопа.

Ключевые слова: машинное обучение, обзор, температурная компенсация, волоконно-оптический гироскоп, алгоритмические методы.

Одной из основных проблем при создании высокоточных волоконно-оптических гироскопов (ВОГ) является чувствительность выходного сигнала к температуре. Для создания ВОГ навигационного класса точности необходимо точно компенсировать влияние температуры на выходной сигнал ВОГ [1].

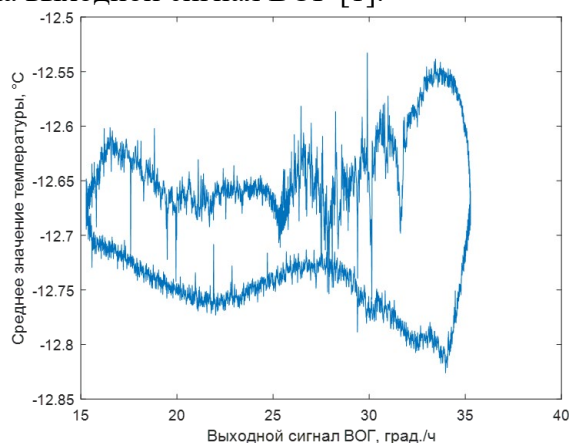


Рис. 1. График зависимости выходного сигнала ВОГ от среднего значения температуры с 8 датчиков

Температура и ее изменение оказывают влияние на такие компоненты ВОГ, как источник света, многофункциональная интегральная оптическая схема, волоконно-оптическая катушка, плата обработки.

Проблема влияния изменения температуры на выходной сигнал ВОГ, помимо различных конструктивных решений и применения термостатирования может быть решена введением алгоритмической температурной компенсации выходного сигнала [2].

На рисунке 1 показана зависимость выходного сигнала ВОГ от температуры, эту зависимость и стремится убрать алгоритмическая компенсация.

Принципы работы систем машинного обучения. Алгоритмические методы включают в себя два способа. Первый – аналитический – подразумевает анализ каждого элемента прибора при внешнем температурном воздействии. Данный процесс является затруднительным ввиду специфики самих элементов гироскопа, требует много ресурсов и времени и усложняется тем, что все компоненты имеют индивидуальные особенности. Второй способ – регрессионный анализ – предполагает получение многомерной функции зависимости выходного сигнала ВОГ от температуры окружающей среды. Такую зависимость получают по реальным данным при проведении климатических испытаний на специальных стендах с термокамерами.

Наиболее распространённые методы регрессионного анализа это:

- Статистические методы: линейная регрессия, полиномиальная регрессия, ридж регрессия, регрессия лассо [3].
- Методы искусственного интеллекта: искусственные нейронные сети, экспертные системы, байесовские сети, генетические алгоритмы, метод роя частиц.

Нейронные сети. Нейронная сеть — это последовательность нейронов, соединенных между собой синапсами. Структура нейронной сети имеет свое начало в биологии, нейросеть, это интерпретация мозга человека, в котором находятся нейроны, передающие информацию с помощью электрических импульсов. Благодаря такой структуре машина обретает способность анализировать, запоминать и воспроизводить из памяти различную информацию [4, 5].

Нейронные сети используются для решения сложных задач, которые требуют аналитических вычислений подобных тем, что делает человеческий мозг.

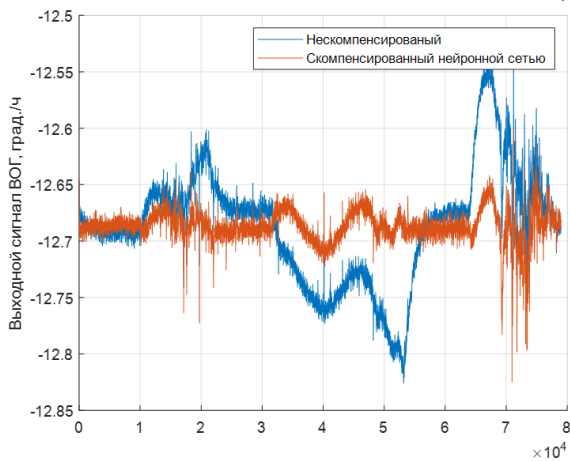


Рис. 2. Скомпенсированный и нескомпенсированный сигнал ВОГ

Байесовские сети. Байесовская сеть — модель, описывающая множество переменных и их вероятностных взаимодействий. Байесовские сети часто применяются для диагностирования заболевания на основе симптомов, диагностирование поломки в технической системе на основе ее проявлений. Пример подобной системы представлен на рисунке 3. В нашем случае могут быть применены для выбора определенного компенсационного алгоритма в зависимости от входных данных/воздействий.

Генетические алгоритмы. Генетический алгоритм — это алгоритм оптимизации, представляющий собой упрощенную программную интерпретацию эволюционных изменений популяции [5].

Генетические алгоритмы представляют собой, скорее, подход, чем единые алгоритмы. Они требуют содержательного наполнения для решения каждой конкретной задачи.

В общем случае алгоритм можно описать так: оптимизируемые параметры представляются в виде векторов, называемых хромосомами, случайным образом создается начальная популяция, начальная популяция проверяется на приспособленность с помощью специальной функции приспособленности, исходя из которой каждой особи в популяции

Самыми распространенными применениями нейронных сетей является: классификация, предсказание и распознавание.

В ходе своей выпускной квалификационной работы я производил компенсацию температурной зависимости выходного сигнала ВОГа с помощью нейронных сетей. Результат моей работы представлен на экране. Теперь я нахожусь в поисках методов, которые могут помочь улучшить полученную регрессионную модель.

Но как вы можете заметить по форме скомпенсированного сигнала, все еще существует место для улучшения модели. Результаты такой компенсации представлены на рисунке 2.

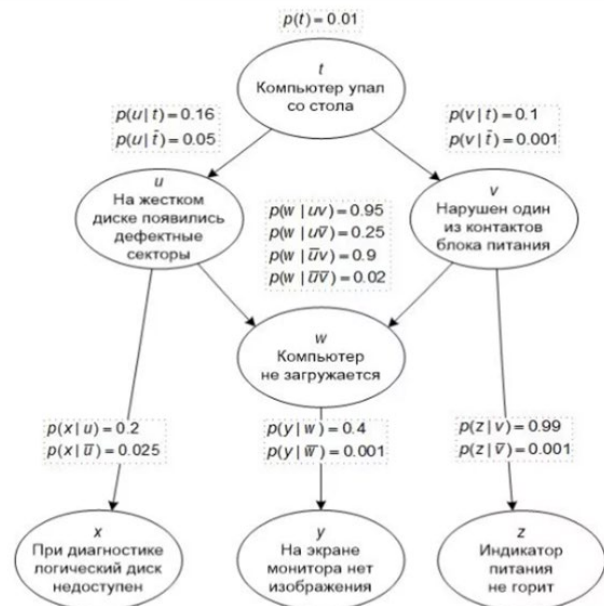


Рис. 3. Пример байесовской сети

присваивается вероятность выживания. Далее определяется набор выживших особей, все выжившие особи допускаются к скрещиванию и/или мутации. После скрещивания появляется новая популяция. Алгоритм повторяется до тех пор, пока не будет удовлетворён критерий останова алгоритма.

В нашем случае могут быть применены для оптимизации и нахождения параметров выбранной модели, для выбора модели компенсации температурной зависимости выходного сигнала ВОГ в зависимости от входных воздействий.

Например: группа авторов из Юго-Восточного университета в Нантине уже применили генетический алгоритм для нахождения весовых коэффициентов и пороговых значений нейронной сети.

Метод роя частиц. Метод роя частиц — это оптимизационный алгоритм, поддерживающий множество решений, называемых популяцией или множеством частиц и перемещающий их в пространстве решений согласно простой формуле.

Текущее состояние отдельной частицы определяются их координатами, направлением и скоростью движения в пространстве решений. Эти параметры выбираются случайным образом на этапе инициализации. Каждая частица хранит свое текущее состояние, лучшее из найденных ей решений и лучшее из найденных всеми частицами решений. Является одним из примеров мультиагентных систем.

Может быть применён для оптимизации параметров нейронной сети или другого регрессионного алгоритма.

Вывод

Анализ литературы показывает, что нейронные сети, используемые вместе с другими алгоритмами искусственного интеллекта, являются более эффективным методом алгоритмической компенсации, чем сети, не использующие оптимизирующих алгоритмов.

Литература

1. Драницына Е.В. Исследование зависимости выходного сигнала волоконно-оптического гироскопа от температуры в составе бескарданного инерциального измерительного модуля // XIV конференция молодых ученых «Навигация и управление движением». – СПб.: ОАО «Концерн «ЦНИИ «Электроприбор». – 2012. – С. 447–452.
2. Драницына Е.В., Егоров Д.А., Унтилов А.А., Дейнека Г.Б., Шарков И.А., Дейнека И.Г. Снижение влияния изменения температуры на выходной. – 2012. – №4(79). – С. 10–20.
3. Дрейпер Н., Смит Г. Прикладной регрессионный анализ. – 2017. – С. 273–340.
4. Гудфеллоу Я., Бенджио И., Курвилль А. Глубокое обучение / пер. с англ. А. А. Слинкина. – 2-е изд., испр. – М.: ДМК Пресс. – 2018. – 652 с.
5. Курейчик В.М., Курейчик В.В., Гладков Л.А. Генетические алгоритмы. – 2017. – 15 с.

УДК 004.032.26

АНАЛИЗ ЭФФЕКТИВНОСТИ ПРИМЕНЕНИЯ НЕЙРОННЫХ СЕТЕЙ В ЗАДАЧАХ КЛАССИФИКАЦИИ ОБЪЕКТОВ КИБЕРАТАКИ НА ОСНОВЕ ОТКРЫТЫХ ИСТОЧНИКОВ ДАННЫХ

Григорьев А.М.¹ (студент)

Научный руководитель – канд. техн. наук Менщиков А.А.¹

¹*Университет ИТМО*

e-mail: figaroalex@gmail.com, menshikov@itmo.ru

В данной работе рассматривается анализ эффективности применения нейронных сетей в задачах классификации объектов кибератаки на основе открытых источников данных. Прогнозирование типа целей при кибератаках может быть очень важным при разработке обширной структуры для борьбы с будущими кибератаками. Различные секторы, основанные на характере их бизнеса, могут быть конкретной целью для некоторых видов кибератак. Различные компании должны учиться на прошлых ошибках и выявлять потенциальные риски и уязвимости, чтобы быть готовыми к будущим кибернарушениям.

Ключевые слова: тип цели, кибератака, классификация, кибернарушение, разведка по открытым источникам.

Киберпреступники ежедневно нападают на различные предприятия. В зависимости от их мотивации и намерений, различные секторы находятся под угрозой кибератак, и в настоящее время все предприятия, от правительственных до розничной торговли, пытаются принять различные стратегии для решения этой новой растущей проблемы. Эти контрмеры и стратегии борьбы с ними могут быть различными в каждом секторе на основе чувствительности характера их работы, будь то предоставление различных услуг для своих клиентов или хранение конфиденциальных и критически важных данных.

Эксперты по кибербезопасности должны определять приоритеты на основе характера целей и чувствительности данных. Одним из способов решения этой проблемы является использование методов интеллектуального анализа данных, которые используются в различных областях, включая кибербезопасность. Это исследование направлено на изучение использования нейронной сети в задачах классификации для прогнозирования типа целей при кибератаках, которые могут помочь экспертам по кибербезопасности идентифицировать и определить приоритеты требований и возможных уязвимостей в любом типе бизнеса.

В качестве исходных данных использовалась база данных Hackmageddon [1]. Пополняемая одним энтузиастом Paolo Passeri, она была выбрана за счет преимуществ в полноте покрытия и непредвзятости открытых данных об кибератаках и в достаточном для классификации категорировании.

Рассматривались следующие категории: автор кибератаки (кто ее совершил), способ кибератаки (малварь, фишинг и пр.), мотивация кибератак (киберпреступление, кибервойна, кибершпионаж, хактивизм) и страна цели. Классификации подвергалась категория типа кибератаки, т.е. ее отрасль, будь это здравоохранение, финансовые и оборонные структуры, или отдельные индивиды. Исходными данными для тренировочных данных выступили атаки за весь 2021 год, для тестовой выборки – первая половина января 2022 года.

Для написания метода классификации, нейронной сети, был использован язык программирования Python и библиотека sklearn. Классификатором был выбран мультислойный

перцептрон, а его оптимальные параметры – функция активации, кол-во слоев и нейронов в них, темп обучения – были подобраны через grid search.

Были получены следующие результаты точности (рис. 1) и матрицы ошибок (рис. 2).

	precision	recall	f1-score	support
1	0.00	0.00	0.00	2
2	0.00	0.00	0.00	3
3	0.00	0.00	0.00	5
7	0.00	0.00	0.00	4
8	0.00	0.00	0.00	2
9	0.31	0.83	0.45	12
10	0.11	0.33	0.17	3
11	0.00	0.00	0.00	3
12	0.00	0.00	0.00	1
13	0.54	0.68	0.60	19
14	0.00	0.00	0.00	4
15	0.00	0.00	0.00	7
16	0.45	0.42	0.43	12
17	0.00	0.00	0.00	1
18	0.00	0.00	0.00	1
19	0.25	1.00	0.40	1
21	0.00	0.00	0.00	3
accuracy			0.36	83
macro avg	0.10	0.19	0.12	83
weighted avg	0.24	0.36	0.28	83

Рис. 1. Параметры точности модели

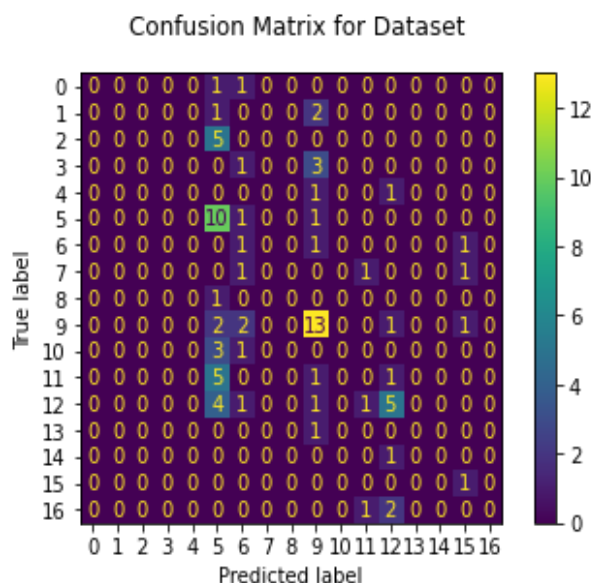


Рис. 2. Матрица ошибок

Это исследование рассмотрело использование нейронную сеть как метод классификации для обучения моделей для прогнозирования типа целей при кибератаках. Показатель точности 36% свидетельствует, что нейронная сеть может использоваться в сочетании с другими подходами экспертов по кибербезопасности для прогнозирования типа целей. Одним из решений для повышения точности прогностической модели является использование более точного и полного набора данных, поскольку это исследование рассматривало только открытые источники данных из-за отсутствия доступа к секретной информации. Наряду с более полным и точным набором данных, будущее рассмотрение проблемы может быть выполнено в направлении анализа временных рядов, потому что время также может влиять на тип прогнозирования цели. Некоторые типы целей становятся более благоприятными по сравнению с другими в отношении времени. Применяя анализ временных рядов, прогнозирование типа целей может быть переменным и динамическим в соответствии с датой.

Это исследование подчеркнуло важность и значимость использования методов классификации для анализа прошлых кибератак для прогнозирования типа целей. Прогнозирование типа целей дает экспертам по кибербезопасности возможность разработать киберстратегию, основанную на различных секторах в отношении прошлого опыта.

Литература

1. Passeri P. HACKMAGEDDON. [Электронный ресурс]. – Режим доступа: <https://www.hackmageddon.com/> (Дата обращения: 20.01.2022).
2. Орельен Ж. Прикладное машинное обучение с помощью Scikit-Learn и TensorFlow. Концепции, инструменты и техники для создания интеллектуальных систем / Ж. Орельен. — 1-е изд. — СПб: ООО "Альфа-книга". – 2018. – 688 с.

УДК 620.179.119

АВТОМАТИЗАЦИЯ ТЕХНОЛОГИИ КОМПЛЕКСНОГО НЕРАЗРУШАЮЩЕГО КОНТРОЛЯ КАЧЕСТВА ПОКРЫТИЙ ИЗДЕЛИЙ СЛОЖНОЙ ГЕОМЕТРИЧЕСКОЙ ФОРМЫ

Губин М.С.¹ (аспирант)

Научный руководитель – д-р техн. наук Федоров А.В.¹

¹Университет ИТМО

e-mail: gubin.maxim@mail.ru

Работа затрагивает вопросы, направленные на автоматизацию комплексного неразрушающего контроля качества покрытий изделий сложной геометрической формы, с целью повышения достоверности контроля. Проанализировано современное состояние технологий неразрушающего контроля качества никель-хромового покрытия. Обоснована необходимость автоматизации комплексного неразрушающего контроля (НК).

Ключевые слова: автоматизация неразрушающего контроля, контроль толщины никель-хрома, контроль дефектов в никель-хромовом покрытии, современное состояние технологий неразрушающего контроля, комплексное применение методов НК.

На данный момент отрасли, связанные с изготовлением высокотехнологичных агрегатов, нацелены на развитие в нашей стране. Особое внимание направлено на стороны, связанные с повышением безотказности, долговечности сложных технических систем. Наиболее важной частью высокотехнологичного агрегата является сложная техническая система, направленная на создание силы тяги, создание которой требует колоссальных временных затрат и ресурсов на разработку и производство. Поэтому решение задач, связанных с повышением долговечности и безотказности, актуально и своевременно на данный момент.

Производство данных технических систем осуществляется в соответствии с предъявляемыми техническими требованиями. Основным этапом в технологической цепочке является нанесение металлического и неметаллического покрытия, основная задача которого заключается в повышении термостойкости внутренних огневых стенок [1]. На безотказность и долговечность влияют такие параметры, как толщина покрытия, поверхностные и подповерхностные дефекты огневой стенки, представляющую изделие сложной геометрической формы. Отклонение параметров от допустимых значений может приводить к преждевременному выходу из строя, что является недопустимым при эксплуатации. Поэтому контроль основных параметров является необходимой задачей. На данный момент существует ряд технологий неразрушающего контроля основных параметров никель-хромового покрытия, которые находят применения на предприятиях, изготавливающих изделия сложной геометрической формы. Например, используются технологии ручного неразрушающего контроля с помощью магнитных толщиномеров толстослойного никелевого покрытия МТП-1М и хромового покрытия на никелевом подслое МТП-ХН1. Реализуемые в данных приборах портативный и магнитодинамический методы контроля позволили обеспечить их портативность, возможность контроля покрытий на криволинейных поверхностях и в труднодоступных местах изделий сложной геометрической формы. Измерения производятся, согласно инструкционной карте контроля, представленной на рисунке 1.

Согласно инструкционной карте контроля сложно-профильного изделия можно определить, что измерения производятся в четырех сечениях в 19-ти точках с шагом от 5 до 15 мм. Следовательно, возникает проблема нахождения толщины никелевого покрытия в допустимом диапазоне по всей поверхности сложно-профильного изделия. Для решения данной проблемы на предприятиях внедряются автоматизированные комплексы НК. Но из-за существенного

разнообразия конструктивных исполнений изделий сложной формы, автоматизированный контроль не проводится для всех существующих конструкций. Ниже приведены примеры автоматизированных установок в части контроля основных параметров никель-хромового покрытия.

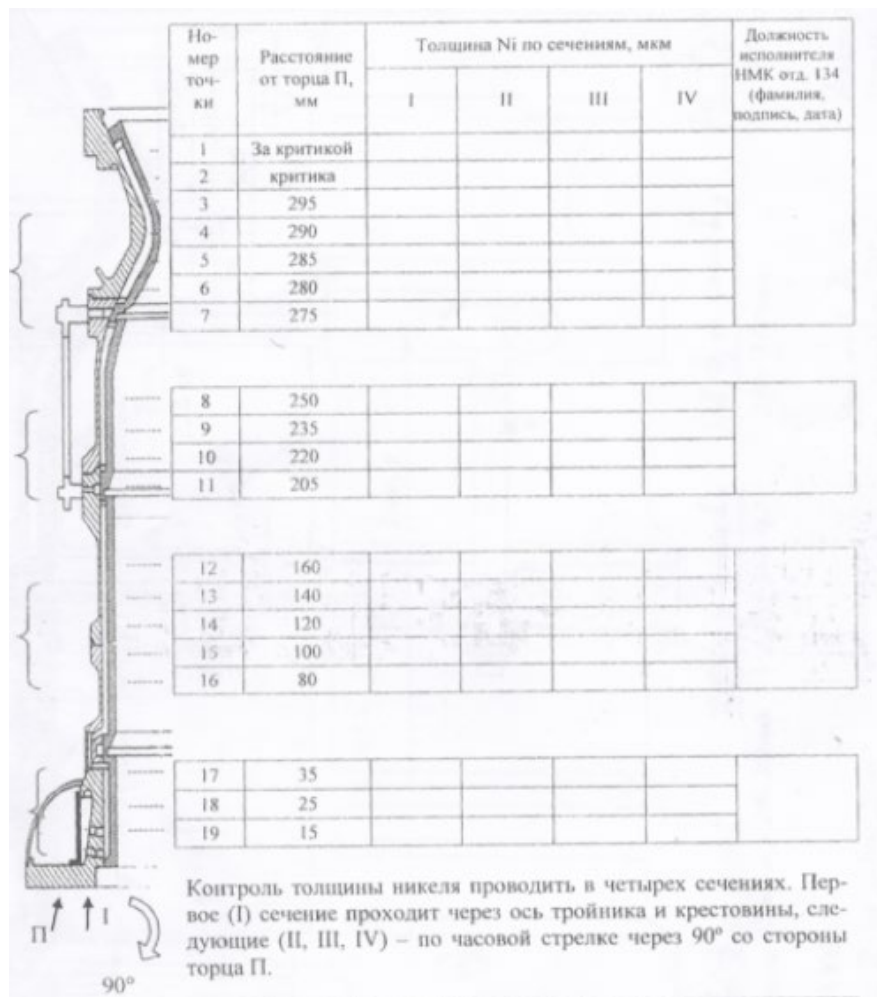
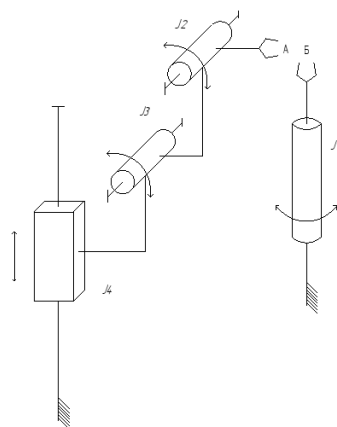


Рис. 1. Карта обмеров толщины никелевого покрытия сложно профильного изделия

Автоматизированная установка неразрушающего контроля толщины никель-хромовых покрытий сборочных единиц предназначена для измерений толщин никелевого и хромового покрытия [2] представлена на рисунке 2, а.



а



б

Рис. 2. Автоматизированная установка НК: а) - автоматизированная установка контроля толщины никель-хромового покрытия; б) – кинематическая схема автоматизированной установки

Данная автоматизированная установка включает в себя модуль позиционирования преобразователя и устройство вращения объекта контроля. Модуль позиционирования имеет в своем составе узел линейного перемещения – J4, узлы вращения J3, J2, место установки преобразователя А и обладает 3 степенями свободы. Устройство вращения объекта контроля включает узел вращения – J1, имеет количество степеней свободы равное 1.

Автоматизированная установка неразрушающего контроля толщины хромового покрытия, представлена на рисунке 3, а.

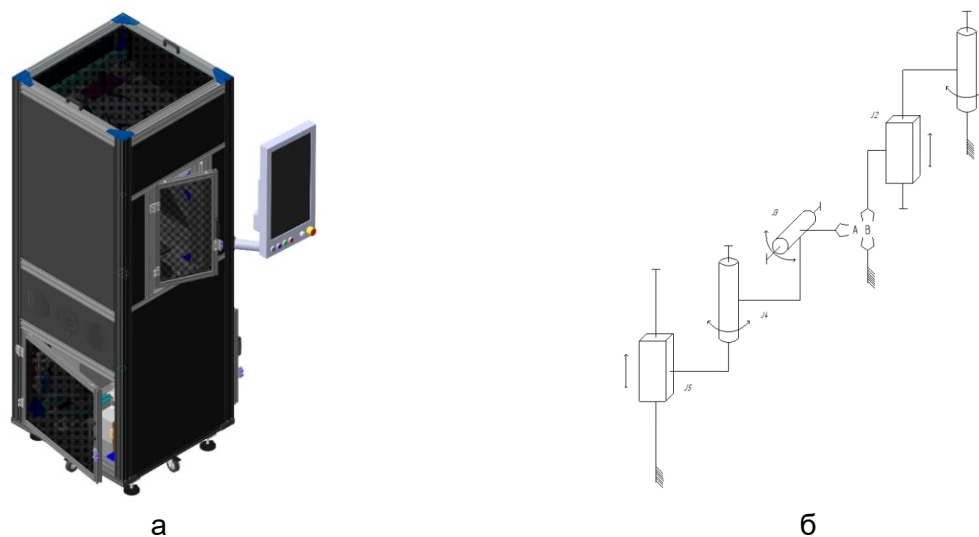


Рис. 3. Автоматизированная установка НК: а) - автоматизированная установка контроля толщины хромового покрытия; б) – кинематическая схема автоматизированной установки

Данный комплекс (рис. 3, б) включает в себя модуль позиционирования преобразователя и устройство вращения объекта контроля. Модуль позиционирования имеет в своем составе узел линейного перемещения – J5, узлы вращения J4, J3, место установки преобразователя А и обладает 3 степенями свободы. Устройство вращения объекта контроля включает узел вращения – J1, узел линейного перемещения – J2 и имеет количество степеней свободы равное 2.

Анализ данных технических систем показывает, что решена проблема контроля всей поверхности изделий сложной геометрической формы толщины никель-хромового покрытия. На качество функционирования сложно-профильных изделий может оказывать влияние нахождение дефектов на поверхности и подповерхностном слое никель-хромового покрытия. Применение комплексного неразрушающего контроля в рамках автоматизированной установки способно определить основные параметры конкретной области сканирования, что значительно повысит качество контроля.

Таким образом, проанализировано современное состояние технологий неразрушающего контроля качества никель-хромового покрытия сложно-профильных изделий, по результатам которого были определены и обоснованы основные недостатки существующих технологий НК на предприятиях, занимающихся производством сложно-профильных изделий, сформулировано и обосновано решение существующих проблем за счет внедрения автоматизированных установок комплексного неразрушающего контроля.

Литература

1. Калошин В.А. Исследование и разработка метода неразрушающего контроля качества никелевых и никель-хромовых покрытий узлов жидкостных ракетных двигателей: дис. ... канд. тех. наук: 05.11.13 – Москва. – 2013.
2. Описание типа средства измерений. Регистрационный № 84038-21 // Федеральный информационный фонд по обеспечению единства измерений. – 2021. – С. 1–6.

УДК 004; 056.55

СРАВНИТЕЛЬНЫЙ АНАЛИЗ ID-BASED СИСТЕМ, ОСНОВАННЫХ НА ФУНКЦИЯХ СПАРИВАНИЯ НА ЭЛЛИПТИЧЕСКИХ КРИВЫХ

Дакуо Ж.-М.Н.¹ (студент), Иогансон И.Д.¹ (студент)

Научный руководитель – Давыдов В.В.¹

¹Университет ИТМО

e-mail: jeandakuo@mail.ru

Работа выполнена в рамках темы НИР №620164 «Методы искусственного интеллекта для киберфизических систем».

Работа частично финансировалась Федеральной программой академического лидерства Приоритет 2030.

Данная работа посвящена рассмотрению различных ID-Based схем подписи с использованием функций спаривания на эллиптических кривых. В ходе работы был проведен сравнительный анализ схем Паттерсона и Хесса.

Ключевые слова: ID-Based криптография, спаривание Вейля, схема Хесса, схема Паттерсона.

В наше время ID-Based схемы подписи, созданные на функциях спаривания Вейля и Тейта, получили новый виток в развитии благодаря тому, что криптография на эллиптических кривых сделала большой шаг вперед. ID Based криптосистемы подходят для внутреннего использования в компаниях, где уже имеется общий доверенный сервер. Данные схемы могут использоваться в рамках государства, где на их основе можно создать новый тип удостоверения личности, с которым можно будет легко и удобно подписывать электронные чеки, квитанции, юридические документы и тому подобное.

ID-Based схемы позволяют любой из сторон создавать открытый ключ из идентификационной информации пользователя и проверять подписи без обмена ключами. Схема предполагает наличие надёжного центра генерации закрытых ключей. Третья доверенная сторона – Private Key Generator (PKG). Целью третьей стороны является предоставление каждому новому пользователю валидного секретного ключа [1–4]. Большим плюсом является то, что секретный ключ не требует обновления при добавлении нового пользователя в систему. ID-Based криптография может успешно использоваться для замкнутого круга пользователей — международная компания, крупный банк и т. п., так как штаб-квартира такой организации способна организовать и администрировать доверенный PKG для данной группы пользователей.

Возможным использование подобных схем сделало наличие такого математического аппарата как спаривание Вейля. Спаривание Вейля над группой n -кращения над эллиптической кривой – сильный и полезный инструмент не только в теоретическом изучении кривых, но и на практике.

Пусть существует эллиптическая кривая E над конечным полем K , и пусть существует натуральное число n , не кратное p характеристике поля K .

Тогда $E[n] \cong Z_n \oplus Z_n$, когда n не кратна p , где $E[n]$ – есть множество всех точек эллиптической кривой, для которых n – есть число вращения.

Пусть $\mu_n = \{x \in \underline{K} \mid x^n = 1\}$, μ_n – группа n -корней из единицы. Тогда спариванием Вейля называется билинейное отображение e_n .

$$e_n: E[n] \times E[n] \rightarrow \mu_n.$$

И это отображение имеет следующие свойства:

- e_n – билинейное:

$$e_n(S_1 + S_2, T) = e_n(S_1, T)e_n(S_2, T),$$

$$e_n(S, T_1 + T_2) = e_n(S, T_1)e_n(S, T_2),$$

для всех $S, S_1, S_2, T, T_1, T_2 \in E[n]$;

- E_n не вырожденное. Это значит если $e_n(S, T) = 1$ для всех $T \in E[n]$, то $S = \infty$ и также наоборот если $e_n(S, T) = 1$ для всех $S \in E[n]$, то $T = \infty$;
- $e_n(T, T) = 1$ для всех $T \in E[n]$;
- $e_n(T, S) = e_n(S, T)^{-1}$ для всех $S, T \in E[n]$;
- практическая эффективность: для всех $T, S \in E[n]$ e_n имеет эффективное вычисление.

Из билинейности следует, что $e_n([m]S, T) = e_n(S, T)^m$. Таким образом, для решения задачи дискретного логарифмирования на эллиптических кривых с помощью спаривания Вейля необходимо вычислить пару $\alpha = e_n(S, T)$ и $\beta = e_n([m]S, T) = \alpha^m$. Обладая двумя этими элементами группы, можно применить алгоритм для поиска n [5].

На рисунке 1 приведен общий вид схемы подписи в ID-Based криптографии.

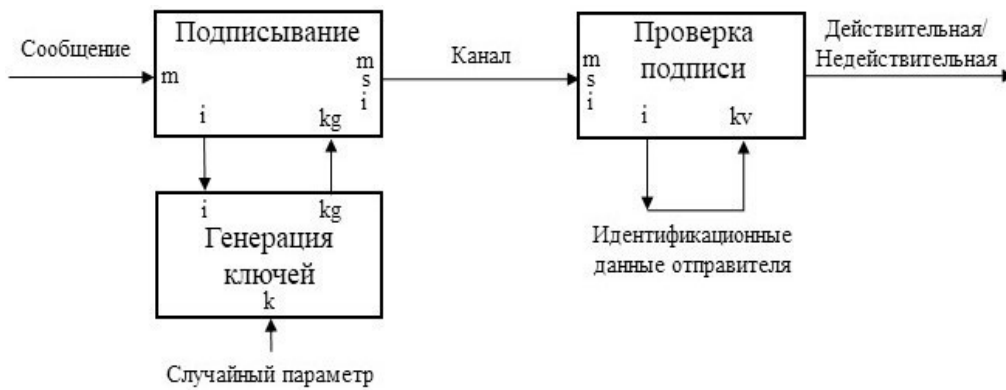


Рис. 1. Схема подписи ID-Based криптографии

Схема Патерсона. Патерсон [2] представил ID-Based схему подписи из билинейных пар, использующих вычислительные примитивы и ключи почти такие же, как и в схеме Боне-Франклина. Предложенная схема подписи состоит из следующих алгоритмов.

Генерация ключей. Пусть G_1 и G_2 – две группы простого порядка q , а P – генератор группы G_1 . Задайте билинейное отображение $e: G_1 \times G_1 \rightarrow G_2$. PKG выбирает $s \in_{\mathcal{R}} Z_q^*$ и устанавливает $P_{pub} = sP$.

Также определим криптографические хэш-функции $H_1: \{0, 1\}^* \rightarrow G_1, H_2: \{0, 1\}^* \rightarrow Z_q$ и $H_3: G_1 \rightarrow Z_q$. P_{pub} – это открытый ключ, а s – главный ключ.

Для заданного публичного идентификатора $ID \in \{0, 1\}$ пользователя PKG вычисляет открытый ключ пользователя $Q_{ID} = H_1(ID)$ и его секретный ключ $S_{ID} = sQ_{ID}$.

Подпись сообщения. Чтобы подписать сообщение $M \in \{0, 1\}$, пользователь сначала выбирает число $k \in Z_q^*$ и вычисляет $S = k^{-1}(H_2(M)P + H_3(R)S_{ID})$, где $R = kP$, подпись на сообщение M – это $(R, S) \in G_1 \times G_1$.

Проверка подписи. Чтобы проверить подпись (R, S) для M , верификатор вычисляет $e(P, P)^{H_2(M)} e(P_{pub}, Q_{ID})^{H_3(R)}$ и принимает подпись тогда и только тогда, когда $e(R, S) = e(P, P)^{H_2(M)} e(P_{pub}, Q_{ID})^{H_3(R)}$

Схема Хесса. Схема подписи Хесса, безопасность которой основана на проблеме Диффи-Хеллмана на функциях спаривания на эллиптических кривых [3]. Кроме того, схема описана в общей базовой ситуации и ее безопасность в первую очередь основана на вычислительной

проблеме. В статье Хесса разобран общий вариант схемы, которую можно использовать с различными криптопримитивами. Утверждается, что эти схемы отличаются от подобных им, а также то, что они имеют преимущества перед этими схемами. Кроме того, представляется возможным, что существуют и другие воплощения общей схемы, использующие, например, RSA вместо спаривания, подобных схеме Шамира. В работе рассматриваются вопросы, связанные с депонированием и распределением ключей среди нескольких доверенных лиц/авторов, и обсуждаются, как схема может быть реализована с использованием эллиптических кривых и спариваний Вейля или Тейта.

Генерация ключей. Пусть G_1 и G_2 — две группы простого порядка q , задайте билинейное отображение $e: G_1 \times G_1 \rightarrow G_2$. P есть генератор группы G_1 . PKG выбирает $s \in_R Z_q^*$ и устанавливает $P_{pub} = sP$.

Также определим криптографические хэш-функции

$$H_1: \{0, 1\}^* \rightarrow G_1^*, H_2: \{0, 1\}^* \times G_2 \rightarrow Z_q^*$$

где $G_1^* = G_1 \setminus \{0\}$. P_{pub} — это открытый ключ, а s — главный ключ.

Для заданного публичного идентификатора $ID \in \{0, 1\}$ пользователя PKG вычисляет открытый ключ пользователя $Q_{ID} = H_1(ID)$ и его секретный ключ $S_{ID} = sQ_{ID}$.

Подпись сообщения. Чтобы подписать сообщение $M \in \{0, 1\}$, пользователь, имеющий свой секретный S_{ID} сначала выбирает точку $P_1 \in G_1^*$, точку $k \in Z_q^*$ и вычисляет:

$$\begin{aligned} r &= e(P_1, P)^k, \\ v &= H_2(m, r), \\ U &= vS_{ID} + kP_1. \end{aligned}$$

Подписью сообщения является пара $(U, v) \in (G_1^*, Z_q^*)$.

Проверка подписи. Для того чтобы проверить подпись (U, v) и с его открытым ключем Q_{ID} сообщения m . Проверяющий вычисляет $r = e(U, P) * e(Q_{ID}, -P_{pub})^v$. И если $v = H_2(m, r)$, то проверяющий принимает подпись.

Схема Патерсона настолько безопасна, насколько любая другая схема подписи, основанная на проблеме Диффи-Хелмана и обобщенной схеме подписи Эль-Гамала. Однако, в своей работе [3] автор не привел никаких формальных доказательств безопасности. Схема Хесса защищена от атак, связанных с подделкой сообщений и отправки специфических сообщений в дополнение к проблеме Диффи-Хелмана.

Сравнивая две схемы, можно заметить, что спаривания $e(P, P)$ и $e(P_{pub}, Q_{ID})$ в схеме Патерсона уникальны и фиксированы для каждого пользователя.

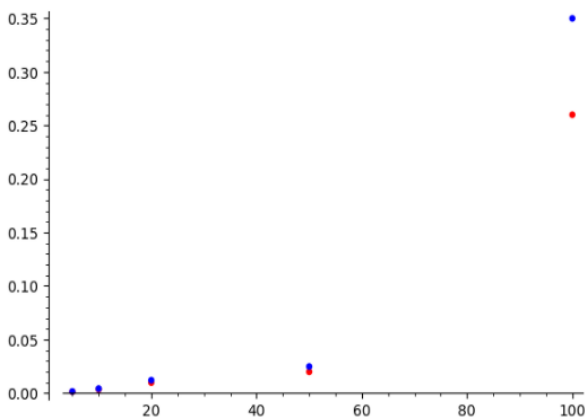


Рис. 2. Быстродействия ID-based схем подписи

Однако, в схеме Хесса $e(P_1, P)$ и $e(Q_{ID}, P_{pub})$ могут быть вычислены заранее, что позволяет подписывать и проверять быстрее. Что говорит нам о том, что схема Хесса эффективнее схемы Патерсона. В равной степени это было проверено на практике. В ПО Sage были смоделированы схемы и посчитаны подписи для одного сообщения для разного количества пользователей. На рисунке 2 представлены результаты. Ось абсцисс — количество посчитанных подписей. Ось ординат — время, затраченное на вычисление. Синие точки — схема Хесса. Красные точки — схема Патерсона.

В ходе работы над статьей были рассмотрены схемы подписи Хесса и Патерсона. Был проведен сравнительный анализ, в ходе которого выяснилось, что схема Хесса эффективнее схемы Патерсона. В дальнейшем планируется разбор дополнительных схем, также разработка модифицированных схем с использованием других типов спаривания.

Литература

1. Ohgishi K., Sakai R., Kasahara M. Cryptosystems based on pairing // The 2000 Symposium on Cryptography and Information Security-SCIS. Okinawa, Japan. – 2000. – Pp. 26–28.
2. Paterson K. G. ID-based signatures from pairings on elliptic curves // Electronics Letters. – 2002. – Т.38. – №18. – Pp. 1025–1026.
3. Hess F. Efficient identity-based signature schemes based on pairings // International Workshop on Selected Areas in Cryptography. – Springer, Berlin, Heidelberg. – 2002. – С. 310–324.
4. Choon J.C., Cheon J.H. An identity-based signature from gap Diffie-Hellman groups // International workshop on public key cryptography. – Springer, Berlin, Heidelberg. – 2003. – С. 18–30.
5. Shamir A. Identity-based cryptosystems and signature schemes // Workshop on the theory and application of cryptographic techniques. – Springer, Berlin, Heidelberg. – 1984. – С. 47–53.

УДК 004.418

СРЕДСТВО ДЛЯ АВТОМАТИЗАЦИИ КОНТЕЙНЕРИЗАЦИИ ПРИЛОЖЕНИЙ С ВОЗМОЖНОСТЬЮ РАСШИРЕНИЯ ИСПОЛЬЗУЕМОГО СТЕКА ТЕХНОЛОГИЙ

Дербенёва Е.С.¹ (студент)

Научный руководитель – канд. пед. наук, доцент Авксентьева Е.Ю.¹

¹Университет ИТМО

e-mail: katderbi@gmail.com

В работе рассматривается проблема создания средства автоматизации контейнеризации приложения. Были проанализированы существующие средства автоматизации контейнеризации, рассмотрены инструменты, которые можно использовать в разрабатываемом инструменте, после чего разработан алгоритм и архитектура системы, с учетом того, что сторонний разработчик может добавить свои модули для работы системы на других стеках технологий.

Ключевые слова: docker, контейнеризация, автоматизация, DevOps, CI/CD.

На пути продвижения каждого продукта через цикл разработки, после передачи его заказчику и при эксплуатации зачастую встречается множество препятствий. Помимо работы по подготовке приложения к эксплуатации в различных условиях, можно также столкнуться с проблемами отслеживания зависимостей, масштабирования приложения и обновлений отдельных компонентов, не затрагивающих непосредственно само приложение.

На данный момент использование технологий контейнеризации для решения этих проблем – скорее правило, чем исключение. Контейнеры используются как в инфраструктуре, так и в процессах разработки программного обеспечения (ПО). Сфера контейнеризации быстро развивается, но в ней уже появилось два наиболее распространенных продукта – это платформа Docker [1] и ПО для оркестрации контейнеров Kubernetes.

Из-за популярности и из-за предлагаемого человекочитаемого формата специального файла Dockerfile, который описывает процесс построения контейнеров, для изучения и использования была выбрана платформа Docker. На практике, процесс первоначальной контейнеризации приложения и дальнейшей его поддержки требует временных затрат от разработчиков, так как зачастую файл Dockerfile содержит большое количество инструкций, которые имеют общую часть между приложениями с одинаковым технологическим стеком, поэтому для новых приложений модифицируются уже существующий.

Для решения этой проблемы разработаны специальные средства автоматизации контейнеризации приложения. Одни средства заточены только под определенный язык программирования, другие уже устарели и не обновляются.

Jib – плагин для систем сборки Maven и Gradle, не требующий долгой конфигурации [2]. Это инструмент с открытым исходным кодом, который поддерживается компанией Google для создания образов Docker Java приложений. Это позволяет упрощать контейнеризацию, так как пропадает необходимость в написании файла Dockerfile. Более того, разработчику даже не нужно устанавливать саму платформу Docker, чтобы создавать и публиковать образы.

Docker-maven-plugin - этот плагин был первым Maven плагином, который использовался в Spotify для создания образов Docker из Java сервисов [3]. Данный плагин позволяет контейнеризировать приложение на основе конфигурации в файле pom.xml, для этого нужно

указать образ и зависимости. Со временем разработчики плагина осознали, что самый простой способ создать образ Docker из проекта Java – с помощью файла Dockerfile. Это привело к созданию второго плагина Maven для создания образов Docker, dockerfile-maven, так же от компании Spotify, который предлагает более простой способ работы с Docker из Maven.

Dockerfile Maven – является продолжением первого плагина, с тем отличием, что делает уклон на создание Dockerfile [4]. Процесс сборки Docker интегрирован с процессом сборки Maven, позволяя привязывать к Maven фазам создание образа Docker. Несмотря на интеграцию с Maven, в плагине была выявлена проблема указания относительных и абсолютных путей: так как для создания образа необходим относительный путь до jar-архива приложения, который в Maven является абсолютным путем.

На данный момент, первый из продуктов устарел и больше не поддерживается, а второй, согласно информации в репозитории проекта, больше не разрабатывается и не принимает запросы на новые функции или исправление не критических ошибок.

Однако, такие инструменты тоже имеют свои конфигурационные файлы, что ведет к проблеме конфигурации ради конфигурации. Очевидно, что для решения данной проблемы, необходим инструмент, который на основе анализа проекта сможет автоматически сформировать требования к образу Docker контейнера, а впоследствии и выполнить сборку образа.

На данный момент уже существует подобное решение проблемы. Инструмент Whales помогает создавать необходимые для контейнеризации файлы на основе статического анализа кода, но у данного решения есть несколько недостатков, из-за которых оно не является удовлетворяющим:

1. Нет возможности расширить поддержку языков стороннему разработчику, то есть данное решение поддерживает только следующие языки программирования и платформы Ruby, PHP, Python и Node.
2. Долгое время проект уже не поддерживается (последнее обновление в мае 2019 года).

Более того выбор языков очень необычен, потому что согласно рейтингу, представленному Институтом инженеров электротехники и электроники (IEEE), наиболее популярными языками программирования в 2017 г., в год начала разработки продукта, являлись:

- Python;
- Java;
- C [5].

Сравнение рассмотренных средств представлено в таблице.

Таблица. Средства автоматизации контейнеризации

Средство	Поддерживаемые языки программирования	Обновляется	Поддерживаемая система сборки	Дополнительная конфигурация
Jib	Java	+	Gradle, Maven	+
Docker-maven-plugin	Java	-	Maven	+
Dockerfile Maven	Java	-	Maven	+
Whales	Ruby, Python, PHP	-	-	-

Jib, Docker-maven-plugin и Dockerfile Maven поддерживают только один язык программирования. Более того, для их работы необходима система сборки приложения, что не подходит для проектов без системы сборки. Whales расширяет количество поддерживаемых языков программирования, но эти языки не являются наиболее распространенными. Кроме того, Whales, Docker-maven-plugin и Dockerfile Maven уже не поддерживаются разработчиками.

После анализа существующих решений, было решено, что разрабатываемый инструмент должен быть расширяемым, чтобы сторонний инженер-программист мог написать расширение, благодаря которому можно было использовать инструмент для стека технологий, на которое он не был рассчитан.

Алгоритм работы инструмента состоит из определения с помощью статического анализа проекта параметров для формирования Dockerfile:

1. Язык программирования.
2. Зависимости приложения.

Согласно рейтингу IEEE на 2021 год, тройка наиболее популярных языков программирования не изменилась, поэтому было решено добавить их поддержку в систему. Данные языки используются в различных сферах разработки ПО: мобильной, веб-разработки, разработки драйверов и софта для микроконтроллеров, но для расширения использования инструмента в фронтенд разработке, было решено вместо C анализатора реализовать Node.js анализатор.

Определение основного языка приложения в общем случае осуществляется за счет анализа расширения файлов внутри проекта. Так можно воспользоваться инструментом Linguist, который на данный момент используется на платформе GitHub для определения используемого языка в репозитории проекта. Данный инструмент использует множество стратегий для определения языка: от расширения файла до наивного байесовского классификатора.

Кроме того, можно воспользоваться инструментом Guesslang, который используется в IDE Visual Studio Code. Для определения языка программирования используется машинное обучение, точность угадывания более 90%.

Проанализировать используемые технологии можно с помощью анализа системы сборки приложения на основе поиска специфичных для системы сборки файлов. Например, для определения того, что проект собирается через Gradle, идет поиск наличия файлов:

1. С расширением .gradle или .gradle.kts.
2. Gradle-wrapper.jar.
3. Gradlew или gradlew.bat.

На основе языка приложения и системы сборки выбирается основной образ для создания образа контейнера, а также шаги для сборки.

В качестве поддерживаемых системы сборки были выбраны Gradle, Maven, npm и venv, так как они являются наиболее распространенными при разработке на Java, Node.js и Python.

При отсутствии системы сборки приложения необходим алгоритм получения зависимостей приложения. Рассмотрим пример, допустим в директории есть файл main.py и файл requirements.txt. На основе файлов в проекте было получено, что в проекте нет системы сборки, а языком приложения является Python 3. Так как используемый язык Python 3, то в качестве базового образа Docker будет выбран python:3-alpine. Так как не было найдено никаких систем сборки, то будет произведен поиск файлов, которые могут содержать зависимости для проекта на Python. В данном случае зависимости проекта лежат в requirements.txt, они будут установлены через pip. Во время старта контейнера будет запущена точка входа приложения на Python, для этого будет предварительно найден файл, который содержит в себе определение функции main(), которая является точкой входа для Python приложений.

С другой стороны, для выявления зависимостей в Java приложениях можно использовать поиск по паттерну в файлах с расширением .java, так как для использования классов из других пакетов (зависимостей), их необходимо указать с помощью директивы "import". После получения списка пакетов (классов) необходимо найти все jar-архивы, распаковать их и провести поиск по файлам на наличие там необходимых зависимостей. При отсутствии всех необходимых архивов, пользователю будет выведена ошибка.

Если же в Java проекте используется система сборки, то в зависимости от нее, происходит поиск файла, где указаны зависимости приложения. Так же, как и в случае с проектом без системы сборки, при отсутствии архивов указанных зависимостей, пользователю будет выведена ошибка.

В рамках анализа зависимостей, если в проекте будут присутствовать неудовлетворенные зависимости, пользователю будет выведена детализированная ошибка. Без данного анализа пользователь получить образ приложения, который не будет работать.

После анализа существующих решений, был разработан алгоритм работы инструмента для автоматической контейнеризации проекта.

Алгоритм работы системы состоит из следующих шагов:

1. Пользователь указывает проект для генерации Dockerfile.
 - а. Пользователь может также указать: стек проекта, путь до собственных реализаций анализаторов для добавления поддержки нового стека технологий, данные для публикации полученного образа.
2. Если пользователь не указал стек проекта, то язык программирования и система сборки выявляется на основе файлов проекта.
3. Проверка есть ли анализатор для данного языка и системы сборки.
 - а. Если анализатор не найден, то выводится сообщение с ошибкой.
4. Генерируется Dockerfile на основе информации, полученной с помощью статического анализа.
5. Если были указаны параметры для публикации образа, то он собирается и публикуется на основе созданного Dockerfile.

Архитектура системы разрабатывалась на основе нескольких требований, главное из которых – возможность подключить расширение, благодаря которому можно использовать инструмент для стека технологий, на которое он не был рассчитан.

Системе было дано название Docker Automation Packer (DAP). Были выделены следующие модули:

- Resolver;
- Handler selector;
- Handler;

Предлагаемая архитектура системы представлена на рисунке 1.

Система получает на вход набор файлов проекта, который поочередно передает различным реализациям модуля Resolver, которые содержатся в системе DAP, в том числе добавленные сторонним разработчиком.

Модуль Resolver служит для определения стека проекта, а именно языка программирования, на котором разрабатывается проект, и системы сборки.

Модуль Handler Selector на основе информации о стеке технологий, который был получен от модуля Resolver выбирает реализацию Handler из списка обработчиков, которые содержатся в DAP и которые определил сторонний разработчик. После чего модуль Handler на основе полученной информации от модуля Resolver создает Dockerfile.

Для поддержки расширяемости были определены следующие интерфейсы, которые должен реализовать разработчик для добавления поддержки нового стека в систему. Интерфейсы системы представлены на рисунке 2.

Класс Resolver имеет метод resolveStack, который определяет стек проекта. Класс Handler генерирует Dockerfile в методе generateDockerFile для тех стеков, который может обработать в методе canHandler. После реализации данных классов, программисту необходимо указать директорию с данными классами при генерации Dockerfile для проекта.

Таким образом, в рамках данной работы был проведен анализ существующих систем автоматизации контейнеризации приложения. Было определено, что существующие системы не являются актуальными или являются очень узкоспециализированными из-за языка программирования, которые они поддерживают. Для уменьшения недостатков существующих

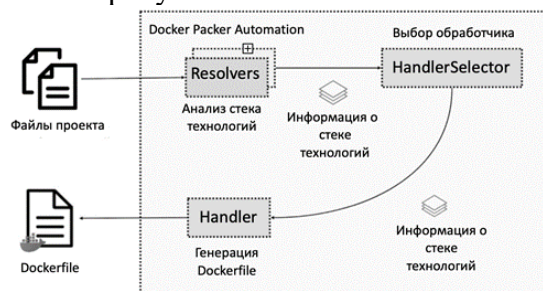


Рис. 1. Архитектура инструмента для автоматизации контейнеризации приложения

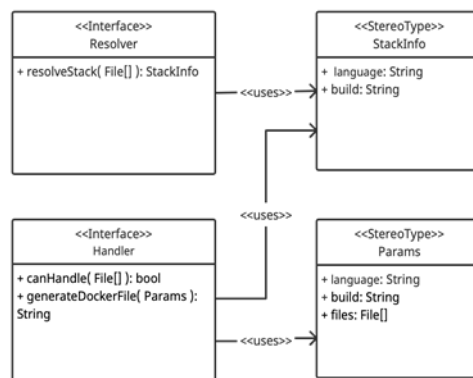


Рис. 2. UML диаграмма системы

систем контейнеризации, был создан алгоритм работы инструмента, который позволит автоматически контейнеризовать приложение. Более того, данный инструмент является расширяемым за счет разработанной архитектуры.

Дальнейшие исследования планируется проводить в области управления контейнеризованным приложением, а также внедрения инструмента в процессы CI/CD.

Литература

1. Docker. [Электронный ресурс]. – Режим доступа: <https://docs.docker.com> (дата обращения: 09.12.2021).
2. Introducing Jib. [Электронный ресурс]. – Режим доступа: <https://cloud.google.com/blog/products/gcp/introducing-jib-build-java-docker-images-better> (дата обращения: 09.12.2021).
3. Docker-maven-plugin. [Электронный ресурс]. – Режим доступа: <https://github.com/spotify/docker-maven-plugin> (дата обращения: 09.12.2021).
4. Dockerfile maven. [Электронный ресурс]. – Режим доступа: <https://github.com/spotify/dockerfile-maven> (дата обращения: 09.12.2021).
5. The 2017 Top Programming Languages. [Электронный ресурс]. – Режим доступа: <https://spectrum.ieee.org/the-2017-top-programming-languages> (дата обращения: 09.12.2021).

УДК 681.78

АНАЛИЗ ИСТОЧНИКОВ ПОГРЕШНОСТЕЙ ЯДЕРНОГО МАГНИТНОГО ГИРОСКОПА ПРИ МЕХАНИЧЕСКИХ ВОЗДЕЙСТВИЯХ

Завитаев А.С.¹ (студент)

Научный руководитель – д-р техн. наук Евстифеев М.И.¹

¹Университет ИТМО

e-mail: zav.al98@mail.ru

В работе рассматривается ядерный магнитный гироскоп. В ходе работы гироскоп испытывает различные физические нагрузки. В результате данных нагрузок возможны смещения и деформации элементов конструкции, что приводит к возникновению погрешностей при работе прибора. В работе рассмотрены влияния деформаций на изменение магнитного поля в ячейке, а также изменение мощностей излучений накачки и детектирования.

Ключевые слова: ядерно-магнитный гироскоп, Стео, метод конечных элементов, погрешности, механические воздействия.

Ядерный магнитный гироскоп (ЯМГ) – квантовый гироскоп, чувствительным элементом которого является ансамбль ориентированных атомных ядер, обладающий макроскопическим магнитным моментом (рисунок). Принцип действия ЯМГ основан на зависимости частоты прецессии магнитного момента ядер в постоянном магнитном поле от угловой скорости вращения гироскопа [1–3].



Рисунок. Ядерно-магнитный гироскоп

С целью расширения эксплуатационных характеристик и использования прибора на высокодинамичных объектах необходимо проанализировать источники погрешностей ядерного магнитного гироскопа.

В качестве источников рассматривались деформации и смещения элементов конструкции. С целью их анализа была построена математическая расчетная модель.

В САПР системе ProEngineer была построена конечно-элементная модель корпуса прибора. В качестве допущения, не влияющего на точность будущих расчетов, из модели были исключены фаски, скругления, крепежные элементы [4].

В результате моделирования были получены резонансные частоты 1229, 2265, 2519, 2857 и 2978 Гц.

Основным параметром, влияющим на точность работы гироскопа, является степень ориентируемости ядер активного вещества в капсуле [5, 6]. Необходимое значение данного параметра достигается за счет однородности магнитного поля и положения лазера накачки относительно ячейки. Следуя из этого, исследовалась деформация и смещение следующих компонентов:

1. Лазер оптической накачки.
2. Лазер опроса.
3. Кольцевые катушки.

При деформациях и горизонтальных смещениях катушек изменяется направление магнитного поля, которое определяет ось чувствительности прибора, внутри ячейки.

При вертикальных смещениях направление магнитного поля не изменяется, однако, внутри ячейки возникает градиент поля. При смещении на 0,1 мм возникает градиент магнитного поля ~ 7 нТл. Деформации экранов увеличивают значения этого градиента. Его наличие влияет на время релаксации атомов. Из-за этого ядра ксенона не приходят в резонанс при определенной частоте. Возникшая неоднородность резонанса приводит к уменьшению предельной чувствительности гироскопа.

При смещениях лазеров накачки уменьшается объем ориентированных атомов цезия. Что влияет на поляризацию ксенона. Смещение лазера детектирования изменяется объем «опрашиваемых» атомов ксенона. Оба этих изменения приводят к снижению предельной чувствительности гироскопа.

Заключение

1. Проведен анализ источников погрешностей ядерного магнитного гироскопа.
2. Построена конечно-элементная модель ЯМГ.
3. Проведен расчет и анализ смещений и деформаций конструкции.
4. Рассмотрены влияния деформаций и смещений катушек прибора на его погрешности.
5. Рассмотрены влияния деформаций и смещений лазеров прибора на его погрешности.

Литература

1. Шевченко А.Н., Захарова Е.А. Исследование зависимости качества резонанса от распределения цезия в ячейке квантового датчика вращения // Научно-технический вестник информационных технологий, механики и оптики. – 2019. – Т.19. – №4. – С. 567–573. DOI: 10.17586/2226-1494-2019-19-4-567-573.
2. Литманович Ю.А., Вершовский А.К., Пешехонов В.Г. Гироскоп на основе явления ядерного магнитного резонанса: прошлое, настоящее, будущее // Материалы пленарного заседания 7-й Российской мультikonференции по проблемам управления ОАО «Концерн «ЦНИИ «Электроприбор». – 2014. – С. 35–42.
3. Larsen M., Bulatowicz M. Nuclear magnetic resonance gyroscope: for DARPA's micro-technology for positioning, navigation and timing program // Proc. IEEE Int. FrequencyControl Symposium. – 2012. DOI: 10.1109/fcs.2012.6243606.
4. Аванесов Ю.Л., Буканова А.Н., Воронов А.С., Евстифеев М.И. Оптимизация конструктивных параметров глубоководного индукционного датчика скорости // Научно-технический вестник информационных технологий, механики и оптики. – 2018. – Т.18. – №1. – С. 140–146. DOI: 10.17586/2226-1494-2018-18-1-140-146.
5. Александров Е.Б., Константинов О.В., Перель В.И. Оптическая ориентация атомов в магнитном поле, перпендикулярном лучу // ЖЭТФ. – 1965. – Т.49. – №1(7). – С.97–106.
6. Вершовский А.К., Литманович Ю.А., Пазгалев А.С., Пешехонов В.Г. Гироскоп на ядерном магнитном резонансе: предельные характеристики // Гироскопия и навигация. – 2018. – Т.26. – №1. – С. 55–80. DOI 10.17285/0869-7035.2018.26.1.055-080.

УДК 004; 056.55

СХЕМА РАЗДЕЛЕНИЯ СЕКРЕТА НА МНОГОМЕРНЫХ УРАВНЕНИЯХ

Иогансон И.Д.¹ (студент), Дакуо Ж.-М.Н.¹ (студент)

Научный руководитель – д-р техн. наук, доцент Беззатеев С.В.¹

¹Университет ИТМО

e-mail: ivan.ioganson@yandex.ru

Работа выполнена в рамках темы НИР №620164 «Методы искусственного интеллекта для киберфизических систем».

Работа частично финансировалась Федеральной программой академического лидерства Приоритет 2030.

В данной работе приведена схема разделения секрета, разработанная на основе схемы Шамира, использующая в качестве секрета многомерные полиномы. На основе представленной схемы в будущем планируется реализовать схему постквантовой пороговой подписи на основе алгоритма Rainbow.

Ключевые слова: схема разделения секрета, протоколы, постквантовая криптография, многомерные уравнения, протокол конфиденциального вычисления.

На сегодняшний день стало ясно, что появление первых квантовых компьютеров с большим количеством кубит всего лишь вопрос времени. Поэтому важно уже сейчас готовиться к этому и изучать постквантовые криптографические схемы. Одной из таких схем является схема электронной подписи на основе многомерных уравнений Rainbow. Данная схема входит в тройку финалистов третьего раунда конкурса NIST (National Institute of Standards and Technology) на звание стандарта электронной подписи, что говорит о высоком уровне доверия к данной схеме.

Схемы пороговой подписи нужны для того, чтобы любая коалиция из уполномоченных пользователей, количеством не менее порогового значения, могла сформировать электронную подпись от лица всех уполномоченных пользователей. Подобные схемы имеют широкое распространение в децентрализованных системах, таких как, например, криптовалюты. Однако, на сегодняшний день не существует полноценных постквантовых схем пороговой подписи.

Разумеется, для классической (доквантовой) криптографии уже существуют подобные схемы. Примером может послужить работа [1], в которой представлена схема пороговой подписи на основе схемы разделения секрета Шамира и подписи RSA (Rivest, Shamir и Adleman). В постквантовой криптографии схем разделения секрета, оперирующих не числами, а многомерными полиномами почти нет. Единственное, что удалось найти, это работа [2], в которой представлена схема разделения секрета на многомерных полиномах, использующая Китайскую теорему об остатках. Однако, данная схема не удовлетворяет некоторым из требований, которые будут представлены далее.

Схема разделения секрета должна удовлетворять следующим требованиям:

1. Размер тени должен быть примерно равен размеру секрета.
 2. Идеальность: сторона, знающая менее k теней, где k – порог, не может получить никакой информации о секрете.
 3. Возможность обновления секрета без привлечения третьей доверенной стороны.
 4. Возможность создания новой тени без разглашения секрета.
 5. Возможность использования секрета в алгоритмах без его разглашения.
-

Разработанная схема основана на схеме разделения секрета Шамира [3] и выглядит следующим образом. Пусть дано поле G . Тогда полиномом $f \in G[X]$, где $X = \{x_1, x_2, \dots, x_m\}$, будет называться выражение (1):

$$f(X) = \sum_i a_i x_1^{e_{i1}} x_2^{e_{i2}} \dots x_m^{e_{im}}, \quad (1)$$

где $a_i \in G$, а $e_{i1}, e_{i2}, \dots, e_{im} \in N \cup \{0\}$.

Для разделения многочлена $s(X) \in G[X]$ между n пользователями, так чтобы любые k из них собравшись смогли восстановить $s(X)$, необходимо выбрать $k - 1$ случайных полиномов $f_2(X), \dots, f_k(X) \in G[X]$.

Тогда тенью пользователя под номером $i \in G/\{0\}$ будет являться полином $s_i(X) \in I$, определенный формулой (2):

$$s_i(X) = s(X) + f_2(X) * i + \dots + f_k(X) * i^{k-1}. \quad (2)$$

Для восстановления $s(X)$ необходимо, собрав k теней $s_{i_1}(X), s_{i_2}(X), \dots, s_{i_k}(X)$, посчитать по формуле (3):

$$s(X) = \sum_{j=1}^k s_{i_j}(X) * l_j, \quad (3)$$

где $l_j = \prod_{t \neq j} \frac{i_t}{i_t - i_j}$.

Так как представленная схема основана на схеме Шамира, то первые два пункта из требований автоматически выполняются. Третий пункт, обновление секрета, происходит следующим образом: для секрета $s(X)$ необходимо выбрать полином $\Delta(X)$ и тогда каждый пользователь i сможет вычислить новую тень $s_i'(X) = s_i(X) + \Delta(X)$. Тогда новый секрет будет равен $s'(X) = s(X) + \Delta(X)$.

Для выполнения четвертого пункта, создания новой тени, необходимо, чтобы собралось минимум k пользователей с идентификаторами i_1, \dots, i_k , уже имеющих тени, и каждый из них должен вычислить $d_{i_j}(X) = s_{i_j}(X) * l_j'(v)$, где $j \in \{1, \dots, k\}$, а $l_j'(v) = \prod_{t \neq j} \frac{v - i_t}{i_j - i_t}$. И тогда, используя протокол безопасной суммы, они вычисляют $s_v(X) = \sum_{j=1}^k d_{i_j}(X)$, так чтобы результат знал только пользователь v .

Что касается пятого требования, то оно зависит от конкретной задачи. В случае пороговой подписи с помощью алгоритма Rainbow, необходимо иметь возможность находить решение системы линейных уравнений, не разглашая при этом саму систему. Назовем это слепым решением системы линейных уравнений.

Пусть дана система линейных уравнений $F(X) \in G[X]^m$, вида (4):

$$F(X) = \begin{cases} f_1(X) \\ \dots \\ f_m(X) \end{cases}. \quad (4)$$

Тогда решением системы линейных уравнений $F(X) = Y'$, где $Y' \in G^m$, назовем такое $X' \in G^m$, что $F(X') = Y'$. Разделим данную систему линейных уравнений по методу, приведенному выше. Тогда пользователь под номером i будет иметь в качестве тени систему линейных уравнений $S_i(X)$ вида (5):

$$S_i(X) = \begin{cases} s_{i1}(X) \\ \dots \\ s_{im}(X) \end{cases}, \quad (5)$$

где $s_{ij}(X)$ – тень полинома $f_j(X)$ для пользователя с номером i .

Пусть есть k участников с тенями $S_{i_1}(X), S_{i_2}(X), \dots, S_{i_k}(X)$ системы линейных уравнений $F(X)$. Тогда для решения системы линейных уравнений $F(X) = Y'$, где $Y' \in G^m$, необходимо произвести следующие действия:

1. Зная значения i_1, i_2, \dots, i_k , каждый участник с номером i_j вычисляет свое значения

$$l_j = \prod_{t \neq j} \frac{i_t}{i_t - i_j}.$$

2. Участники совместно генерируют случайное обратимое аффинное преобразование $L: G^m \rightarrow G^m$.
3. Каждый участник с номером i_j вычисляет (6):

$$\bar{S}_{i_j}(X) = \begin{cases} \bar{s}_{i_j 1}(X) = l_j * s_{i_j 1}(L(X)) \\ \dots \\ \bar{s}_{i_j m}(X) = l_j * s_{i_j m}(L(X)) \end{cases}. \quad (6)$$

4. Каждый участник посылает свой $\bar{S}_{i_j}(X)$ специально выделенному участнику, не входящему в множество участников с номерами i_1, i_2, \dots, i_k (назовем его Решатель). У него может вообще не быть своей тени секрета. Важно чтобы Решатель не знал преобразование L .
5. Решатель вычисляет систему $\bar{F}(X)$ по формуле (7):

$$\bar{F}(X) = \begin{cases} \bar{f}_1(X) = \sum_{j=1}^k \bar{s}_{i_j 1}(X) \\ \dots \\ \bar{f}_m(X) = \sum_{j=1}^k \bar{s}_{i_j m}(X) \end{cases}. \quad (7)$$

6. Решатель находит решение \bar{X}' для системы $\bar{F}(X) = Y'$.
7. Решатель отправляет \bar{X}' остальным участникам.
8. Каждый участник вычисляет $X' = L(\bar{X}')$.

X' и является решением системы линейных уравнений $F(X) = Y'$.

В данной схеме все участники используют общее секретное преобразование L , чтобы «засолить» свои тени. Благодаря этому Решатель не сможет ничего узнать ни о тенях участников, ни о секретной системе $F(X)$.

Приведенное выше решение позволяет разделить систему многомерных уравнений между участниками так, чтобы любые k , собравшись вместе, смогли восстановить секрет.

В дальнейшем планируется разработать на основе данной схемы протокол, позволяющий подписывать сообщения с помощью постквантового алгоритма Rainbow, не разглашая при этом секрет.

Литература

1. Shoup V. Practical threshold signatures // International Conference on the Theory and Applications of Cryptographic Techniques. – Springer, Berlin, Heidelberg. – 2000. – С. 207–220.
2. Galibus T., Matveev G. Finite Fields, Gröbner Bases and Modular Secret Sharing // Journal of Discrete Mathematical Sciences and Cryptography. – 2012. – Т.15. – №6. – С. 339–348.
3. Shamir A. How to share a secret // Communications of the ACM. – 1979. – Т.22. – №11. – С. 612–613.

УДК 67.05

ПРОБЛЕМЫ ПРИМЕНЕНИЯ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА В АДДИТИВНОМ ПРОИЗВОДСТВЕ

Зименко К.В.¹ (аспирант)

Научный руководитель – канд. техн. наук, доцент Афанасьев М.Я.¹

¹*Университет ИТМО*

e-mail: kvzimenko@itmo.ru

Работа выполнена в рамках темы НИР №620164 «Методы искусственного интеллекта для киберфизических систем».

В статье рассматриваются основные области применения и методы искусственного интеллекта в аддитивном производстве, а также описаны текущие проблемы и препятствия на пути распространения технологий искусственного интеллекта в данной области, а также направления для дальнейшего развития.

Ключевые слова: ЧПУ система, искусственный интеллект, аддитивное производство, трехмерная печать, машинное обучение, генетические алгоритмы, экспертные системы.

В настоящий момент наблюдается стремительное развитие аддитивного производства (АП) [1]. АП демонстрировало активный рост в течение последних 40 лет, превратившись из многообещающей группы некоммерческих технологий в начале 1980-х годов в рынок, с оценочной стоимостью более 21 миллиардов долларов в 2020 году [2]. Благодаря своим многочисленным преимуществам, таким как расширенные возможности дизайна и быстрого прототипирования, экономия ресурсов, минимизация процесса подготовки производства, АП стало популярной тенденцией в производственных процессах. Оборудование для АП на сегодняшний день является одной из наиболее распространенных и активно развивающихся технологических систем с числовым программным управлением (ЧПУ). Однако, существует ряд ограничений данной группы технологий, препятствующих ее широкому распространению в приборостроении, к которым относятся различные дефекты печати, а также непостоянство качества получаемых изделий [1]. Этих проблем можно избежать, внедряя алгоритмы обнаружения дефектов и прогнозирования параметров печати, что спровоцировало распространение в данной области методов искусственного интеллекта (ИИ). В настоящей статье приведены основные области применения и методы ИИ в АП, а также описаны текущие проблемы и препятствия на пути распространения технологий ИИ в данной отрасли приборостроения, а также направления для дальнейшего развития.

В статье под понятием ИИ понимаются алгоритмы нечеткой логики, экспертные системы, генетические алгоритмы и методы машинного обучения (МО). Последние получили наибольшее распространение в силу разнообразия различных алгоритмов и технологий, а также доступности данных для обучения. Задачи машинного обучения можно разделить на три основные категории: обучение с учителем, обучение без учителя и обучение с подкреплением. Среди них в области АП преобладают категории обучения с учителем и без учителя из-за большого объема потенциально доступных данных.

Методы обучения с учителем применяются для задачи прогнозирования каких-либо значений, к которым относится определение свойств изделий или оптимизация параметров, а также для классификации, например, для оценки качества или обнаружения дефектов напечатанных изделий. Алгоритмы обучения без учителя работают с незамеченными

данными, например, с изображениями или акустическими данными для решения задач кластеризации, например, при оценке стоимости и контроля отклонений геометрии.

Для того, чтобы охарактеризовать области применения методов ИИ в рассматриваемой области, цикл АП был классифицирован на четыре этапа: проектирование, оптимизация процесса, контроль процесса производства (в том числе и в режиме реального времени) и проверка.

На этапе проектирования важными задачами являются определение конструкции изделия и его требуемых свойств. Для их решения применяют экспертные системы, методы кластеризации и классификации, а также методы МО для предсказания параметров детали. Например, применяются нейронные сети для прогнозирования массы детали, требуемых свойств материала и времени построения [3]. Также применяются алгоритмы для разработки схем оценки затрат, используя доступные данные об аналогичных продуктах и производстве. Распространены объединения нескольких методов ИИ для решения конкретной проблемы, например, был разработан алгоритм, в котором были объединены генетическое программирование и нейронная сеть. Эта комбинированная модель была применена для прогнозирования износостойкости аэрокосмических деталей, изготовленных методом экструзии материала [3].

На этапе оптимизации процесса производства необходимо определить взаимосвязи между параметрами процесса, конечной структурой материала, свойствами и производительностью. Хотя методы конечно-элементного моделирования (FEM) обеспечили определенный успех в отображении данных взаимосвязей, точное представление процессов АП с использованием высокоточного моделирования затруднено. Модели, основанные на физике, сложны и требуют глубокого понимания свойств материалов и физических законов, регулирующих процесс АД, а модели с низкой точностью страдают от недостатка информации о физических свойствах, особенно из-за различий в оборудовании и материалах. С другой стороны, методы машинного обучения могут успешно обнаруживать сложные отношения, преодолевая многие ограничения, связанные с методами, перечисленными выше. Основной задачей ИИ на данном этапе является прогнозирование различных параметров процесса. Для этого применяются методы регрессии, нейронные сети и генетические алгоритмы. Например, была успешно спрогнозирована ширина расплава при лазерном спекании сочетанием методов регрессии и генетических алгоритмов [3].

На этапе мониторинга процесса необходимы методы, которые способны работать с неупорядоченными данными, например, с изображениями, в том числе и в режиме реального времени. Например, здесь нашли применение методы МО с подкреплением и без учителя. А также применяются методы принятия решений для отбраковки и контроля качества продукции. Например, была представлена модель «квалификации по мере строительства», в которой методы МО используют данные датчиков в реальном времени для определения условий процесса, которые могут вызвать пористость [4].

Наконец, методы ИИ используются для окончательной проверки полученных деталей. Основное внимание уделяется метрологии поверхности, а также обнаружению и классификации дефектов с использованием измерений после производства. Например, был разработан новый метод классификации вариаций размеров деталей, изготовленных методом экструзии материала, на основе теории спектральных графов [4]. После извлечения признаков из данных трехмерного облака точек, отсканированных лазером, были применены контролируемые методы МО для классификации размерных вариаций, включая разреженное представление, нейронную сеть, наивную байесовскую модель, метод опорных векторов и дерево решений.

Несмотря на широкое распространение ИИ в АП, остается еще большое число областей и задач, в которых возможности и преимущества применения технологий ИИ остаются неисследованными. Например, возможно дальнейшее изучение вопроса проектирования материалов для трехмерной печати на основе МО с целью прогнозирования нелинейного поведения материала, включая деформацию, развитие трещин и повреждений, где могут потребоваться сложные функции отображения [5]. Так, было проведено успешное

предсказание механики жидкости, используя нейронные сети, обученные уравнениям Навье-Стокса [5]. Потенциально аналогичный метод можно использовать для экономии вычислительных затрат при моделировании поведения материала.

Область мониторинга и управления на месте может извлечь выгоду из преимуществ методов глубокого обучения для использования при обнаружении ошибок и сбоях сборки. Сверточные нейронные сети, например, могут обнаруживать и классифицировать как макроскопические, так и микроскопические неточности, используя послойные данные оптических датчиков.

Несмотря на многочисленные преимущества применения методов ИИ в рассматриваемой области, на данный момент существует несколько серьезных ограничений, тормозящих их распространение в АП.

Первой проблемой можно назвать малоэффективность моделей МО при диагностике состояний, с которыми они ранее не сталкивались. Для АП это серьезная проблема, так как одним из его недостатков является непостоянство выходных свойств продукции, а в данный момент не существует доступных и обширных баз данных для процессов, продуктов и материалов, позволяющих закрыть все пробелы. Это ограничение делает упор на сбор данных для обучения путем создания сценариев, учитывающих широкий диапазон рабочих условий и пространственных измерений.

Вторым ограничением является отсутствие общей структуры данных и стандартных методов их интеграции. Несмотря на то, что каждый процесс создания изделия может генерировать большие объемы данных, например, может быть собрано до 0.5 Тб за построение, регистрироваться до 600 переменных и 75 Гб данных изображений в секунду, отсутствуют стандартные методы обработки, характеризующиеся большим объемом и скоростью в режиме реального времени [4].

Это приводит к проблемам доступности и качества данных. В настоящий момент, эффективность моделей МО во многом зависит от объема и качества данных, применяемых для их обучения. Обучающие данные в контексте параметров процесса АП можно получить численными или эмпирическими методами. Численные методы позволяют получать данные на основе моделирования процессов, а эмпирические данные получают путем проведения экспериментов. Однако, каждый из методов имеет свои недостатки. Моделирование представляется сложной задачей из-за большого числа параметров процесса и их вариативности, а получаемые результаты могут оказаться искаженными за счет неизбежного упрощения модели. С другой стороны, проведение многочисленных повторяющихся экспериментов по сбору данных может занимать много времени. Кроме того, данные, получаемые эмпирическим методом, могут иметь низкое качество, например, недостаточное качество и разрешение оптических данных при сборе в режиме реального времени. На данный момент исследования показали многообещающие результаты по сбору удовлетворительного количества обучающих данных с использованием высокопроизводительных методов АП, однако, это до сих пор остается нерешенной проблемой. Также необходимы дальнейшие исследования для объединения датчиков данных на месте. Объединение тепловых, акустических, оптических и других данных об окружающей среде сборки может создать более целостный, надежный и точный источник информации для обнаружения и исправления дефектов в реальном времени. Здесь также играют важную роль системы принятия решений.

Еще одна проблема связана с тем, что вычислительные затраты на обучение модели МО стремительно растут при увеличении количества задействованных в обучении переменных. А так как во всех процессах АП задействовано большое число различных параметров, вычислительные затраты, особенно на контроль в реальном времени, сильно возрастают. Решением данной проблемы видится использование методов извлечения признаков, которые сокращают проблемную область, например метод главных компонент (principal component analysis, PCA), а также стратегии расширенного обучения, например, байесовскую оптимизацию, чтобы сэкономить вычислительные затраты и время, необходимые для обучения и функционирования моделей МО [5].

Наконец, на развитие методов ИИ также негативно сказывается отсутствие библиотек для характеристик функций АП. Их разработка помогла бы облегчить выбор подходящего алгоритма, совместимого с доступными данными.

Таким образом, были приведены основные области применения современных методов ИИ в области АП. Основными тенденциями являются объединение нескольких методов и алгоритмов для решения конкретной задачи, а также стремление к разработке алгоритмов, способных работать с оптическими и акустическими данными в режиме реального времени. Также проанализированы основные ограничения, которые существуют в применении ИИ в области трехмерной печати. Среди них можно выделить проблему диагностики новых состояний, отсутствие стандартных методов интеграции и анализа данных, сложность получения качественных данных для обучения, проблему выбора метода для конкретной функции, большое число переменных процесса.

На сегодняшний день продолжают открываться новые возможности и области применения методов ИИ в АП. Однако, выявление новых возможностей в жизненном цикле АП является предвестником проблем с данными. Исходя из вышесказанного, основными задачами на данный момент являются: получение способа сбора актуальных и качественных данных, библиотек для характеристики функций АП, а также методов интеграции данных путем создания сценариев, учитывающих широкий диапазон рабочих условий и измерений.

Литература

1. Praveena B.A., Lokesh N., Abdulrajak Buradi, Santhosh N., Praveena B.L., Vignesh R.A. Comprehensive review of emerging additive manufacturing (3D printing technology): Methods, materials, applications, challenges, trends and future potential // *Materials Today: Proceedings*. – 2022. – Pp. 1–5.
2. Thompson M.K., Moroni G., Vaneker T., Fadel G., Campbell R.I., Gibson I., Bernard A., Schulz J., Graf P., Ahuja B., Martina F. Design for Additive Manufacturing: Trends, opportunities, considerations, and constraints // *CIRP Annals - Manufacturing Technology*. – 2016. – Vol.65. – Pp. 737–760.
3. Meng L., McWilliams B., Jarosinski W., Park H.-Y., Jung Y.-G., Lee J., Zhang J. Machine Learning in Additive Manufacturing: A Review // *JOM*. – 2020. – Vol.72. – Pp. 2363–2377.
4. Razvi S.S., Witherell P., Narayanan A. A Review of Machine Learning Applications in Additive Manufacturing // *Proceedings of the ASME 2019 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*. – 2019. – Pp. 1–11.
5. Zeqing Jin, Zhizhou Zhang, Kahraman Demir, Grace X. Gu Machine Learning for Advanced Additive Manufacturing // *Matter*. 2020. – Vol.3. – Pp. 1541–1556.

УДК 681.7.07

ИССЛЕДОВАНИЕ ВОЗМОЖНОСТИ ИСПОЛЬЗОВАНИЯ АМОРФНОГО ЖЕЛЕЗА В КАЧЕСТВЕ МАГНИТНОГО ЭКРАНА ВОГ

Климова Е.Н.^{1,2} (студент), Шевченко А.Н.²

Научный руководитель – канд. техн. наук Драницына Е.В.^{1,2}

¹Университет ИТМО

²АО «Концерн «ЦНИИ «Электроприбор»

e-mail: elizaveta5555@mail.ru

Работа выполнена в рамках темы НИР №620164 «Методы искусственного интеллекта для киберфизических систем».

В данной работе анализируются магнитные параметры аморфного сплава на основе железа. Экспериментально проверены коэффициенты экранирования магнитного экрана для волоконно-оптического гироскопа из пермаллоя и макета экрана из аморфного железа. Результаты коэффициентов верифицированы для нескольких вариантов конструкции. Проведен сравнительный анализ материалов. **Ключевые слова:** аморфное железо, коэффициент магнитного экранирования, магнитный экран, волоконно-оптический гироскоп, ВОГ, пермаллой.

Наиболее перспективными приборами для решения задач автономной навигации являются волоконно-оптические гироскопы (ВОГ), которые широко применяются для построения малогабаритных инерциальных навигационных систем среднего класса точности. Следует учесть, что в магнитном поле из-за эффекта Фарадея появляется оптический сдвиг нуля гироскопа. Путем введения магнитного экрана можно снизить погрешности измерений, обусловленные этим эффектом и таким образом улучшить точностные параметры гироскопа [1]. Для расширения областей применения ВОГ и улучшения их эксплуатационных характеристик требуется повышение коэффициента магнитного экранирования (КМЭ), которое может быть достигнуто путем оптимизации конструкции магнитных экранов. Добавление аморфного железа (АЖ) в конструкцию магнитного экрана ВОГ – это возможность увеличить КМЭ без существенного изменения конструкции.

Аморфные сплавы – новый особый класс прецизионных сплавов, отличающийся от кристаллических сплавов структурой, способом изготовления, свойствами и областью существования на температурно-временной диаграмме. На данный момент аморфные металлы приобретают все большую популярность и уже используются в следующих устройствах: системы телекоммуникаций, солнечные генераторы, электромагнитные экраны, электроизмерительная техника и др. [2].

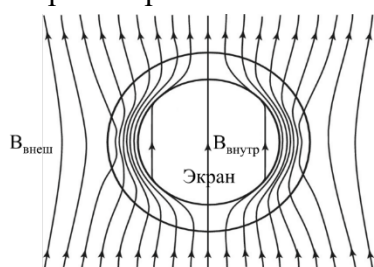


Рис. 1. Схема магнитного экранирования

Целью данной работы является исследование возможности использования АЖ в качестве магнитного экрана ВОГ для повышения КМЭ.

Методика определения коэффициента экранирования
Магнитное экранирование предназначено для уменьшения напряжённости постоянного или низкочастотного магнитного поля в какой-либо области пространства, а также для локализации магнитного поля в заданной области пространства (рис. 1) [3].

КМЭ представляет собой отношение индукций магнитных полей внутри и снаружи экрана [4].

$$K = \frac{B_{ext}}{B_{int}},$$

где B_{ext} – индукция магнитного поля, создаваемого вне экрана;

B_{int} – индукция магнитного поля внутри экрана.

Для определения напряженностей была сделана установка (рис. 2), включающая следующие устройства:

- катушку для задания постоянного поля с проводом ПЭТВ-2 сечением 3,0 мм, общее сопротивление 2,2 Ом;
- источник питания Rigol DP832, работающий в режиме стабилизации тока;
- немагнитную подставку высотой 420 мм;
- магнитометр НМС5983 (на эффекте Холла);
- персональный компьютер для получения показаний от магнитометра.

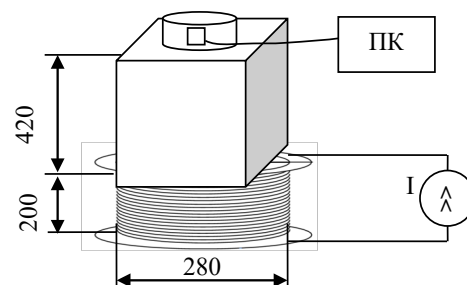


Рис. 2. Установка для испытаний

Экраны размещались на немагнитной подставке, а центр каждого экрана совмещался с осью симметрии катушки. Измерения напряженности магнитного поля в экранах производилось при подаче на катушку тока 2,5 А, 3 А и 0 А. Далее были рассчитаны приращение созданного катушкой магнитного поля (для исключения магнитного поля Земли и сдвига нуля магнитометра) и приращение остаточной намагниченности магнитного экрана. Для этого из показаний магнитометра при каждом значении тока в катушке были вычтены показания при нулевом токе. КМЭ для каждого из экранов рассчитывался как отношение приращения напряженности поля вне экрана к приращению напряженности поля внутри экрана. КМЭ рассчитывался для оси Y, расположенной вертикально.

Исследование экранирующих свойств АЖ

Для изучения экранирующих свойств АЖ были рассмотрены экраны с разным положением лент материала – вдоль и поперек линий магнитного поля (рис. 3).

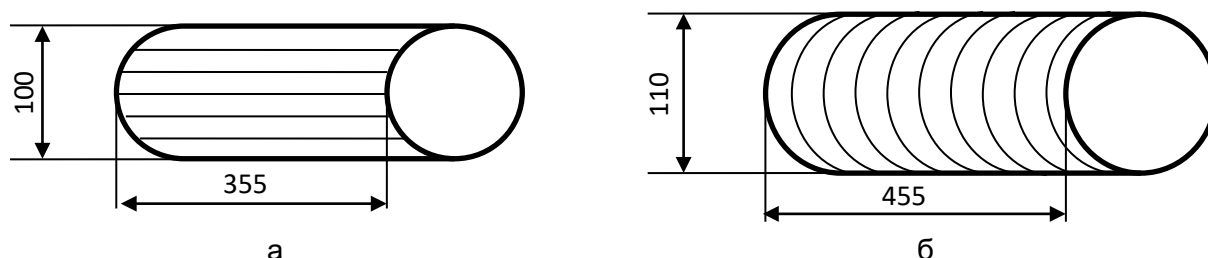


Рис. 3. Многослойные магнитные экраны: а) экран №1 – лист АЖ свернут спиралью продольно в 2 слоя; б) экран №2 – лист АЖ свернут спиралью поперечно в 3 слоя

Для изучения влияния многослойности экранов также был проведен эксперимент совместного использования экранов – экран №3 (экран №1 вставлен внутрь экрана №2). Перекрестное расположение экранов можно использовать в сложно ориентированных полях.

В таблице 1 представлены вычисленные по экспериментальным данным коэффициенты для трех вариантов экранов.

Таблица 1. КМЭ аморфного сплава на основе железа

Внешнее магнитное поле, мкТл	КМЭ экрана №1	КМЭ экрана №2	КМЭ экрана №3
от -50 до 150	7,8	28,3	78,1
180	3,2	10,4	38,3
200	2,7	8,6	27,4

Проведенные испытания показали, что направление намотки лент имеет значение – при их поперечном расположении магнитное поле лучше экранируется. Также получено, что в низкочастотных полях КМЭ выше, так как экраны имеют насыщение, сверх которого они не обеспечивают экранирующие свойства.

Конструкция магнитного экрана ВОГ

Магнитный экран ВОГ состоит из двух отдельных полых элементов, образующих при совмещении форму цилиндра, внутри которого размещается чувствительный элемент гироскопа (рис. 4). Его прочную фиксацию и магнитное экранирование обеспечивает соединение внахлест двух полуцилиндров, изготовленных из сплава железа и никеля (пермаллой 79НМ), который обладает высокой прочностью, коррозионной устойчивостью и пластичностью, а также имеет высокие показатели удельного сопротивления и магнитной проницаемости [5]. Таким образом, основные элементы, чувствительные к магнитному полю, надежно защищены как от механических, так и от магнитных воздействий в процессе эксплуатации.

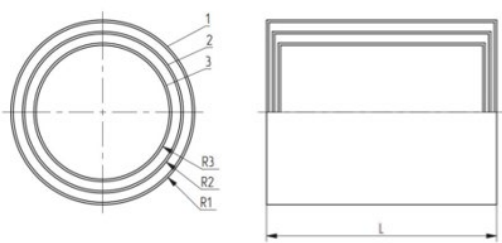


Рис. 4. Схема многослойного магнитного экрана ВОГ

Повысить КМЭ экрана из пермаллоя можно за счет дополнительного экрана из АЖ с внешней или внутренней стороны. Основной экран обеспечивает жесткость конструкции, а сортament АЖ ограничен (изготавливаются тонкие ленты, фольга, нити или порошок), поэтому конструкция из них не будет достаточно жесткой.

В предыдущем разделе показано, что в сильных магнитных полях коэффициент экранирования

снижается, следовательно, целесообразно разместить экран из АЖ внутри экрана из пермаллоя.

В процессе работы был изготовлен макет из отдельных склеенных между собой частей, образующих замкнутый контур, и произведено экспериментальное определение его коэффициента экранирования (рис. 5).

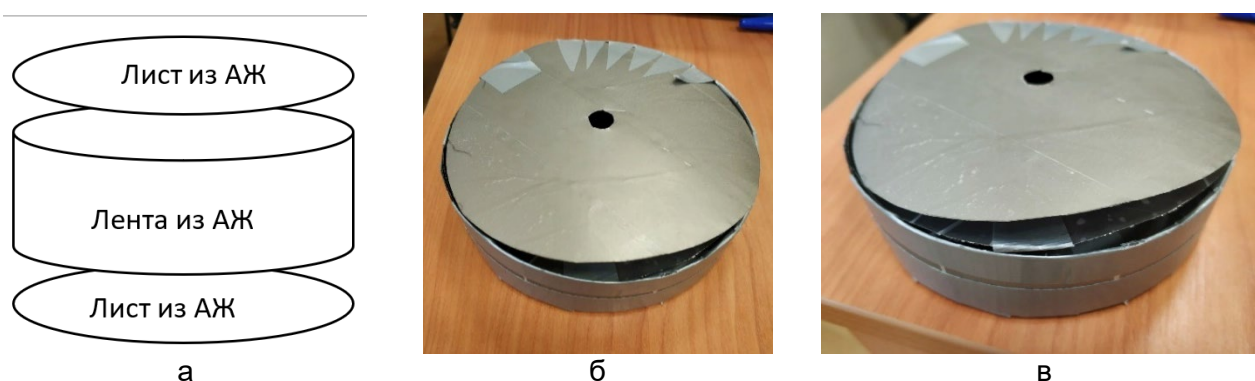


Рис. 5. Макет магнитного экрана из АЖ: а) схема макета экрана из АЖ; б) макет экрана из 1 слоя аморфного железа; в) макет экрана из 3 слоев аморфного железа

Экспериментальные исследования

Для сравнительного анализа КМЭ материалов были изучены однослойный и трехслойный экраны из пермаллоя, однослойный и трехслойный макеты экранов из АЖ, а также варианты конструкции однослойного и трехслойного экрана из пермаллоя с вкладкой трехслойного макета экрана из АЖ. Для каждой конструкции проводились испытания – определялись напряженности внешнего и внутреннего магнитных полей и коэффициент экранирования как отношение этих напряженностей.

В таблице 2 приведены вычисленные КМЭ для всех рассмотренных вариантов конструкций экранов.

Таблица 2. Сравнительная таблица КМЭ экрана из пермаллоя и макета экрана из АЖ

	Однослойный экран	Многослойный экран
КМЭ экрана из АЖ	2,1	2,95
КМЭ экрана из пермаллоя	34,72	147,06
КМЭ экрана из пермаллоя с вкладкой многослойного макета экрана из АЖ	38,46	125

В ходе данного исследования было выявлено, что макет экрана из АЖ имеет КМЭ от постоянных магнитных полей существенно ниже, чем у пермаллового экрана, что может быть связано с неидеальной конструкцией сделанного макета (экран должен плотно прикрываться и плотно соприкасаться). КМЭ слабо зависит от количества слоев АЖ. Поэтому доработка пермалловых экранов вставкой из аморфного сплава на основе железа нецелесообразна. В дальнейшем планируется исследовать свойства экранов из других аморфных сплавов (например, на основе кобальта) и изучить зависимость КМЭ от переменных магнитных полей.

Литература

1. Пат. RU 88797, МПК G01C19/72. Волоконно-оптический гироскоп. // Рамзаев А.П. (RU); правообладатель «Саратовский государственный технический университет» (СГТУ) (RU). – No. 2009126488/22; заявл. 10.07.2009; опубл. 20.11.2009.
2. Аморфные и нанокристаллические магнитомягкие сплавы. [Электронный ресурс]. – Режим доступа: <https://mstator.ru/ru/products/amorf> (дата обращения: 25.02.2022).
3. Перов Н.С. Магнитное экранирование. [Электронный ресурс]. – Режим доступа: <https://bigenc.ru> (дата обращения: 20.02.2022).
4. Donley E.A., Hodby E., Hollberg L., Kitching J. Demonstration of high-performance compact magnetic shields for chip-scale atomic devices // Review of Scientific Instruments 78, 083102. – 2007. DOI:10.1063/1.2767533.
5. Кулаченков Н.К., Шевченко А.Н., Безмен Г.В. Исследование магнитных экранов с использованием современных методов проектирования // Материалы XVIII конференции молодых ученых «Навигация и управление движением». – 2016. – С. 678–682.

УДК 620.178.151.2

ОБОСНОВАНИЕ ВЫБОРА МЕТОДА РЕГИСТРАЦИИ ПАРАМЕТРОВ ДВИЖЕНИЯ УДАРНИКА ПРИ ДИНАМИЧЕСКОМ ИНДЕНТИРОВАНИИ

Колганов О.А.¹ (аспирант), Федоров А.В.¹

Научный руководитель – д-р техн. наук, профессор Лукьянов Г.Н.¹

¹*Университет ИТМО*

e-mail: kolganoff2014@yandex.ru, afedor62@yandex.ru, gen-lukjanow@yandex.ru

Работа выполнена в рамках темы НИР №620164 «Методы искусственного интеллекта для киберфизических систем».

В данной работе рассматривается вопрос выбора метода регистрации параметров движения ударника при динамическом индентировании. Представлены основные требования к разработке системы регистрации движения индентора. Предложено использование магнитоиндукционного метода для регистрации движения ударника. Разработан прототип датчика динамического индентирования на основе магнитоиндукционного с двумя катушками (LVT).

Ключевые слова: динамическое индентирование, контактно-ударное взаимодействие, система регистрации, параметры движения, магнитоиндукционный метод, катушка, магнит.

Контроль физико-механических свойств материалов на различных стадиях жизненного цикла изделия является актуальной и востребованной задачей как на производстве, так и в эксплуатации. В настоящее время для решения данной задачи начинают широко применяться без образцовые методы и портативные средства. Применение этих методов и средств, там, где это возможно снижает конечную стоимость продукции и повышает оперативность контроля.

Одним из перспективных поверхностных методов без образцового контроля физико-механических свойств материала образца или изделия является метод динамического индентирования (ДИ).

Сущность метода ДИ заключается в регистрации параметров движения индентора в процессе его контактно ударного взаимодействия (КУВ) с испытуемым материалом. Основой приборной реализации метода является конструкция датчика для прибора ДИ, который состоит из следующих основных элементов (рис. 1, а):

1. Устройство сброса ударника.
2. Система регистрации параметров движения индентора.
3. Ударник.

Вышеперечисленные элементы датчика уникальны по своему функциональному назначению и состоят из различных деталей. Например, устройство сброса ударника предназначено для обеспечения гравитационного или принудительного разгона ударника. На рисунке 1.б показаны основные элементы устройства сброса ударника. Ударник же состоит из индентора, магнита и корпуса (рис. 1, в).

В настоящее время регистрацию параметров движения индентора выполняют на основе магнитоиндукционного метода. Применение этого метода регламентируется ГОСТ Р 56474-2015. В ходе КУВ движение магнита, жестко закреплённого с индентором, вызывает изменение магнитного потока, проходящего через катушку индуктивности. При этом, в последней наводится сигнал электродвижущей силы (ЭДС), пропорциональный скорости движения индентора на всем временном интервале его КУВ [1].

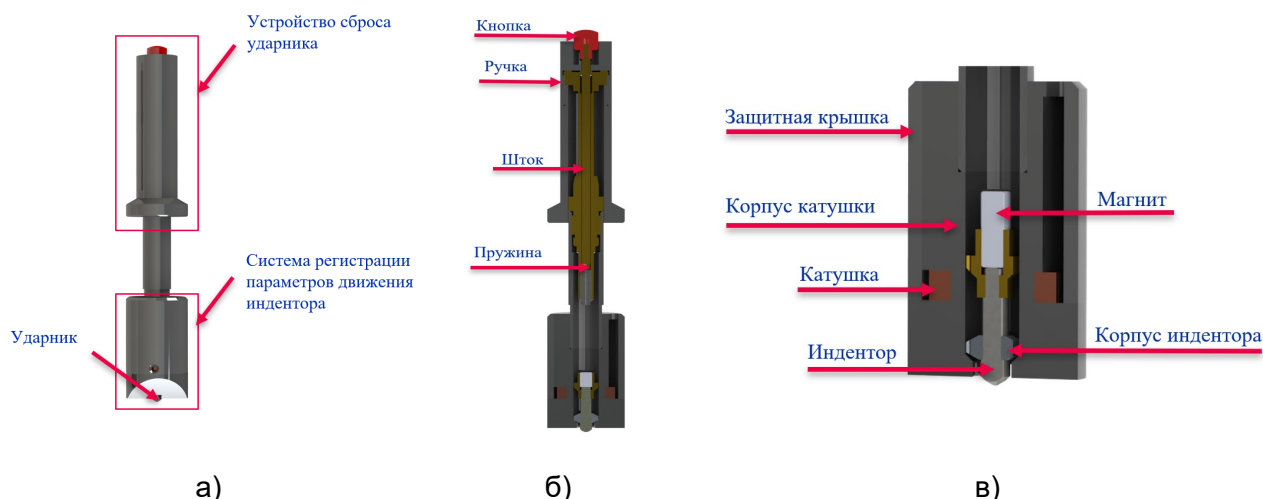


Рис. 1. 3D модель датчика для прибора ДИ: а) 3D модель датчика для прибора ДИ; б) 3D модель датчика для прибора ДИ в разрезе; в) 3D модель системы регистрации движения индентора

Главное достоинство преобразователей, которые основаны на магнитоиндукционном методе заключается в том, что их выходной сигнал ЭДС $\varepsilon(t)$ в допустимом приближении можно считать линейно зависимым от скорости внедрения индентора во время КУВ:

$$\varepsilon(t) = kV(t), \quad (1)$$

где k – коэффициент пропорциональности, который зависит от конструктивных параметров преобразователя [2].

На рисунке 2 показан пример полученного в эксперименте графика изменения ЭДС в катушке индуктивности. До момента начала КУВ на индентор действует только сила тяжести. С момента касания поверхности объекта контроля скорость индентора убывает от максимального значения до нуля, а затем меняя знак, поскольку меняется направление движения индентора, скорость достигает своего минимального значения называемого скоростью отскока.

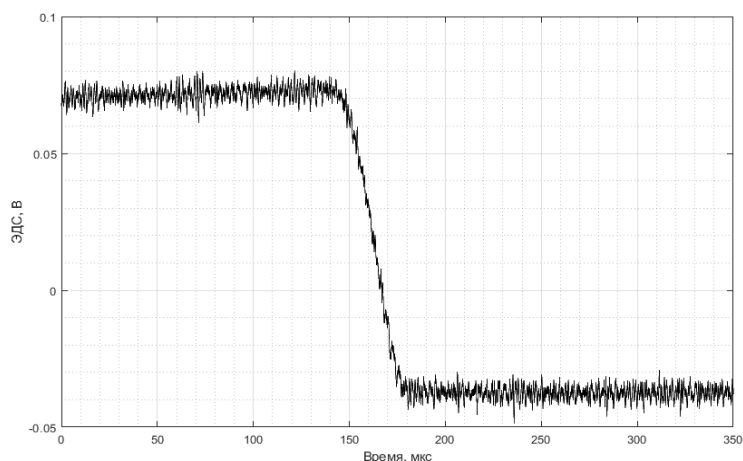


Рис. 2. График изменения ЭДС в катушке индуктивности

Наиболее важным элементом датчика ДИ является система регистрации параметров движения. Данное устройство обеспечивает получение первичной информации при КУВ, а именно выполняет регистрацию движения ударника.

Однако, следует отметить, что значения параметров первичного преобразователя и устройства его разгона в ГОСТ Р 56474–2015 не нормированы [3].

На сегодняшний день большинство исследований в области данного метода сосредоточено на решении вопросов перехода от значений характеристик КУВ индентора с испытываемым материалом, а именно массива данных, позволяющих построить диаграмму зависимости контактного усилия от глубины внедрения индентора, к значениям конкретных механических характеристик. При этом вопросы точности и достоверности получения первичной измерительной информации связанные с системой регистрации движения, не получили должной проработки.

Одним из главных требований при разработке системы регистрации движения индентора, чтобы на всем участке контактно-ударного взаимодействия (КУВ) наблюдалась линейная

зависимость скорости внедрения. Данное требование достигается конструктивными особенностями системы регистрации и электрическими параметрами катушки.

Несмотря на широкое применение магнитоиндукционного метода для регистрации параметров движения ударника, магнитоиндукционный метод с одной катушкой имеет недостатки, которые влияют на форму получаемого первичного сигнала. Одним из недостатков преобразователей такого типа является то, что количество линий потока, пересекающих катушку, меняются в зависимости от положения и выходное напряжение является функцией как скорости, так и положения. Вторым недостатком является то, что, когда ударник центрирован относительно катушки индуктивности, линии потока, которые обусловлены отрицательным полюсом, также могут индуцировать напряжение в катушке противоположной полярности.

Для устранения этих недостатков предлагается реализация магнитоиндукционного метода с использованием двух катушек. Катушка индуктивности здесь окружает магнит, который жестко закреплён с индентором, во время его движения. В рабочем диапазоне преобразователя оба конца магнита находятся внутри катушки, которая разделена на две секции. В каждой секции катушки расположен свой полюс магнита. Эти две катушки затем соединяются последовательно, чтобы получить выходное напряжение постоянного тока, пропорциональное скорости магнита (рис. 3) [4, 5].

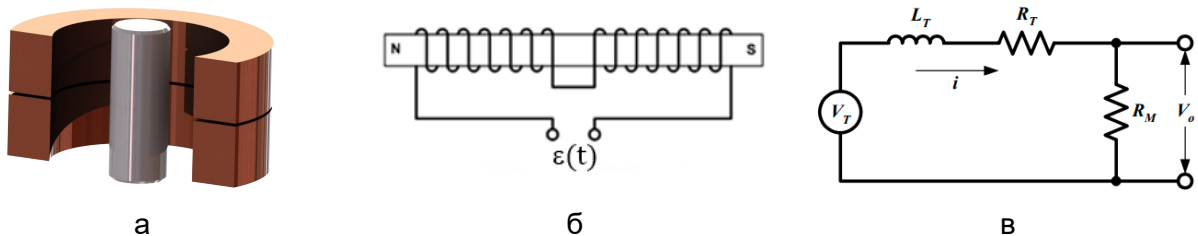


Рис. 3. Магнитоиндукционный метод с двумя катушками (LVT): а) 3D визуализация; б) схема устройства преобразователя линейной скорости; в) эквивалентная схема преобразователя линейной скорости



Рис. 4. 3D модель системы регистрации с двумя катушками и магнитом: а) без защитной крышки; б) без защитной крышки в разрезе; в) разрез системы регистрации с двумя катушками

Достоинством данного метода является, то, что при соединении катушек с противоположной полярностью генерируется напряжение пропорциональное скорости, и это напряжение не зависит от положения в ограниченном диапазоне:

$$L_T \frac{di}{dt} + (R_T + R_M)i = \varepsilon(t) = kV(t), \quad (2)$$

где $\varepsilon(t)$ – ЭДС,

$V(t)$ – скорость,

k – коэффициент пропорциональности,

i – ток, протекающий в цепи,

R_T – сопротивление,

R_M – сопротивление измеряемого прибора [6].

Для отработки регистрации движения ударника с помощью метода с двумя катушками индуктивности был разработан прототип датчика динамического индентирования (рис. 4). Прототип разработан модульным способом, что позволяет менять конструкцию датчика, использовать различные методы регистрации и изменять высоту падения индентора менять скорость движения.

Таким образом, был рассмотрен датчик для прибора динамического индентирования и выделены основные элементы конструкции. Предложено использование магнитоиндукционного метода с двумя катушками для регистрации движения ударника. Показана актуальность и применимость решения задачи оптимизации системы регистрации параметров движения индентора. Разработан прототип датчика динамического индентирования на основе магнитоиндукционного с двумя катушками (LVT).

Литература

1. Крень А.П., Рабцевич А.В., Зинкевич Н.В. Моделирование ударного контакта индетора с нежесткими стальными конструкциями при изменении твердости // Известия национальной академии наук Беларуси. Серия Физико-технических наук. – 2017. – №4. – С. 38–45.
2. Крень А.П., Рудницкий В.А., Делендик М.Н. Оценка погрешности определения физико-механических характеристик материалов при их контроле методом индентирования // Приборы и методы измерений. – 2018. – Т.9. – №3. – С. 263–271.
3. ГОСТ Р 56474-2015 Системы космические. Контроль неразрушающий физико-механических свойств материалов и покрытий космической техники методом динамического индентирования. Общие требования.
4. Medital. [Электронный ресурс]. – Режим доступа: <https://www.medital.com/products/lvt-linear-velocity-transducers>. – (дата обращения: 02.03.2022).
5. Electricaldesk. [Электронный ресурс]. – Режим доступа: <https://www.electricaldeck.com/2021/07/measurement-of-linear-velocity-using-velocity-transducers.html>. – (дата обращения: 02.03.2022).
6. Егоров Р.А. Разработка алгоритмического и программно-технического обеспечения первичной обработки при динамическом индентировании: диссерт. Канд. Техн. Наук. – СПб. – 2021. – 170 с.

УДК 004; 056.55

ГИПЕРАВТОМАТИЗАЦИЯ КАК СПОСОБ ПОВЫШЕНИЯ РЕНТАБЕЛЬНОСТИ ПРОИЗВОДСТВА

Колесникова Д.В.¹ (аспирант)

Научный руководитель – канд. техн. наук, доцент Юрьева Р.А.¹

¹Университет ИТМО

e-mail: kolesnikova_d@itmo.ru, raddayurieva@gmail.com

В работе проанализирован вопрос повышения рентабельности и экономической эффективности при автоматизации и гиперавтоматизации производств. Проведено сравнение понятий автоматизации и гиперавтоматизации в рамках производств, а также определены их преимущества. Рассмотрены влияния данных процессов на экономическую эффективность производства.

Ключевые слова: управление предприятием, оптимизация производственного процесса, управление, автоматизация, гиперавтоматизация.

Повышенное внимание к росту, цифровизации и операционному совершенству обратило внимание на потребность в более качественной автоматизации. Так проверка и автоматизация как можно большего количества процессов требует использования новых технологических инструментов и платформ, включая роботизированную автоматизацию процессов (RPA), платформы с малым количеством программного кода, инструменты анализа процессов, искусственный интеллект (ИИ) и виртуальных помощников. Такие инструменты считаются не зависящими от процессов, а это подразумевает, что их можно использовать в любой организации в различных сценариях.

Программное обеспечение наиболее востребовано как ключевой фактор автоматизации и гиперавтоматизации. Другие инструменты, используемые для автоматизации более специфических задач, такие как планирование ресурсов предприятия (Enterprise Resource Planning, ERP), системы управления цепочками поставок и управление отношениями с клиентами (Customer Relationship Management, CRM), также будет способствовать росту рынка.

Целью работы является исследования вопроса повышения рентабельности и экономической эффективности при автоматизации и гиперавтоматизации производств.

Задачи работы включают в себя исследование определений и значение автоматизации и гиперавтоматизации, а также определение экономического влияния данных процессов на работу предприятий и производств.

Практическая значимость работы заключается в исследовании вопроса оценки экономического влияния процессов автоматизации и гиперавтоматизации в рамках работы предприятия.

Чтобы определить понятие гиперавтоматизации необходимо дать определение «обычной» автоматизации.

Автоматизация – это процесс в развитии производства, при котором функции управления и контроля, ранее выполнявшиеся человеком, передаются приборам и автоматическим устройствам. Но, когда мы говорим об автоматизации, это также включает в себя автоматизацию, которая происходит внутри компьютеров, программ и систем управления. Эти два типа автоматизации служат разным целям, но предлагают одни и те же преимущества: чтобы процессы выполнялись быстрее, эффективнее и с большей точностью. Таким образом, на уровне бизнеса автоматизация способствует повышению производительности, снижению затрат и снижению рисков.

Гиперавтоматизация – это ориентированный на бизнес дисциплинированный подход, который организации используют для быстрой идентификации, проверки и автоматизации как можно большего количества бизнес- и IT-процессов.

Гиперавтоматизация связана с применением передовых технологий, включая роботизированную автоматизацию процессов, искусственный интеллект и машинное обучение (МО), для большей автоматизации процессов и расширения возможностей людей. Гиперавтоматизация распространяется на ряд инструментов, которые можно автоматизировать, но также влияет на те области, с которыми это делать сложно (например, обнаружение и анализ уязвимостей, отказов; проектирование и измерение деталей, мониторинг и переоценка производственного процесса). Другими словами, гиперавтоматизация – это расширение автоматизации; добавление нового уровня передовых технологий [1].

На рисунке представлена схема перехода от автоматизации до гиперавтоматизации.

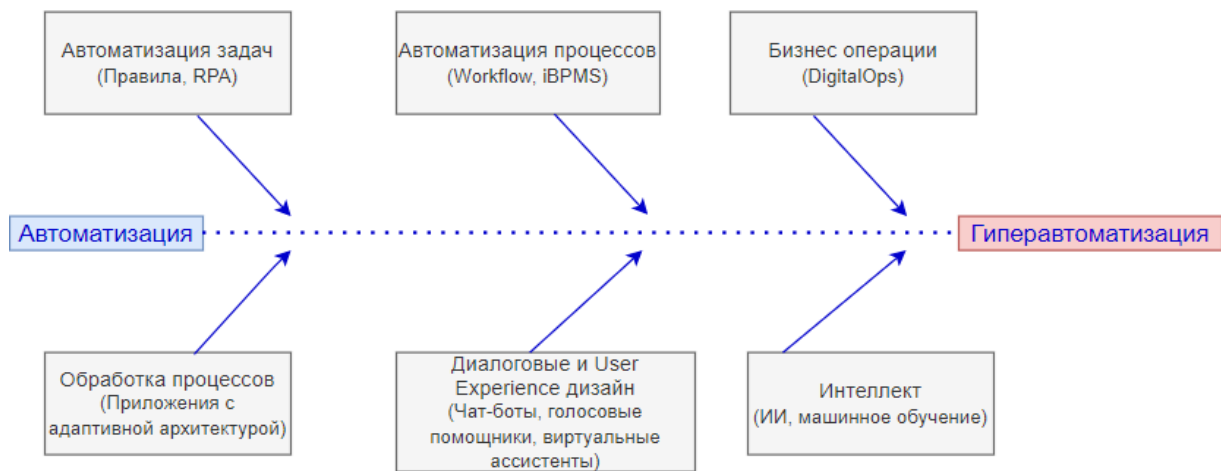


Рисунок. Переход от автоматизации до гиперавтоматизации

В таблице представлены сравнения по пяти ключевым параметрам для определения ключевых различий.

Таблица. Сравнение ключевых параметров автоматизации и гиперавтоматизации

Критерий	Автоматизация	Гиперавтоматизация
Технологии, необходимые для выполнения	Выполняется средствами автоматизации	Выполняется с помощью машинного обучения, пакетного программного обеспечения и инструментов автоматизации
Сложность технологий	RPA и автоматизация задач	Сложная автоматизация процессов на основе ИИ
Результат	Эффективные операции	Умные и эффективные операции
Область применения	Где уместно: «Какие процессы мы можем автоматизировать?»	Всеохватывающий: «Все, что можно автоматизировать, будет автоматизировано».
Сфера	Проводится с одной платформы	Это экосистема платформ, систем и технологий

К основным преимуществам гиперавтоматизации можно отнести [2]:

1. Сокращение времени выполнения задач – она поддерживает автоматизацию все более сложной работы, основанной на знаниях, поступающих от людей.
2. Развертывание цифровых работников – реализация взаимодействия Роботизированной автоматизации процессов с ИИ создает интеллектуальную цифровую рабочую силу, которая может выполнять повторяющиеся задачи для повышения производительности сотрудников. Эти цифровые работники являются агентами гиперавтоматизации, способными подключаться к различным бизнес-приложениям, работать со структурированными и неструктурированными данными, анализировать их и принимать решения, а также обнаруживать процессы и новые возможности автоматизации.

3. Увеличение прибыли – она помогает максимизировать прибыль за счет автоматизации бизнес-процессов и сокращения времени, затрачиваемого на низкоуровневые цифровые задачи и работу.
4. Сокращение расходов – она помогает сократить постоянные и переменные расходы, связанные с расходами на ИТ и телекоммуникации, такими как тарифы на мобильные данные, управление сетями, лицензии и облачные услуги. Кроме того, это влияет на администрирование информационных технологий, что уменьшает потребность в избыточных процессах или ручной работе, обслуживание которой обходится дорого.
5. Обеспечение соответствия – она может соответствовать нормативным требованиям и может использоваться в кибербезопасности, защите данных, управлении контрактами, утверждении счетов поставщиков, финансовых правилах и политиках хранения.

При рассмотрении экономической эффективности производства можно рассматривать несколько определений, выделить можно следующее: количественное соотношение двух величин, которые описывают результат эффекта деятельности и затраты. В качестве затрат принимаются производственные силы и степень их использования. Для определения экономической эффективности можно использовать критерии эффективности, которые устанавливаются для достижения определенных задач в производственном процессе (например, снижение затрат, материалов, ресурсов и другое) [3].

Основным критерием экономической эффективности можно выделить обеспечение максимальных результатов, которые достигаются при минимуме затрат, что определяется из качественно-количественного определения понятия. В качестве критериев также могут выступать различные экономические показатели, которые наглядно дают оценить влияния определенных действий или изменений. Например, рентабельность, прибыль, финансовая устойчивость, производительность и другое.

Рассмотрим непосредственно понятие рентабельности. Это относительный показатель экономической эффективности, который представляет собой степень эффективного использования различных ресурсов предприятия: деньги, люди, материалы. Для подсчета показателя необходимо составить отношение прибыли ко всем активам, ресурсам, потокам, которые ее формируют [4]. Рентабельность можно сравнить с коэффициентом полезного действия в физике, только при использовании финансовых данных.

Можно определить два основных пути повышения рентабельности: снижение издержек и повышение оборота. Однако, данные методы очень обширны и имеют разные ответвления, взаимосвязь и способы реализации. Повышение рентабельности на практике возможно при снижении себестоимости или расширении рынка сбыта продукции. На данные возможности влияет множество факторов, которые могут быть как внутри компании, так и внешними. Соответственно, это дает возможность рассматривать рентабельность различных сфер: активов, производства, продаж, капитала и другое.

Основным фактором производства, который влияет на рентабельность, является производительность труда. Данный параметр влияет как на себестоимость, так и на капитал, которым владеет предприятие. Если работник при фиксированной зарплате сможет изготавливать в несколько раз больше деталей, при этом затраты материалов вырастут соответственно, это все равно принесет компании сверхприбыль. На данный заработок не влияет даже то, что материалы или сырье возрастают пропорционально объему изготовленных деталей.

Следующим фактором можно выделить автоматизацию, механизацию и оптимизацию производственных и технологических процессов. Автоматизируя производство, управляющие структуры могут снижать вовлеченность персонала, что сказывается на затратах, а также исключает возможность ошибок и проблем, связанных с человеческим фактором. В современном производстве такие изменения позволяют увеличивать производственные силы при меньших трудовых затратах, однако, чаще всего требуется более квалифицированный персонал для управления новыми станками или программами.

Кроме того, при автоматизации может повышаться эффективность процессов, так как технологические процессы становятся качественными и обладают меньшим числом операций

(или проверочных, контрольных работ). На данных производствах повышается уровень автоматизации, что позволяет рассматривать гиперавтоматизацию как следующий шаг развития, а также влияет на экономические и технические показатели.

Также влияние оказывается на трудозатраты и реализацию самого производственного процесса, так как многие показатели, как экономические, так и хозяйственные, изменяются и оказывают влияние на работу в целом.

То есть можно отметить, что автоматизация и следующая за ней гиперавтоматизация позволяют повышать рентабельность производства за счет предлагаемых инструментов.

Оценка эффективности автоматизации может проводиться следующим образом [5]:

1. Собрать данные о производственных процессах, ресурсах, обобщить их.
2. Использовать экспертные системы или базы знаний для их анализа или оценки.
3. Определить возможности автоматизации: рассчитать необходимые затраты для внедрения, эксплуатацию, ремонт и утилизацию.
4. Оценить влияния внедряемых средств, произвести качественную и количественную оценки.
5. Определить эффект от изменений (экономические показатели).
6. Экономическую эффективность автоматизации технологических процессов производства можно анализировать в следующей последовательности.

В результате автоматизации и гиперавтоматизации производств экономическое влияние можно рассматривать в рамках материальных аспектов, так и в натуральном выражении. Далее представлена классификация источников экономического эффекта:

1. Трудовые ресурсы: сбалансированное использование трудовых сил, оптимизация рабочего процесса, времени, графика работы и простаивания, улучшение и укрепление состава.
2. Производственные ресурсы, оборудование: распределение загрузки и простоя, предупреждение наладки, ремонта и оценка состояния оборудования.
3. Производственный процесс: оптимизация производственных циклов, маршрутов, процессов, технологическая подготовка производства, а также оптимизация других средств, которые влияют на процесс.
4. Качество изготавливаемого продукта: оценка брака, дефектов, возможных потерь времени и ресурсов.

Кроме того, автоматизация циклов выставления счетов, контактов с клиентами и сбора платежей осуществляется с помощью гиперавтоматизации. Она также может вести записи, собирать данные и предоставлять полезную информацию для более точного планирования производственного процесса. Гиперавтоматизация также широко используется для обеспечения соблюдения важнейших правил устойчивости и прибыльности организации. Ее можно использовать для управления запасами и закупкой сырья, а также для планирования рабочих и других ресурсов. Применение гиперавтоматизации в секторе промышленности безгранично, а преимущества могут улучшить производственные процессы, связи с клиентами, а также работу сотрудников.

В заключении необходимо отметить, что в процессе автоматизации производств создается большая основа для качественного развития производства и его возможностей. Кроме того, это создает большие выгоды не только в рамках производственных активностей, но и в экономической сфере. Возможно, выгода часто выглядит косвенной, однако, в перспективе все внедрения окупаются и создают прочную базу для стабильного развития. Задачи выполняются быстрее, уменьшается время простоя или вероятность внештатной ситуации или сбоя, так как идет постоянный контроль оборудования, производства. Люди могут оперативно обрабатывать поступающую информацию с помощью систем поддержки решений или экспертных систем.

Однако, автоматизация может касаться не только производственных процессов, но и сфер управления, отчетности, бухгалтерии: автоматическое оформление отчетов, контроль расходов, принятие решений.

Важным свойством автоматизации можно считать то, что чем больше времени и сил потратить на внедрение, тем больше преимуществ можно получить. Так как стабильная структура и введение во множество областей – позволяет создавать единую базу для дальнейшей работы.

Литература

1. Haleem A. et al. Hyperautomation for the enhancement of automation in industries // *Sensors International*. – 2021. – Т.2. – С. 100–124.
2. Bornet P., Barkin I., Wirtz J. INTELLIGENT AUTOMATION: Welcome to the World of HYPERAUTOMATION: Learn How to Harness Artificial Intelligence to Boost Business & Make Our World More Human. – 2021. – Pp. 300–305.
3. Леонова Н.В. Теоретические основы экономической эффективности производства // *Современные организационно-экономические проблемы развития АПК: Материалы научно-практической конференции, посвященной*. – 2015. – С. 99–102.
4. Косенко Т.Г. Оценка факторов рентабельности производства // *Успехи современной науки*. – 2015. – №3. – С. 19–20.
5. Борщ Л.М., Герасимова С.В., Тюлин А.С. О вопросах трансформации экономики и модернизации технологических процессов в России // *Креативная экономика*. – 2018. – Т.12. – №6. – С. 717–732.

УДК 004.8

ОБЕСПЕЧЕНИЕ ИНФОРМИРОВАНИЯ ЖИТЕЛЕЙ ОБЛАСТИ О ХОДЕ РЕАЛИЗАЦИИ ПРОГРАММЫ БЛАГОУСТРОЙСТВА ДВОРОВЫХ ТЕРРИТОРИЙ ПРИ ПОМОЩИ ТЕХНОЛОГИЙ МАШИННОГО ОБУЧЕНИЯ

Кочарян А.М.¹ (студент), Мирошник А.А.², Хмелев А.П.¹ (студент), Москальчук В.Р.¹ (студент)

Научный руководитель – канд. физ.-мат. наук, доцент Еремина В.В.¹

¹ *Амурский Государственный Университет*

² *Автономная некоммерческая организация «Центр развития территорий»*

e-mail: armen0101017@gmail.com

В работе рассмотрено применение последних достижений машинного обучения и NLP в информировании граждан о ходе реализации проекта благоустройства дворовых территорий, а также преимущества данного подхода над традиционными способами взаимодействия с жителями в ходе реализации программ развития городской среды.

Ключевые слова: чат-бот, векторное представление, NLP, вопросно-ответная система, языковая модель.

В ноябре 2021 года все регионы Дальнего Востока начали участвовать в программе «1000 дворов», который должен ознаменовать переход от примитивных детских площадок к пространствам, актуальных для всех возрастных групп. В Амурской области в течение 2022 года должны будут реализованы 100 современных дворовых территорий. Важное условие при реализации программы – вовлечение жителей в работы над проектом и информирование их о ходе реализации проекта, а также освещение в СМИ. Чтобы отвечать на возникающие запросы граждан и принятие запросов о возможных нарушениях на строительных площадках от жителей, на базе АНО «ЦРТ» (Автономная некоммерческая организация «Центр развития территорий») было решено создать профильное направление – Ситуационный центр. В задачи Ситуационного центра входят:

1. Ответы на вопросы граждан касательно реализации программы в их дворе (в частности вопросы о сроках реализации программы, наполнении дворовой территории и т. д.).
2. Разъяснение о системе голосования и выбора территорий, участвующих в программе.
3. Сбор обратной связи о ходе реализации программы от жителей (для дальнейшей коррекции условий и сроков реализации программы).

В состав Ситуационного центра входят операторы, вопрос о содержании которых и их оснащения техникой и мебелью является наиболее острым. В последующие годы стоимость содержания таких специалистов будет только возрастать. Предполагается, что опыт внедрения бота может сократить расходы на содержание дополнительной штатной единицы и расширить охват жителей, которые могут обращаться с запросами в Ситуационный центр.

Для автоматизации процесса информирования граждан о ходе реализации программы благоустройства было решено использовать последние достижения в области машинного обучения и NLP. Простейшая схема функционирования системы представлена на рисунке 1. Вначале пользователю предлагается выбрать сценарий обращения: разговор с оператором, заявление о нарушении на строительных площадках, общение с виртуальным собеседником, в

котором предусмотрен блок выхода общения с ботом, если пользователь хочет продолжить разговор непосредственно с оператором. При выборе бота происходит переход к FAQ-системе, который является своего рода детектором намерений пользователя [2]. При этом из запроса пользователя выявляются и удаляются именованные сущности, такие как улица, дом, двор, и заполняется контекст разговора. Переработанный вопрос преобразуется в числовой вектор и находится наиболее близкий шаблон этого вопроса, что аналогично системе ранжирования.

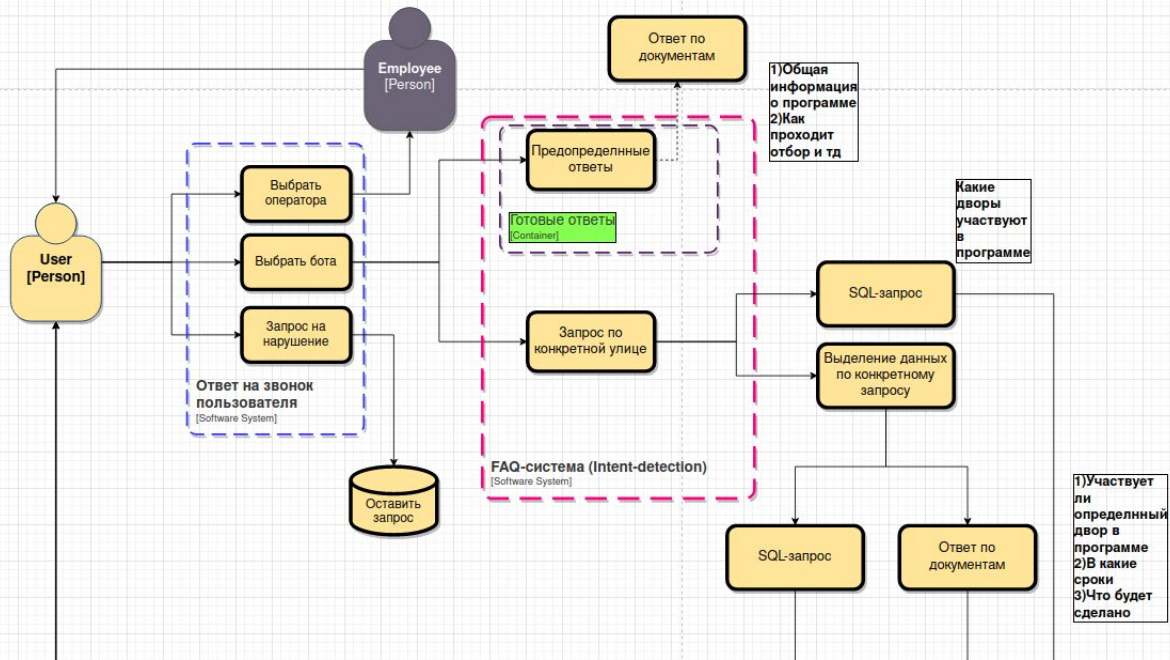


Рис. 1. Схема функционирования системы

Процесс преобразования документа в векторное представление может происходить с помощью меры $tf-idf$, отражающий важность использования каждого слова из коллекции документов:

$$tf-idf(t, d, D) = tf(t, d) \times idf(t, D) = \frac{n_t}{\sum_k n_k} \times \log \frac{|D|}{|\{d_i \in D | t \in d_i\}|},$$

где n_t – число вхождений слова t в документ,

$\sum_k n_k$ – общее число слов в документе,

$|D|$ – число документов в коллекции,

$|\{d_i \in D | t \in d_i\}|$ – число документов, в которых встречается t .

Также альтернативой может являться технология Word2Vec, основанная на моделях искусственных нейронных сетей и в случае архитектуры CBOW минимизирующей представленную ниже функцию ошибки, либо ее модификация FastText [3].

$$\mathcal{L} = -\log P(\omega_o | \omega_{c,1}, \omega_{c,2}, \dots, \omega_{c,C}) = -u_{j^*} + \log \sum_i e^{u_i},$$

где ω_o – целевое слово,

$\omega_{c,i}$ – контекстное слово,

$$u = W'^T W^T \frac{1}{C} \sum_{c=1}^C x^{(c)},$$

j^* – ожидаемая позиция слова,

x – входной вектор,

C – количество контекстных слов,

W', W – весовые матрицы.

В таком случае нахождение семантической близости слов становится простой задачей вычисления косинусного расстояния между двумя векторами, представляющими слова:

$$\text{similarity}(A, B) = \frac{A \cdot B}{\|A\| \|B\|} = \frac{\sum_i A_i B_i}{\sqrt{\sum_i A_i^2} \sqrt{\sum_i B_i^2}},$$

где A, B – вектора, представляющие слова.

Говоря абстрактно, вопросы пользователей делятся на вопросы общего характера с predefined ответами и на вопросы по конкретному двору. Второй случай более сложный и требует поиск в базе данных или нахождение ответа в документах организации. Так как вопросно-ответные системы на табличных данных на данный момент развиты плохо, было решено заготовить SQL-запросы для наиболее типовых вопросов, параметры которых заполняются по словарю контекста (например, сроки реализации программы на конкретной дворовой территории). В случае, если вопрос пользователя не найден в FAQ-системе, он передается блоку, который находит ответ в документах программы (примером может являться вопрос о наполняемости дворовой территории).

Это достигается с помощью архитектуры, состоящей из системы ранжирования и считывателя документов (рис. 2), который находит ответ на вопрос в документе и работает на основе продвинутой языковой модели BERT. Модель BERT в данном случае является наиболее предпочтительной благодаря идеям, заложенным в ее обучение: обучаясь, она параллельно выполняет задачу предсказания маскированных слов предложения и определения порядка следования предложений [1].

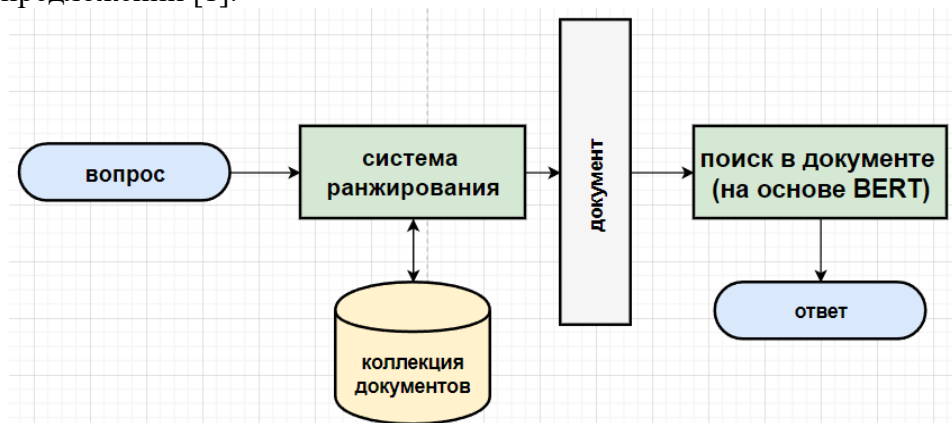


Рис. 2. Архитектура системы поиска ответа в документах

Успешный опыт реализации данного виртуального собеседника может дать стимул к расширению функций центра и подключению его к курированию других программ развития городской среды и взаимодействия с жителями, без необходимости повышать расходы. Помимо возможности отвечать в чате, планируется внедрение возможности отвечать людям по телефону, так как у категории лиц, для которой дворовая территория является важной частью их жизни - людей пожилого возраста, отсутствуют устройства коммуникации через интернет, и им важно позвонить и быть услышанными.

Литература

1. BERT, ELMO и Co в картинках (как в NLP пришло трансферное обучение). [Электронный ресурс]. – Режим доступа: <https://habr.com/ru/post/487358/> (дата обращения: 01.12.2021).
2. Simple intent recognition and question answering with DeepPavlov. [Электронный ресурс]. – Режим доступа: <https://medium.com/deeppavlov/simple-intent-recognition-and-question-answering-with-deeppavlov-c54ccf5339a9> (дата обращения: 03.12.2021).
3. Алгоритм обратного распространения ошибки на примере Word2Vec. [Электронный ресурс]. – Режим доступа: <https://habr.com/ru/post/496186/> (дата обращения: 03.12.2021).

УДК 004; 056.55

ИССЛЕДОВАНИЕ И РАЗРАБОТКА ОБЛАЧНОЙ АРХИТЕКТУРЫ РАСПРЕДЕЛЕННОГО ПРЕДПРИЯТИЯ

Кузнецова В.Е.¹ (студент)

Научный руководитель – канд. техн. наук Юрьева Р.А.¹

¹*Университет ИТМО*

e-mail: vlada0997@yandex.ru

В статье разработана имитационная модель предприятия по производству и монтажу печатных плат, архитектура которого основана на облачных вычислениях. Актуальность работы обусловлена повышенной уязвимостью PoT-устройств к информационным атакам и необходимостью повышения надежности производства печатных плат. Разработанная архитектура отличается от существующих внедрением контроля версионности, обеспечением целостности передаваемой информации посредством технологии блокчейн, а также позволяет собирать данные и контролировать состояние станков в режиме реального времени в облаке.

Ключевые слова: интернет вещей, блокчейн, архитектура, диаграмма потоков данных, имитационная модель.

Целью данной работы является исследование гипотезы о возможности совершенствования технологического процесса изготовления печатных плат путем использования облачной архитектуры предприятия, позволяющей собирать большие объемы данных о состоянии оборудования в режиме реального времени.

Для достижения поставленной цели были проведены обзор технологии блокчейн, а также анализ существующих архитектур PoT [1, 2]. На основе анализа была разработана архитектура предприятия по производству и монтажу печатных плат, отличающаяся от типовых архитектур внедрением технологии блокчейн, что увеличивает безопасность и скорость передачи данных, и обеспечением контроля версионности, что помогает отследить все изменения в файловых ресурсах предприятия в режиме реального времени (рис. 1).

Для изображения логики бизнес-процесса на распределенном предприятии была разработана диаграмма потоков данных [3], которая помогла наглядно отобразить потоки информации в облаке (рис. 2).

Внешние сущности – это источники и пункты доставки информации, которая приходит или уходит из системы. В данном случае это менеджеры, отдел бухгалтерии, инженеры, разработчики, служба безопасности.

Процессы — любые процессы, которые ведут к изменению информации и созданию выходных данных. Здесь, например, это оформление заказа, изготовление и монтаж печатных плат, проектирование печатных плат, защита данных и др.

Хранилища данных — файлы или репозитории, где хранится информация для последующего использования. В данном случае нашим репозиторием является облачное хранилище данных.

Потоки данных — маршруты, по которым информация перемещается между внешними сущностями, процессами и хранилищами данных. Они обозначены стрелками.

На основе разработанных архитектуры и диаграммы потоков данных была реализована модель взаимодействия датчиков в PoT в среде программного обеспечения AnyLogic. Самый нижний в нашей архитектуре агент — датчик. В нём происходит генерация сообщения, случайная выдача параметров и сбор статистических данных.

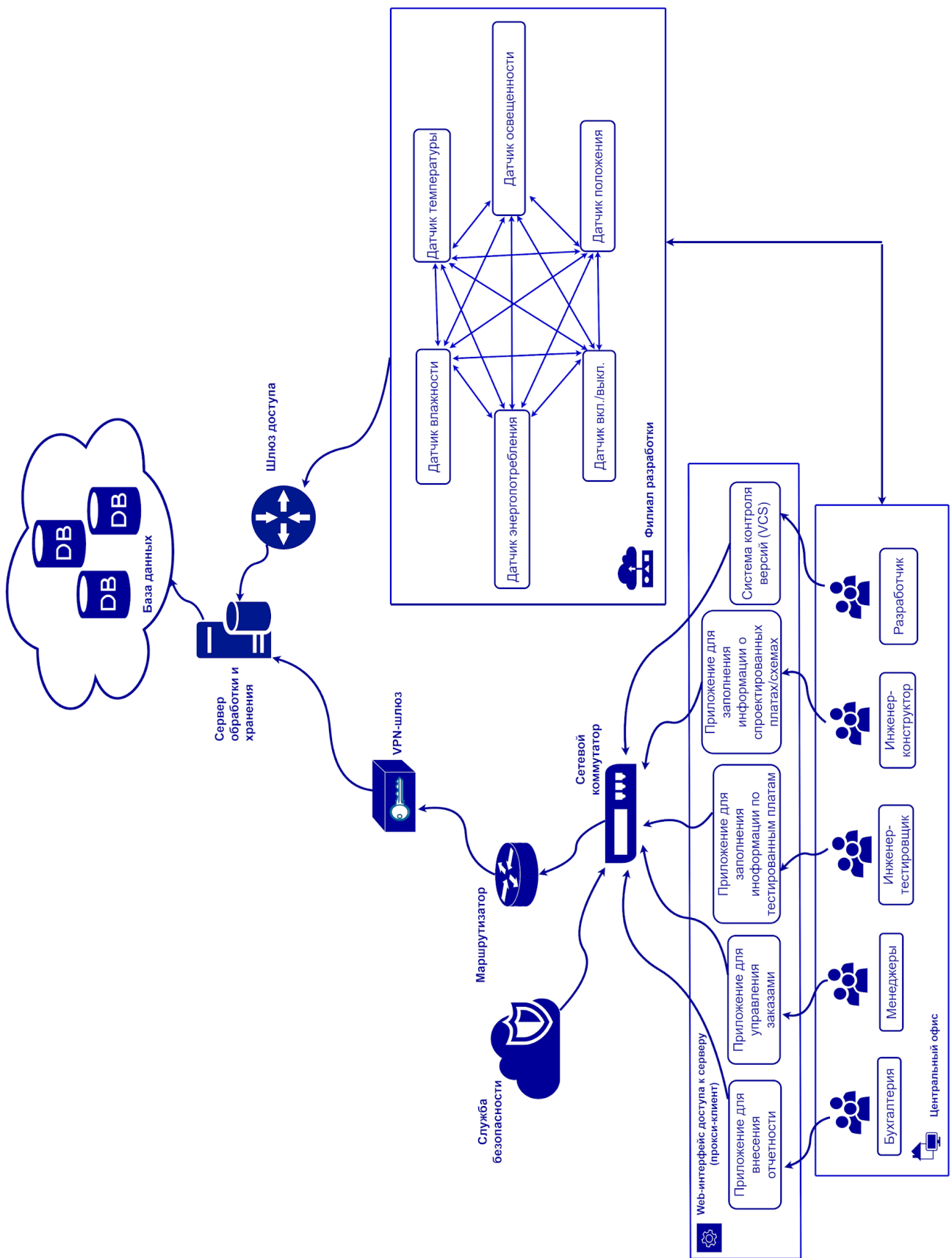


Рис. 1. Архитектура распределенного предприятия по производству и монтажу печатных плат

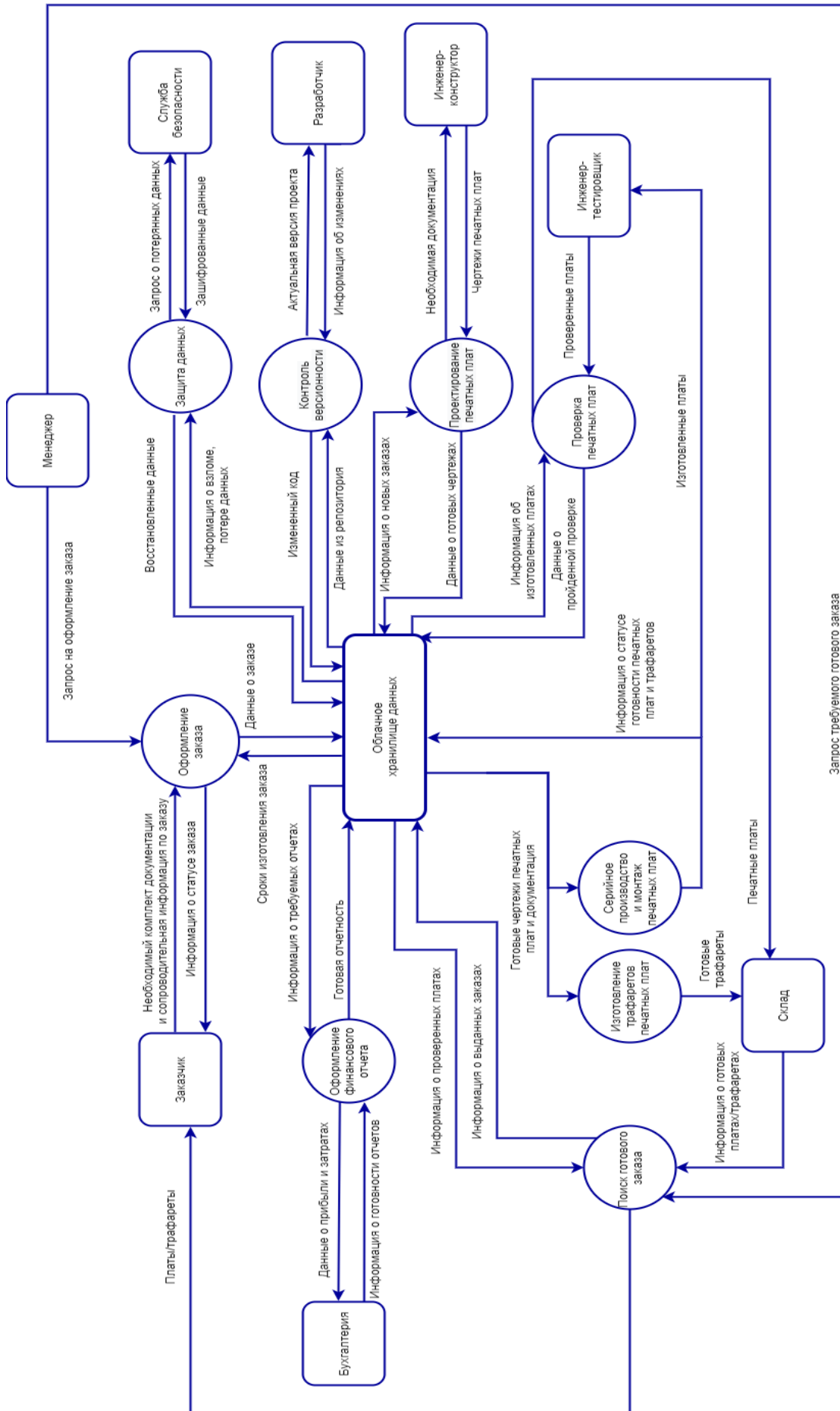


Рис. 2. Диаграмма потоков данных

Датчики будут общаться с друг другом, передавая сообщения. В данных сообщениях будет находиться информация: об отправителе, получателе, адресе получателя, а также полезная нагрузка. Для этого были заданы соответствующие параметры (рис. 3).

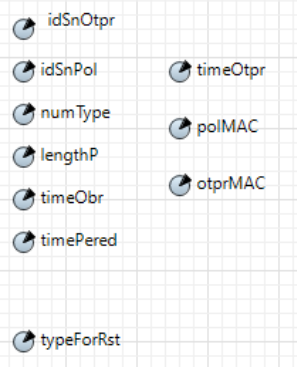


Рис. 3. Параметры сообщений

Параметры idSnOtptr (номер датчика отправителя) и idSnPol (номер датчика получателя) будут отвечать за адресацию. Параметр numType будет задавать тип сообщения, lengthP указывает длину сообщения в байтах, timeObr указывает время обработки сообщения, timePered — время передачи, timeOtptr — время отправления сообщения датчиком. Параметр typeForRest (является частью полезной нагрузки) будет задействован в сообщениях сброса, указывающих на то, что конкретное сообщение было получено, но вследствие перегруженности узла, часть текста отсутствует и невозможно правильно его обработать. Два последних параметра polMAC и otrpMAC указывают адрес подсети, в которой находится агент.

Отправка сообщений осуществляется через порт вых (рис. 4). Приём и обработка сообщений проходит через порт вх [4, 5, 6].

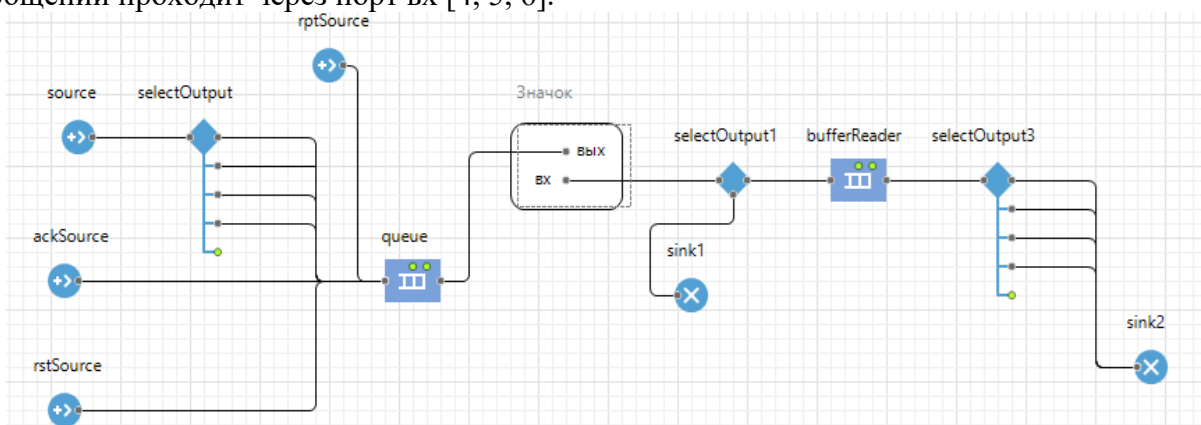


Рис. 4. Генерация и прием сообщений

Блок Source создает агентов. Используется в качестве начальной точки потока данных.

Блок Queue в нашей схеме используется для того, чтобы процессы (например, нанесение SMD-компонентов на плату, травление платы и т.д.) на производстве происходили не одновременно. Для этого необходимо задать определенное время, через которое произойдет переход на следующий этап процесса по изготовлению печатных плат.

Блок bufferReader позволяет построчно считывать информацию из базы данных на нашем производстве. Блоки SelectOutput участвуют в сборе статистики, а также генерируют параметры нового сообщения через Java код. Код будет исполняться в момент, когда сгенерированное сообщение будет проходить через блоки.

Блоки Sink уничтожают поступивших агентов, используются в качестве конечной точки потока агентов (завершают процесс).

Таким образом, реализованы 2 основных вида сообщений: Confirmable (подтверждаемое) и Non-confirmable (неподтверждаемое).

Для того чтобы получить ответ датчика получателя на Confirmable (Acknowledgement - подтверждение), создан блок Source (ackSource). А блок rstSource необходим для того, чтобы в ответ на сообщение сброса (Reset - сброс) выслать копию, когда буфер будет переполнен.

Далее была создана среда, в которой датчики обмениваются сообщениями, для этого необходим маршрутизатор. На рисунке 5 представлена модель маршрутизатора. Изображены входы датчиков на данный агент. Сообщение, поступающее на маршрутизатор, попадает в буфер и вычислительный комплекс. Сообщение может повредиться в случае, если вычислительных мощностей недостаточно (узел перегружен), либо если маршрутизатор вышел из строя.

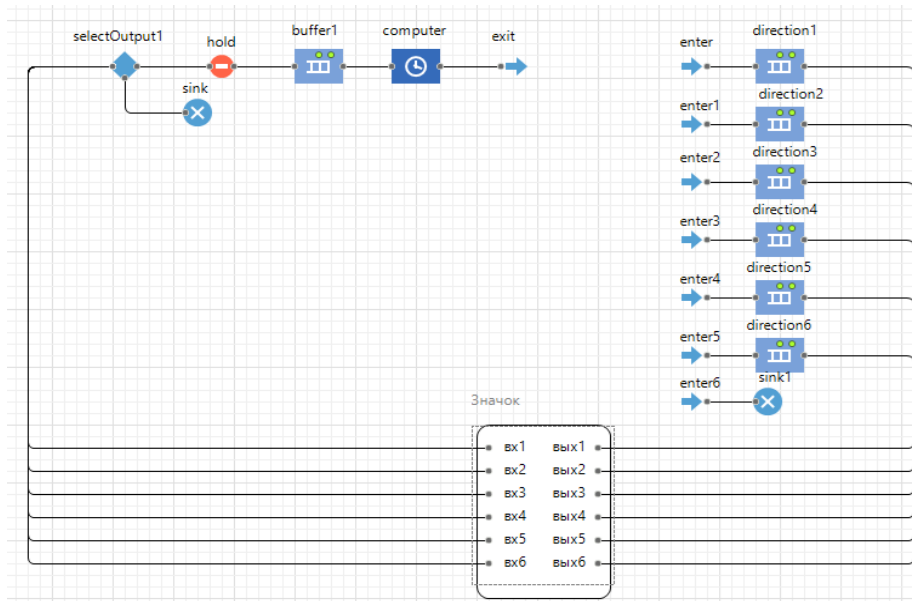


Рис. 5. Вычислительный сегмента маршрутизатора и система маршрутизации сообщений

Для перенаправления сообщений на нужный путь используются блоки Exit и Enter. Сообщение будет отсылаться на нужный путь в зависимости от идентификатора получателя (датчика). Для этого используются блоки Enter и Queue (direction 1-6). Блок Hold блокирует/снимает блокировку с потока агентов на определенном участке блок-схемы. Блок Computer задерживает агентов на заданный период времени. Блок Exit используется для перемещения агентов в другую часть модели.

Маршрутизатор может выходить из строя, чтобы показать данную особенность, была создана система имитации отказа - имитация перезагрузки роутера в случае сбоя (рис. 6).

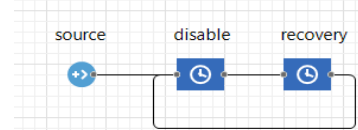


Рис. 6. Система имитации отказа

На рисунке 7 представлена модель взаимодействия датчиков в подсети через маршрутизатор.

Таким образом, мы подготовили датчики. Подсетей на предприятии несколько, поэтому был создан механизм перенаправления сообщений в другую подсеть.

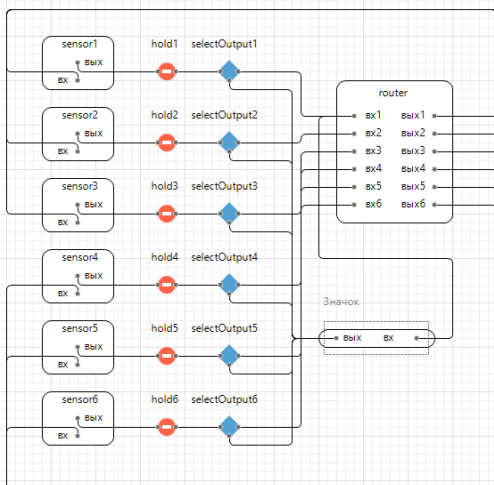


Рис. 7. Взаимодействие датчиков в подсети

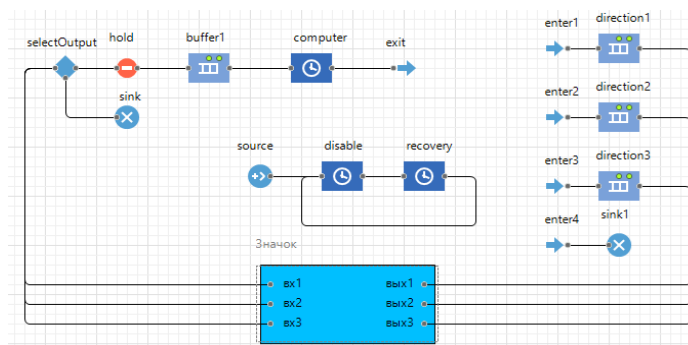


Рис. 8. Взаимодействие подсетей

Таким образом, мы получили работающую подсеть, далее был настроен процесс обмена сообщениями между подсетями, изображенный на рисунке 8.

Далее создан самый верхний сегмент нашей модели. В нём происходит обмен сообщениями между подсетями, а также посылка сообщений в Интернет. Конечный вид данного сегмента показан на рисунке 9.

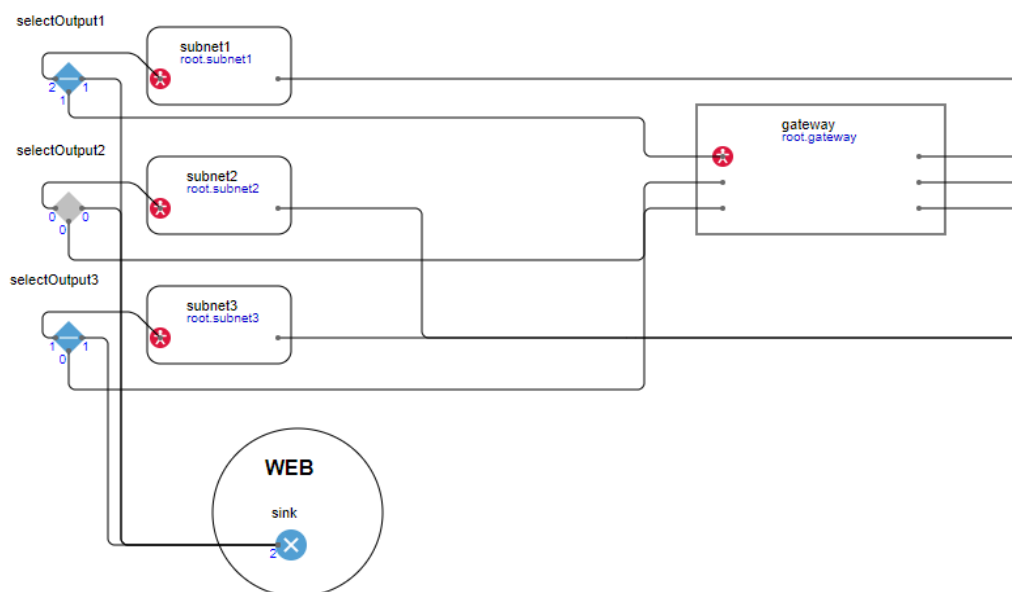


Рис. 9. Моделирование взаимодействия датчиков в среде IIoT

Таким образом, была разработана архитектура распределенного предприятия по производству и монтажу печатных плат, отличающаяся от существующих повышенным уровнем безопасности и обеспечением снижения количества брака в режиме реального времени.

Разработанная имитационная модель доказала гипотезу о целесообразности использования блокчейна в IIoT. Одной из самых важных особенностей данной разработки является возможность изготавливать платы и вести контроль процесса в режиме реального времени. Также благодаря электронному документообороту и блокчейну можно контролировать целостность чертежей и документации. В дальнейшем модель может быть использована в качестве цифрового двойника и для гиперавтоматизации производства печатных плат, для снижения количества бракованных изделий и повышения качества монтажа.

Литература

1. Перри Ли, Архитектура Интернета Вещей // Разработка архитектуры систем интернета вещей с применением датчиков, информационно-коммуникационной инфраструктуры, граничных вычислений, анализа и защиты данных. – М.: ДМК Пресс. – 2019. – 454 с.
2. Андреев Ю.С. Промышленный Интернет вещей: Учебное пособие / Ю.С. Андреев, С.Д. Третьяков – СПб: Университет ИТМО. – 2019. – 54 с.
3. Использование DFD: как описать движение данных в бизнес-процессах. [Электронный ресурс]. Режим доступа: <https://systems.education/data-flow-diagrams> (дата обращения: 09.03.2021).
4. AnyLogic Help. [Электронный ресурс]. – Режим доступа: <https://anylogic.help/anylogic/index.html> (дата обращения: 15.12.2021).
5. Григорьев И. AnyLogic за 3 дня // Практическое пособие по имитационному моделированию. – 2021. [Электронный ресурс]. – Режим доступа: [anylogic_in_three_days\(rus\).pdf](#) (дата обращения: 17.12.2021).
6. Юрьева Р.А. Основы моделирования умного производства в среде AnyLogic // Учебное пособие. – СПб: Университет ИТМО. – 2022. – 59 с.

УДК 004; 056.55

ОБЗОР И АНАЛИЗ СХЕМ РАЗДЕЛЕНИЯ СЕКРЕТА, ПОСТРОЕННЫХ НА РЕШЕТКАХ

Кустов Е.Ф.¹ (аспирант), Леевик А.Г.¹ (студент)

Научный руководитель – д-р техн. наук, профессор Беззатеев С.В.¹

¹Университет ИТМО

e-mail: elizarkustov@mail.ru, anton.leevik@gmail.com

Работа выполнена в рамках темы НИР №620164 «Методы искусственного интеллекта для киберфизических систем».

Работа частично финансировалась Федеральной программой академического лидерства Приоритет 2030.

Статья посвящена исследованию постквантовых схем разделения на решётках. Основная цель данной работы заключается в оценке их применимости на практике и выявлении недостатков. Оценка производилась в соответствии с критериями Шамира.

Ключевые слова: схемы разделения секрета, теория решёток, постквантовые алгоритмы, задача нахождения кратчайшего вектора, Обучение с ошибками.

Схемы разделения секрета используются в ситуациях, когда существует большая вероятность компрометации хранителей секрета, однако, вероятность сговора большинства участников принимается маловероятной.

Большинство современных схем разделения секрета основаны на задаче факторизации и дискретного логарифмирования. После изобретения Питером Шором квантового алгоритма возник вопрос о создании новых систем аутентификации устойчивых к атаке с использованием квантового компьютера. На данный момент компания ИВМ имеет в своём распоряжении квантовую систему, состоящую из 20 кубитов. Можно предположить, что развитие квантового компьютера до состояния, которое представляло бы угрозу распространённым сегодня алгоритмам, это вопрос ближайшего времени. Одним из возможных решений описанной проблемы может являться схема, основанная на решетках.

В данной работе будут рассмотрены две пороговые схемы и одна многоступенчатая и проведён их сравнительный анализ по критериям Шамира.

Схема разделения секрета — это способ распределения секрета среди всех пользователей, где все участники получают свою часть секрета (тень секрета). Восстановление секрета возможно только при объединении всех уникальных долей секрета пользователей.

Пороговая схема (t, n) , где n — это количество долей, на которые был разделён секрет и переданные пользователям, а t — это количество долей, которые нужны для восстановления общего секрета.

В многоступенчатой схеме участники должны делиться более чем одним секретом, и предоставить участникам только один общий доступ для восстановления всех секретов. Восстановленные секреты не дают никакой информации о раскрытых секретах.

Решетки часто используются при создании схем обмена секретами. Первая такая схема была представлена в статье [1], где используется задача CVP. Современные схемы обмена секретами в основном основаны на проблемах SVP и LWE, решение которых является NP-сложным.

Во-первых, давайте рассмотрим общее описание решеток. Обозначим решетку как L , которая является дискретной абелевой подгруппой R^m . Базис этой решетки обозначим как $B = [b_1, \dots, b_n] \in R^{m \times n}$. L состоит из $d \leq n$ линейно независимых векторов b_i , где

$$L(B) = \{ \sum_{i=1}^n t_i b_i \mid t_1, \dots, t_n \in Z \} \in R^m, \quad (1)$$

где $L(B)$ – множество всех целых комбинаций векторов в B .
 $\lambda(L(B))$ – минимальное расстояние решетки.

$$\lambda(L(B)) = \min\{\|x\| : x \in L(B)\}. \quad (2)$$

Одна из первых схем, основанных на решетках и задаче SVP (Shortest vector problem) была предложена в статье [2]. Это пороговая схема разделения секретов $(n1, n2)$, где $n1 = n2$, когда для восстановления секрета требуются все доли участников. Основная идея этой схемы состоит в том, чтобы восстановить базовую матрицу B , которая требуется для вычисления секретного вектора. Модификации этой схемы используются во многих работах.

Задача о кратчайшем векторе (SVP) – задача, когда нам нужно найти вектор x в $L(B)$ такой, что

$$\|x\| = \lambda(L(B)), \text{ где } \lambda(L(B)) \in R^n. \quad (3)$$

Рассмотрим схему не интерактивного проверяемого протокола обмена секретами на основе решетки, описанной в [2]. В этой схеме первое лицо, называемое дилером, выбирает несколько разных векторов длиной n и весом $m > 1$, затем создает секретные доли длины n , где $n > 0$. Чем больше вес, тем выше безопасность протокола, поскольку общий секрет представляет собой линейную комбинацию векторов.

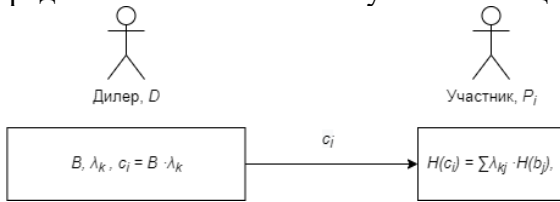


Рис. 1. Фаза распространения схемы на SVP

В фазе распространения секретными данными будут являться $B \in Z^{n \times n}, s = B * v$, а публичными $v \in Z^n, H(s), H(b_j), \lambda_k, 1 \leq j, k \leq n$, где v – случайный вектор, s – наш секрет, H – хэш-функция, b_j – строка матрицы B , λ_k – случайный вектор. Фаза распространения представлена на рисунке 1.

В фазе комбинации секретными данными будут являться c_n , а публичными $v \in Z^n, H(s), H(s), H(b_j), 1 \leq j, k \leq n$, где c_n – доли секрета участников. Фаза комбинации представлена на рисунке 2.

Безопасность этой схемы основана на используемой хэш-функции, которая скрывает параметры. В [2] авторы используют хэш-функцию Айтая, Любашевского и Микчанчио и Пейкерта и Розена.

В [3] впервые была представлена схема совместного использования секретов, основанная на задаче LWE (Learning with errors). Обучение с ошибками (LWE) - проблема заключается в задаче восстановления секрета $s \in Z_q^n$, заданного произвольным числом случайных линейных уравнений, каждое из которых может быть исправлено с точностью до заданной ошибки. Данная схема представляет собой (n, n) пороговую схему, поэтому для восстановления секрета все стороны должны участвовать со своими долями.

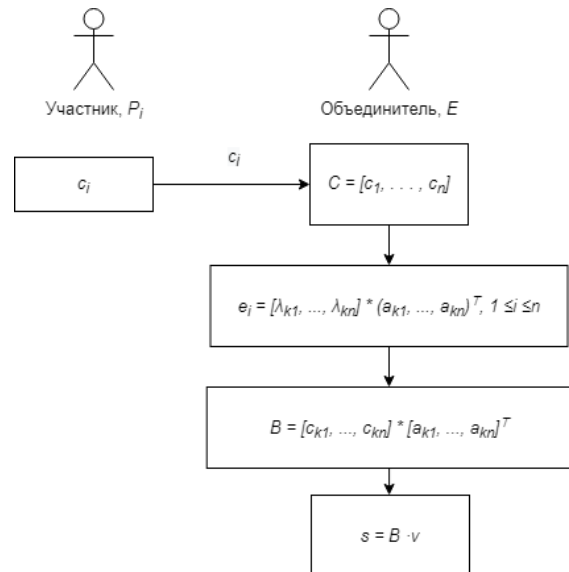


Рис. 2. Фаза комбинации схемы на SVP

В данной схеме секретными данными будут являться $s \in Z_q^n$, а публичными q, p, g генератор группы $GF(p), \chi \in Z^q$, где q, p – простые числа, s – наш секрет, χ – вероятность ошибки. Алгоритм схемы представлен на рисунках 3 и 4.

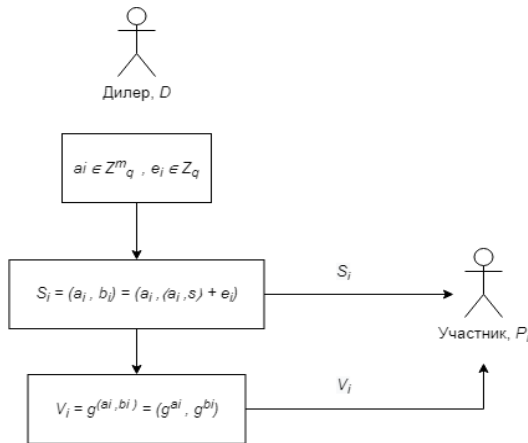


Рис. 3. Фаза проектирования

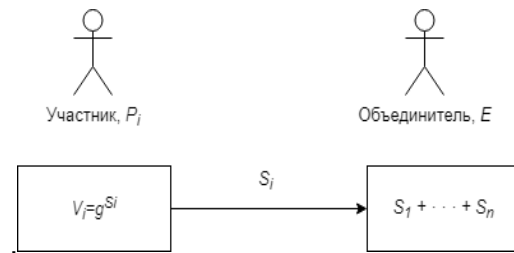


Рис. 4. Фаза проверки и восстановления

Пороговая схема многоступенчатая схема (MSSS) (t, n) представленная в [4], где t участников могут восстановить каждый из секретов. Эта схема позволяет участникам независимо восстановить любое количество секретов, и использование полученной информации для восстановления других долей секрета. В этой схеме у нас есть m секретов.

В данной схеме секретными данными будут являться $c_j \in \{0,1\}^r$, $B_i \in Z_q^{t \times t}$, $s_i \in Z^t$, а публичными $\lambda_j \in Z_q^t$, $A_i \in Z_q^{t \times r}$, $F \in Z_q^{t \times r}$, $H(s_i)$; $j = 1, \dots, n$, $i = 1, \dots, m$, где λ_j – случайный целочисленный вектор, s_i – наш секрет, c_j – часть секрета, B_i, A_i, F – матрица базисов и две вычисляемые матрицы, соответственно. Алгоритм схемы представлен на рисунках 5 и 6.

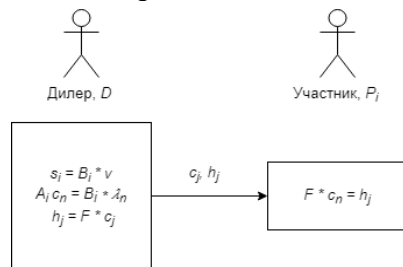


Рис. 5. Фаза распространения

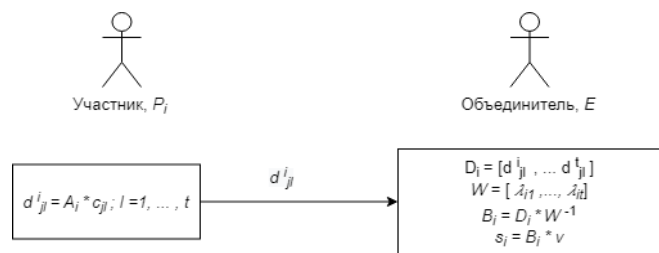


Рис. 6. Фаза комбинирования

Анализ рассмотренных схем был проведен с использованием критерий Шамира. Результаты представлены в таблице.

Таблица. Анализ схем разделения секрета

Критерии Шамира для схем разделения секрета	LWE	SVP	MSSS
Размер доли меньше или равен самому секрету	+	+	+
Возможность повторного использования секрета	-	-	-
Невозможность проанализировать секрет	+	+	+
Возможность добавления нового участника	+	+	+
Возможность обновить секрет	+	-	+
Возможность изменения веса долей	-	-	-

Как видно из результатов анализа ни одна из современных постквантовых схем разделения секрета на решетках не отвечает заявленным требованиям. Основная проблема всех схем на решётках заключается в том, что алгоритмы разделения секрета не предполагают его повторного использования и способов изменения веса долей секрета. Это происходит из-за того, что во всех схемах на решётках присутствует дилер для раздачи частей секрета и объединитель для их комбинирования, а значит, есть участник, которому известны все тени секрета.

Подводя итоги, можно сказать, что на данный момент постквантовые схемы разделения секрета на решетках не могут заменить их непостквантовые аналоги. Основная проблема заключается в наличии дилера и объединителя в данных схемах, что вынуждает продолжить исследования в данной области. Также можно сказать, что наиболее перспективными являются схемы, основанные на задаче LWE так, как они оказывают более высокий уровень безучастности.

Литература

1. Steinfeld R., Pieprzyk J., Wang H. Lattice-based threshold changeability for standard shamir secret-sharing schemes // IEEE transactions on information theory. – 2007. – Т.53. – №7. – С. 2542–2559.
2. Bansarkhani R.E., Meziani M. An efficient lattice-based secret sharing construction // IFIP International workshop on information security theory and practice. – Springer, Berlin, Heidelberg. – 2012. – С. 160–168.
3. Khorasgani H.A. et al. A lattice-based threshold secret sharing scheme // 2014 11th International ISC Conference on Information Security and Cryptology. – IEEE. – 2014. – С. 173–179.
4. Piliaram H., Eghlidos T. An efficient lattice based multi-stage secret sharing scheme // IEEE Transactions on Dependable and Secure Computing. – 2015. – Т.14. – №1. – С. 2–8.

УДК 004; 056

АНАЛИЗ СОВРЕМЕННЫХ СХЕМ ДОКАЗАТЕЛЬСТВА С НУЛЕВЫМ РАЗГЛАШЕНИЕМ, ОСНОВАННЫХ НА ТЕОРИИ РЕШЕТОК

Леевик А.Г.¹ (студент), Кустов Е.Ф.¹ (аспирант)

Научный руководитель – Давыдов В.В.¹

¹Университет ИТМО

e-mail: anton.leevik@gmail.com

Работа выполнена в рамках темы НИР №620164 «Методы искусственного интеллекта для киберфизических систем».

Работа частично финансировалась Федеральной программой академического лидерства Приоритет 2030.

В работе рассмотрены схемы доказательства с нулевым разглашением, построенные на алгоритмах теории решеток. Рассмотрены и проанализированы последние разработанные алгоритмы, проведено сравнение размеров доказательств при использовании различных схем.

Ключевые слова: криптография, постквантовая криптография, схемы доказательства с нулевым разглашением, теория решеток, криптографические протоколы.

Протокол доказательства с нулевым разглашением – это криптографический протокол, который позволяет одной стороне (доказывающему) подтвердить истинность утверждения другой стороне (верификатору), при этом не раскрывая никакой дополнительной информации об утверждении. Такое определение было дано в 1989 году в работе [1]. Простыми словами можно сказать, что данный протокол позволяет одной стороне, обладающей неким секретом, доказать другой стороне факт обладания этим секретом, не раскрывая никакой информации о секрете. Основными свойствами протокола являются:

1. Полнота (англ. Completeness). Данное свойство означает, что если доказательство, предоставляемое верификатору, корректно, то значит доказывающий обладает секретом.
2. Обоснованность (англ. Soundness). Свойство значит, что верификатора невозможно обмануть.
3. Нулевое знание (англ. Zero knowledge). Это означает, что верификатор не получает дополнительной информации о секрете.

На данный момент протоколы доказательства с нулевым разглашением нашли широкое применение в различных областях. Прежде всего они активно используются в блокчейн-системах и криптовалютах. Например, криптовалюта Zcash построена целиком и полностью на протоколе с нулевым разглашением zk-SNARK, благодаря которому соблюдается конфиденциальность транзакций, отправляемых пользователями. Также протоколы с нулевым разглашением могут использоваться в протоколах электронного голосования, электронных аукционов, систем аутентификации и так далее.

С появлением квантовых компьютеров и алгоритмов Шора и Гровера нынешние криптографические алгоритмы в будущем подвергнутся угрозе взлома. Алгоритмы, используемые в протоколах доказательства с нулевым разглашением, также являются неустойчивыми к атакам с использованием квантового компьютера. В связи с этим, важным направлением исследований является разработка таких алгоритмов, которые не могут быть взломаны с помощью квантового компьютера. Такие алгоритмы называются постквантовыми.

Одним из направлений постквантовой криптографии является криптография, основанная на теории решеток.

Решетка – это множество комбинаций линейно независимых векторов с целыми коэффициентами. Решетку может задать с помощью следующей формулы:

$$L = \left\{ \sum a_i \vec{b}_i : a_i \in Z, \vec{b}_i \in R^n \right\}.$$

Основными вычислительно сложными проблемами, на которых строятся криптосистемы, являются две проблемы:

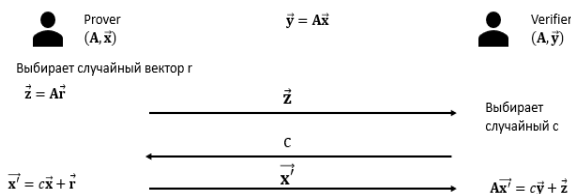
1. Проблема «короткого целочисленного решения» (англ. Short Integer Solution). Дана матрица $A \in Z_q^{n \times m}$ и вектор $\vec{t} \in Z_q^n$, где Z_q – поле целых чисел по модулю q . Требуется найти кратчайший вектор $\vec{s} \in Z_q^m$, такой что

$$A\vec{s} = \vec{t}.$$

2. Проблема «обучения с ошибками» (англ. Learning With Errors). Дана матрица $A \in Z_q^{n \times m}$, вектор $\vec{t} \in Z_q^n$ и задан вектор ошибок \vec{e} малого размера, обычно $\vec{e} \in \{-1, 0, 1\}^n$. Требуется найти вектор $\vec{s} \in Z_q^m$, такой что

$$A\vec{s} + \vec{e} = \vec{t}.$$

Данные проблемы признаны NP-сложными и не могут быть взломаны при помощи квантового компьютера.



Одной из первых схем доказательства с нулевым разглашением на решетках была схема Любашевского, опубликованная в 2009 году [2]. Данная схема построена на проблеме обучения с ошибками (LWE), схема действия данного алгоритма наглядно представлена на рисунке 1.

Рис. 1. Схема доказательства с нулевым разглашением Любашевского

Секретным сообщением, обладание которым хочет доказать «Prover», является вектор \vec{x} . Данный алгоритм является интерактивным и для улучшения его безопасности, то есть уменьшения вероятности успешного доказывания секрета при отсутствии секрета, данная процедура повторяется несколько раз, причем на каждом раунде алгоритма снова выбираются случайным образом вектор \vec{r} и константа c . Но поскольку вероятность обмануть протокол в одном раунде не является достаточно малой, то достижение необходимого уровня безопасности происходит за счет повторения алгоритма, что сильно увеличивает общий размер доказательства, что сильно влияет на производительность системы. С учетом этого факта были разработаны современные схемы доказательства с нулевым разглашением.

В 2019 году была разработана новая схема доказательства с нулевым разглашением [3]. В ней авторы сильно расширили и усложнили алгоритм, но главным отличием от оригинальной схемы стало использование решеток из элементов кольца многочленов по модулю q . Данные нововведения позволили уменьшить вероятность взлома схемы на каждом раунде. Полная схема протокола представлена на рисунке 2.

В 2020 году другой группой авторов предыдущая схема была переработана и уменьшен размер доказательства, создаваемого алгоритмом [4]. Для уменьшения размеров доказательства и ускорения работы алгоритма использовались дискретные преобразования Фурье в конечных полях (NTT). В одной из последних разработанных схем LNS21 [5] упор был сделан на формировании ключей с особой структурой и связанных с ними проверок, которые значительно уменьшают вероятность ошибки.

Для всех упомянутых схем был посчитан размер доказательства с учетом 128-битной безопасности, то есть вероятность обмана схемы равна 2^{-128} . В качестве сравнения был определен размер доказательства для непостквантовой схемы BulletProofs, полученные результаты представлены в таблице.

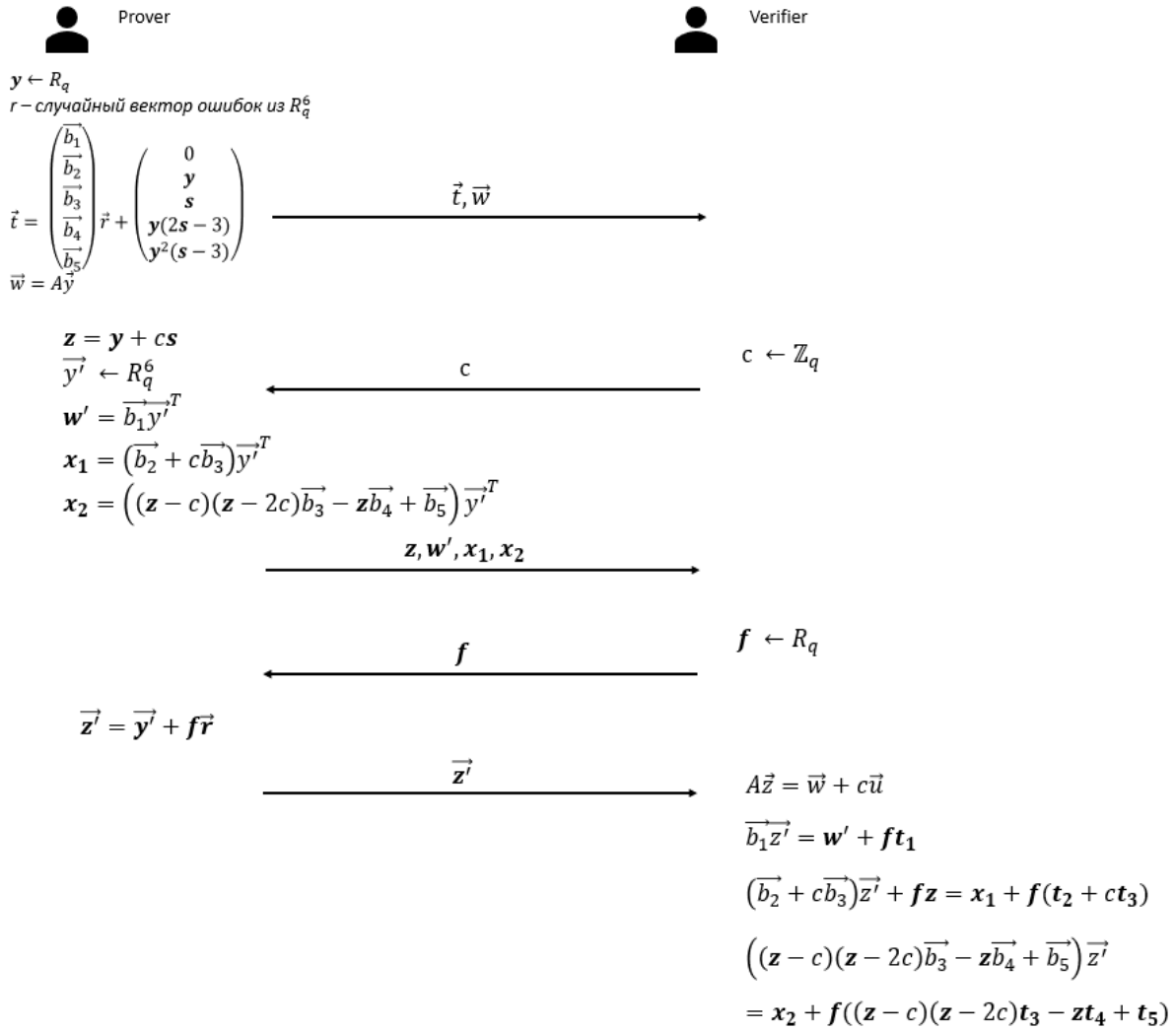


Рис. 2. Схема доказательства с нулевым разглашением BLS19

Таблица. Размеры доказательства протоколов с нулевым разглашением

Тип схемы	Название схемы	Размер доказательства, Кб
Постквантовая	BLS19 [3]	384
	ENS20 [4]	47
	LNS21 [5]	34
Непостквантовая	BulletProofs	2

Как видно из таблицы, размер доказательства постквантовых схем сильно превосходит размер доказательства непостквантовой, и на данном этапе применение постквантовых систем не является целесообразным. Однако, следует отметить, что прогресс в данном направлении есть, и за небольшое время удалось значительно снизить размеры доказательств.

Подводя итоги, можно сказать, что хотя на данный момент постквантовые схемы доказательства с нулевым разглашением не могут заменить их непостквантовые аналоги, был достигнут прогресс в области уменьшения размеров доказательств. Такие результаты были достигнуты путем уменьшения вероятности ошибки на каждом раунде алгоритма, что соответственно повлияло на уменьшение итераций выполнения алгоритма. Можно сделать вывод, что такой способ уменьшения доказательства является перспективным и может быть применен в дальнейших исследованиях.

Литература

- Goldwasser S., Micali S., Rackoff C. The knowledge complexity of interactive proof systems // SIAM Journal on computing. – 1989. – Т.18. – №1. – С. 186–208.

2. Lyubashevsky V. Fiat-Shamir with aborts: Applications to lattice and factoring-based signatures //International Conference on the Theory and Application of Cryptology and Information Security. – Springer, Berlin, Heidelberg. – 2009. – С. 598–616.
3. Bootle J., Lyubashevsky V., Seiler G. Algebraic techniques for short (er) exact lattice-based zero-knowledge proofs //Annual International Cryptology Conference. – Springer, Cham. – 2019. – С. 176–202.
4. Esgin M.F., Nguyen N.K., Seiler G. Practical exact proofs from lattices: New techniques to exploit fully-splitting rings //International Conference on the Theory and Application of Cryptology and Information Security. – Springer, Cham. – 2020. – С. 259–288.
5. Lyubashevsky V., Nguyen N.K., Seiler G. Shorter lattice-based zero-knowledge proofs via one-time commitments //IACR International Conference on Public-Key Cryptography. – Springer, Cham. – 2021. – С. 215–241.

УДК 004.7

КВАНТОВОЕ РАСПРЕДЕЛЕНИЕ КЛЮЧЕЙ: ОБЗОР РЕШАЕМЫХ ЗАДАЧ

Лихтенберг А.М.¹ (аспирант)

Научный руководитель – д-р техн. наук, доцент Беззатеев С.В.¹

¹Университет ИТМО

e-mail: bel.kavalini@yandex.ru

Работа выполнена в рамках темы НИР №620164 «Методы искусственного интеллекта для киберфизических систем».

Квантовая криптография – наука, основанная на принципах квантовой физики. Физические законы позволяют передавать информацию по квантовым каналам с возможностью детектирования злоумышленника, однако, законы физики также накладывают ограничения по передаче данных. В работе рассмотрены задачи, решаемые учёными в области квантового распределения ключей.

Ключевые слова: квантовая криптография, протоколы распределения ключей, информационная безопасность, архитектура сетей, квантовые каналы.

Открытие квантовых частиц стало катализатором к новому скачку развития технологий. Исследование скорости света и её сравнение со скоростью электрона в проводнике позволило создать принципиально новые способы связи: квантовые сети, передача информации в которых основана на световых импульсах и перемещении фотонов.

В настоящий момент учёные исследуют и развивают технологии, основанные на квантовых вычислениях. Новые технологии требуют новых способов защиты информации. Многие существующие криптографические алгоритмы перестают быть устойчивыми перед квантовым компьютером. В результате достижений последних нескольких лет в данной области была создана новая дисциплина – квантовая криптография.

Как и в любой новой дисциплине, в квантовой криптографии существуют нерешённые вопросы, такие как недостатки протоколов квантового распределения ключей, защита от атак на протоколы квантового распределения, ограничения дальности распространения квантового сигнала.

В данной статье рассмотрены основные направления развития квантовой криптографии, некоторые протоколы квантового распределения ключей, а также рассмотрена основная проблематика дальнейшей исследовательской работы.

Квантовая криптография: определение, основные направления развития

Квантовая криптография – это наука о защите коммуникаций, основанных на принципах квантовой физики. История квантовой криптографии начинается в 1970 году, когда учёными была предложена идея квантовых денег, основанная на принципе неопределённости Гейзенберга [1]. Согласно этому принципу, несколько квантовых состояний одной частицы не могут быть измерены с одинаковой точностью. Таким образом, опираясь на этот принцип, можно помешать злоумышленнику точно повторить так называемую «квантовую купюру». К сожалению, на практике данное решение было неосуществимо, так как «квантовому банку» приходилось бы долгое время пытаться сохранить состояние фотонов в «квантовой купюре». Однако, эта идея подтолкнула Чарльза Беннетта и Жиля Brassara к изобретению квантовой криптографии в 1984 году [2], когда опубликовали протокол BB84, в котором было предложено

передавать фотоны определённого состояния поляризации по одномодовому оптоволокну, которое позволяет сохранять это состояние. Впоследствии этот протокол был усовершенствован: появился протокол BB84 на фазовом кодировании. Также учёными были предложены и другие квантовые протоколы обмена ключом, почти все из которых основаны на кодировании квантового состояния одиночной частицы и базируются на принципе невозможности различить абсолютно надёжно два неортогональных квантовых состояния. Только один из предложенных протоколов, E91, основан на эффекте квантового запутывания.

На данный момент направления развития квантовой криптографии условно делят на четыре направления [3]:

- квантовые коммуникационные технологии, в которые входят протоколы обмена ключами между собеседниками;
- технологии квантовой обработки информации – к ним относятся шифрование информации на квантовых алгоритмах;
- технологии квантовых вычислений, в которых изучают работу квантового компьютера;
- постквантовая криптография – создание криптографических алгоритмов, устойчивых к квантовым алгоритмам Шора и Гровера.

Основное внимание в данной работе уделяется первому направлению: квантовым коммуникационным технологиям. К нему относят протоколы квантового распределения ключей (КРК) – это протоколы обмена ключами между недоверяющими друг другу собеседниками по недоверенному каналу с использованием квантовых явлений в канале.

Задачи квантовых коммуникационных технологий

Учёные, изучающие сферу квантовой криптографии, сталкиваются с множеством вопросов и нерешённых проблем в этой области [4]. Квантовая криптография – это молодая наука, достижения которой ещё не вышли на промышленный уровень.

К основным вопросам для изучения относят, например, следующие:

- атаки на системы КРК;
- ограниченная дальность канала;
- эффективное разветвление квантовых сетей.

Атаки на системы КРК

Как и в любых других системах, в системах квантовой криптографии существуют атаки на криптографические механизмы. Большинство из них основаны на физическом вторжении в систему, когда злоумышленник тем или иным образом нарушает работу системы: это может быть ослепление фотодетекторов или лавинных фотодиодов, а также отражение света от источника или приёмника с целью считывания передаваемой информации или перепутывания отправленных фотонов с добавленным злоумышленником.

На данный момент времени далеко не от всех атак на квантовые системы распределения ключей можно экранироваться или каким-то образом защититься: многие из атак можно только детектировать. Например, при использовании протокола BB84 после получения ключа выполняют также операции просеивания и операции проверки на стороннее детектирование, что приводит, как минимум, к потере времени, так как отправителю приходится отправлять в четыре раза больше фотонов, чтобы в результате получить ключ с необходимым количеством бит.

Ограниченная дальность канала

Учёные утверждают, что ограничение обмена ключом в волоконно-оптических линиях связи составляет сто километров [5]. Увеличение дальности квантового канала невозможно без потерь информации и дорогостоящего оборудования. Например, использование одномодового волокна, которое лучше помогает сохранить поляризацию вследствие невозможности отражения светового пучка от стенок канала, сильно уступает в скорости многомодовому волокну, в котором, в свою очередь, много шумов из-за явления дисперсии внутри волокна. Сварка волокна также приводит к некоторым потерям, а также изменение поляризации света с дальностью волокна и его возможными изгибами сильно увеличивается. Учёные

исследовательского центра Toshiba опубликовали статью «Twin-field quantum key distribution over a 511 km optical fibre linking two distant metropolitan areas» с предложением использовать так называемые поля-близнецы, что, по их словам, позволит увеличить передачу данных по оптоволокну в несколько раз. Добиваются таких рекордных дальностей взяв волокно с меньшими удельными потерями (на 1 км) и используя сверхпроводящие детекторы. В реализациях есть свои трудности: источники от А и Б должны быть заранее синхронизированы, а также непрерывный дрейф фазы в волокне должен компенсироваться, это очень непростая задача, и крупных практических реализаций на данный момент нет.

Эффективное разветвление квантовых сетей

Ограниченная дальность канала становится большой проблемой, если принимать во внимание тот факт, что системы КРК позволяют произвести обмен ключами только между двумя участниками. Это означает, что, если Алиса и Боб находятся на расстоянии большем друг от друга, чем позволяет квантовый канал, они не смогут обменяться ключом без какого-то посредника, что не всегда возможно. На данный момент этот вопрос предлагается решать несколькими способами:

- топология сети «звезда»;
- смешанные системы: ключи для обмена информацией защищаются на квантовых ключах и затем используются в обычной сети;
- обмен ключами через доверенный узел, чаще всего некий «чёрный ящик» без возможности доступа к нему извне для возможности исключения вторжений злоумышленников.

К сожалению, каждый из представленных способов имеет некоторые недостатки: топология «звезда» является дорогостоящей и неудобной для использования в квантовых сетях, к тому же, она не решает проблемы дальности канала.

Смешанные системы решают проблему дальности, к тому же они позволяют внедрить квантовые системы в промышленность с наименьшими потерями и затратами на оборудование и обучение сотрудников. Однако, вопрос, является ли такая система отвечающей определению квантовой сети, остаётся открытым.

Обмен ключами через доверенный узел кажется наиболее приемлемым среди всех представленных решений, однако, он не исключает возможные атаки на этот доверенный узел или его возможную подмену, что перестаёт делать систему полностью доверенной.

Квантовая криптография – новая, стремительно развивающаяся наука, перспективы развития которой позволят человеку достичь новых высот в сфере квантовой передачи информации наиболее безопасным путём.

Однако, в этой науке присутствует также много нерешенных вопросов, над которыми сейчас работают учёные: например, дальность квантового канала ограничена, не все атаки в канале возможно предотвратить или прекратить, а обмен ключами доступен только между двумя участниками информационного обмена.

Кроме того, в отличие от типичных условий в задачах криптографии, квантовую криптографию невозможно рассматривать без учета особенностей физической реализации, так как квантовая криптография базируется на законах квантовой физики и её стойкость опирается на некоторые постулаты квантовой механики, например, на запрет клонирования состояния фотона.

Все эти факторы делают данную науку актуальной и востребованной в обществе, ведь человек нуждается в наиболее быстрых и безопасных способах обмена информацией.

Литература

1. Слепов Н. Квантовая криптография: передача квантового ключа. Проблемы и решения. [Электронный ресурс]. – Режим доступа: <https://www.electronics.ru/journal/article/705> (дата обращения: 06.03.2022).
2. Курочкин Ю. Квантовая криптография. Как получить шифр, который невозможно взломать (быстро). [Электронный ресурс]. – Режим доступа: <https://postnauka.ru/longreads/156066> (дата обращения: 06.03.2022).

3. Корольков А. О современном этапе развития прикладной квантовой криптографии. [Электронный ресурс]. – Режим доступа: <http://lib.itsec.ru/articles2/crypto/o-sovremennom-etape-razvitiya-prikladnoy-kvantovoy-kriptografii> (дата обращения: 06.03.2022).
4. Пространственные электрооптические модуляторы: лабораторные заметки. [Электронный ресурс]. – Режим доступа: https://in-science.ru/library/article_post/prostranstvennye-elektroopticheskie-modulyatory-laboratornye-zametki (дата обращения: 06.03.2022).
5. Gisin Nicolas, Grégoire Ribordy, Wolfgang Tittel, and Hugo Zbinden. «Quantum cryptography» // *Reviews of modern physics*. – 2002. – №74. – С. 145–196.

УДК 004.942

ИССЛЕДОВАНИЕ ВЛИЯНИЯ HUMAN FACTORS НА СКОРОСТЬ ДВИЖЕНИЯ РЕЛЬСОВОГО ГОРОДСКОГО ТРАНСПОРТА

Ляховенко Ю.А.¹ (студент)

Научный руководитель – канд. техн. наук Попов И.Ю.¹

¹Университет ИТМО

e-mail: lyakhovenko.kat@gmail.com

Работа выполнена в рамках проекта 2019-0898 «Многоуровневое управление сложными техническими системами».

Представлена модель оценки состояния транспортных узлов движения общественного рельсового транспорта. Исследована зависимость скорости движения рельсового городского транспорта от влияния внешних случайных факторов, зависящих от человека (human factors). Учтены такие факторы, как движение других транспортных средств и пешеходов, ремонтные работы на участках дорог, а также плотность потока транспортных средств и пешеходов. Полученный подход может быть использован при проектировании транспортных узлов беспрепятственного движения беспилотного рельсового городского транспорта.

Ключевые слова: мультиагентные системы, моделирование, железнодорожный транспорт, human factors, моделирование транспортных потоков.

Влияние человека способствует большому количеству инцидентов, которые в том числе происходят в железнодорожной системе [1]. Оперативное выявление и устранение человеческих факторов, а также сдерживание последствий, возникающих в результате их возникновения, позволяют сделать железнодорожный транспорт безопаснее [2]. Первый шаг в процессе обеспечения безопасности – выявление факторов, которые часто приводят к возникновению инцидентов, что может позволить разработать соответствующие стратегии предотвращения и/или смягчения последствий. Существует широкая область, охватывающая бдительность и восприятие машинистов поездов, их распознавание и действия в соответствии со знаками и сигналами. Это включает также исследования дизайна систем указателей и сигнализации. Из всех вопросов человеческого фактора на железнодорожном транспорте эта область была наиболее изучена на протяжении многих лет [3, 4]. Одна из причин данного изучения – связь с сильными исследовательскими программами, охватывающими те же темы, что и вождение автомобиля. В значительной степени связанный с подходом к общей системной эргономике, человеческий фактор на железнодорожном транспорте также дал дальнейшее развитие понятию и практике интеграции человеческого фактора при разработке систем. Достижения в этой области часто скрываются в документации по коммерческим проектам [5].

Скорость движения трамвая в контексте представленной работы зависит от следующих факторов: от других участников дорожного движения (транспортных средств (ТС) и пешеходов (пассажиры), нуждающихся в посадке/высадке в/из трамвая); ремонтные работы и уровень загруженности участка дорожного движения.

Трамвай описывается следующим набором атрибутов параметров движения: координаты и направление; средняя скорость на участке; наличие остановки на участке.

Разрабатываемая для исследования модель имитирует процессы дорожного движения трамвая с влиянием на него всех перечисленных выше факторов.

Трамваи могут двигаться только по рельсам (на дороге или выделенных полосах), останавливаясь на перекрестках при запрещающем сигнале светофора, а также при высадке и посадке пассажиров.

Автомобили могут двигаться по любым участкам дороги, кроме тротуаров и выделенных полос, уступая дорогу трамваям.

На рисунке 1 представлена блок-схема алгоритма движения трамвая в условиях влияния human factors.

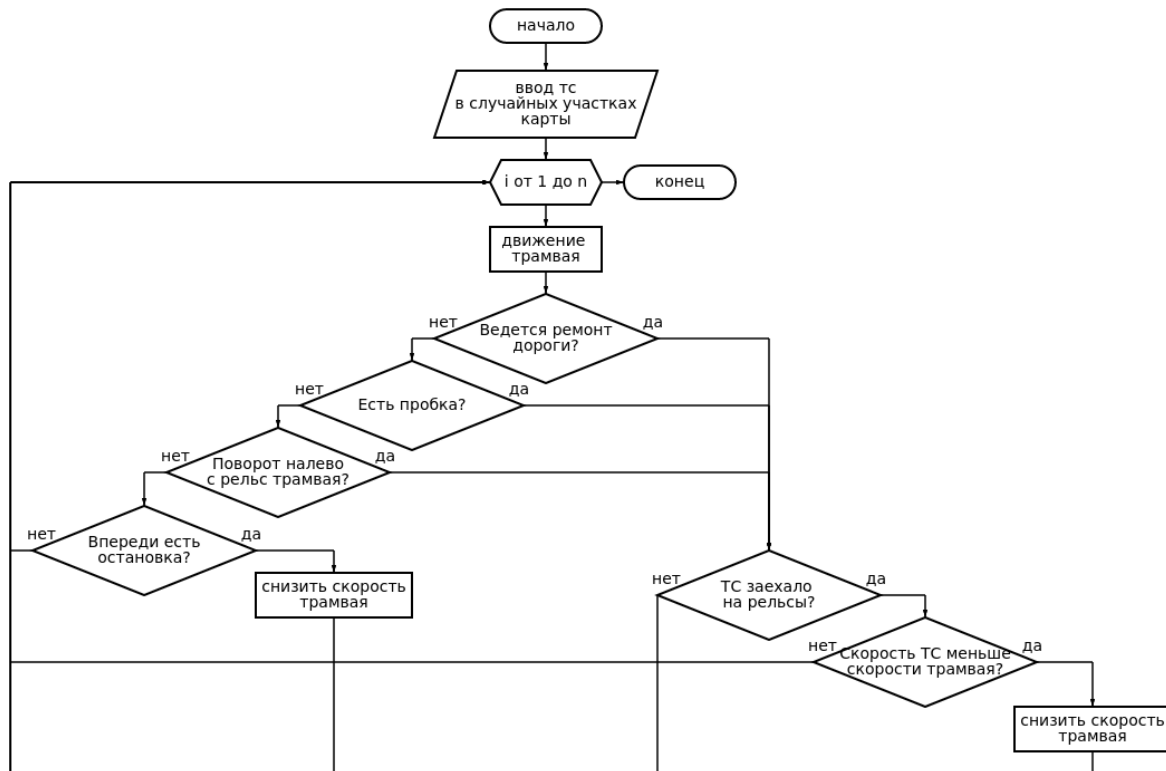


Рис. 1. Блок-схема алгоритма движения трамвая в условиях влияния human factors, где i – индекс итерации цикла движения

Система дорожного движения с участием трамваев, ТС и пешеходов (далее – система) описана с помощью мультиагентного подхода. Предлагаемая система рассматривает процесс дорожного движения как набор временных интервалов $T = \{t_0, t_1, \dots, t_m\}$, т. е. можно предположить о существовании системы в момент времени t_i и в предыдущий момент времени t_{i-1} .

Система имеет следующие виды агентов: трамваи; ТС; пешеходы.

Пусть группа трамваев состоит из n элементов: $A = \{a_1, \dots, a_n\}$. Каждый агент трамвай в момент времени t_j описывается следующим набором атрибутов:

$$a_{i,t_j} = \{(x, y)_{i,t_j}, v_{i,t_j}, stop_{i,t_{j+1}}\},$$

где $(x, y)_{i,t_j}$ и v_{i,t_j} – местоположение и скорость агента a_i в момент времени t_j ;

$stop_{t_{j+1}}$ – наличие остановки в момент времени t_{j+1} для агента a_i .

Отметим, что обозначение fun представляет некоторую зависимость параметров, т. е. $b = fun(c)$ – зависимость параметра b от параметра c .

Определим состояние участка дороги в момент времени t_j :

$$S_{x,y,t_j} = \{traffic_jam_{roal_line_{x,y,t_j}}, repair_{x,y,t_j}\},$$

где $traffic_jam_{roal_line_{x,y,t_j}}$ – плотность потока на дороге в момент времени t_j в баллах на дороге $roal_line_{x,y,t_j}$;

$repair_{x,y,t_j}$ – коэффициент, показывающий наличие ремонтных работ на участке (x, y) при 1 – ведется ремонт дороги, 0 – в обратном случае.

Зависимость скорости от плотности потока многополосной и однополосной дорог приведена в работе [6]. Авторы предложили переход к усредненному представлению скорости и плотности потока дороги, по аналогии с многокомпонентной односкоростной моделью

баротропного газа и решения численного интегрирования системы из $S+1$ уравнений для многоскоростной или односкоростной систем, где S – количество полос на дороге. Так, перенос концентрации ТС между разными полосами (в том числе заезд на полосы движения трамвая) описаны уравнениями конвективно-диффузного переноса.

В рамках настоящей работы под состоянием дороги понимается величина, сильно влияющая на скорость движения трамвая. Тогда скорость движения трамвая a_i в момент времени t_j на участке (x, y) определяется состоянием трамвая и внешними факторами (состоянием дороги S_{x,y,t_j}):

$$v_{i,x,y,t_j} = \text{fun}(a_{i,t_j}, S_{x,y,t_j}).$$

Получим величину скорости, на которую влияет S_{x,y,t_j} :

$$v_{i,x,y,t_j} = \text{fun}(S_{x,y,t_j}) \times \text{fun}(a_{i,t_j}) = Z_{x,y,t_j} \times \text{fun}(\text{repair}_{x,y,t_j}) \times \text{fun}(a_{i,t_j}),$$

где $Z_{x,y,t} = \text{fun}(\text{traffic_jam}_{\text{road_line}_{x,y,t_j}})$ – коэффициент линейной зависимости скорости движения трамвая от вероятности наезда ТС на трамвайные рельсы.

Определим коэффициент Z_{x,y,t_j} на основании исследования Яндекса изменении скорости движения транспорта от уровня загруженности дороги [7]:

$$Z_{x,y,t_j} = \begin{cases} \text{traffic_jam}_{t_j} \in [1; 3]: & [0,7; 1] \\ \text{traffic_jam}_{t_j} \in [4; 6]: & [0,3; 0,7). \\ \text{traffic_jam}_{t_j} \in [7; 10]: & [0; 0,3) \end{cases}$$

Аналогично, в контексте представляемой системы, можно сказать, что $k_{\text{repair}_{x,y,t_j}} = \text{fun}(\text{repair}_{x,y,t_j})$ – является коэффициентом зависимости наличия ремонтных работ на дороге от времени t_j и местоположения участка дороги (x, y) :

$$v_{i,x,y,t_j} = k_{\text{repair}_{x,y,t_j}} \times \text{fun}(a_i) \times Z_{x,y,t_j}.$$

Так как $a_{i,t_j} = \{(x, y)_{i,t_j}, v_{i,x,y,t_j}, \text{stop}_{i,t_{j+1}}\}$, тогда:

$$v_{i,x,y,t_j} = Z_{x,y,t_j} \times k_{\text{repair}_{x,y,t_j}} \times v_{x,y} \times k_{\text{stop}_{t_{j+1}}},$$

где $v_{x,y}$ – средняя скорость агента a_i на участке (x, y) в условиях отсутствия влияния human factors; $k_{\text{stop}_{t_{j+1}}}$ – коэффициент скорости трамвая при наличии остановки трамвая в момент времени t_{j+1} :

1. $\text{stop}_{t_{j+1}} = \text{False}: k_{\text{stop}_{t_{j+1}}} = 1.$
2. $\text{stop}_{t_{j+1}} = \text{True}: k_{\text{stop}_{t_{j+1}}} \in [0, 1).$

Для построения моделей транспортных потоков на микроуровне применяются средства: PTV Vissim, Quadstone Paramics и TTS Aimsun [8].

В качестве программной платформы для реализации рассматриваемой модели зависимости влияния внешних human factors на скорость движения трамвая выбран инструмент имитационного моделирования PTV Vissim Demo, позволяющий строить модели движения транспортных потоков. Система PTV Vissim позволяет имитировать и визуализировать взаимодействие любых видов транспорта в рамках единой модели с использованием правил дорожного движения, что позволяет говорить о приближении проведенного эксперимента к реальным условиям.

Система PTV Vissim обеспечивает реалистичное и всестороннее представление транспортных потоков, а также всех аспектов управления этими потоками, которые динамически адаптируются к дорожно-транспортной обстановке, определяя или воспроизводя индивидуальные параметры моделируемого участка.

Количество данных, используемых в симуляторе: трамваев – 5; ТС – 80; пешеходов – 80; участков ремонтных работ – 2.

На рисунке 2 представлен план карты дороги с участками 1–4. Участок 1 – дорога только для ТС, участки 2, 3 – трамвайные рельсы, на которые не могут заехать ТС. Участок дороги 4 является круговым по две полосы в каждую сторону: по обеим полосам могут двигаться ТС, а по внутренним полосам – трамваи.

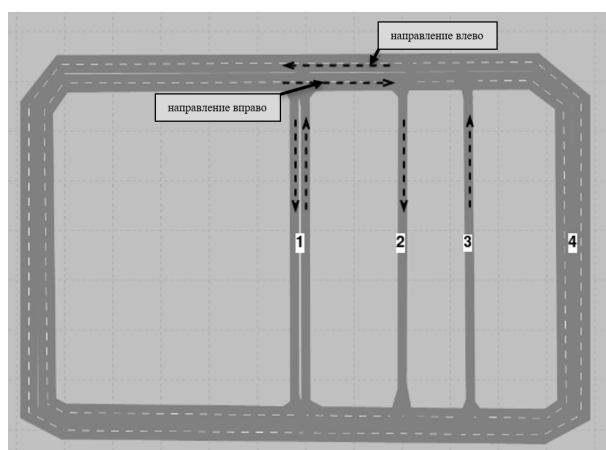


Рис. 2. План карты дороги. Движение правостороннее (направления движения вправо и влево обозначены пунктирными стрелками)

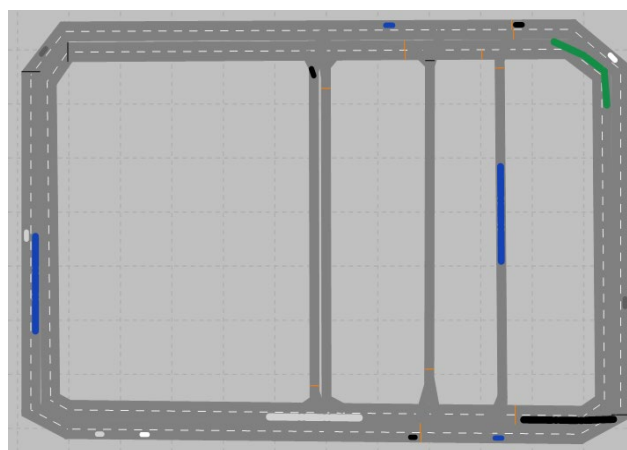


Рис. 3. Схема движения трамваев и транспортных средств по дороге

Симуляция начинается с равномерного распределения агентов в помещении. На рисунке 3 показана работа симулятора: длинными овалами синего, зеленого и черного цвета обозначены трамваи, короткими овалами черного, белого, серого и синего – ТС, желтыми линиями обозначены стоп-линии на второстепенных участках, черными линиями обозначены места появления агентов симуляции.

На первом этапе выполнено 10 запусков системы с 5 агентами трамваями для получения значений средних скоростей на участках дороги.

На втором этапе проведено 10 запусков системы с 5 агентами трамваями, 20 агентами ТС, 80 агентами пешеходами для проверки влияния на скорость трамваев human factors без пробок и ремонтных работ, но с моделью поведения агентов ТС с заездом на трамвайные рельсы.

На третьем этапе проведено 10 запусков системы с 5 агентами трамваями, 40 агентами ТС, 80 агентами пешеходами, а также с ремонтными работами на одном участке дороги для проверки влияния на скорость human factors с небольшими пробками (не более 5 баллов), а также влияния ремонтных работ.

На четвертом этапе выполнено 10 запусков системы с 5 агентами трамваями, 80 агентами ТС, 80 агентами пешеходами, а также с ремонтными работами на двух участках дороги для проверки влияния на скорость human factors на загруженной дороге (пробки более 5 баллов), а также влияния ремонтных работ на нескольких участках дороги.

На пятом этапе проведена агрегация полученных результатов, оценка качества полученных в результате эксперимента результатов, а также формирование выводов.

В ходе проведения эксперимента произведено по 10 запусков модели для каждого этапа. На третьем этапе на участок 4 был добавлен один участок ремонтных работ (направление вправо (рис. 2)). На четвертом этапе – добавлены два участка ремонтных работ (обе стороны направления (рис. 2)).

В таблице представлены средние скорости для каждого участка дороги в зависимости от этапа эксперимента. На рисунке 4 представлены графики изменения скорости на каждом участке дороги в зависимости от факторов, влияющих на скорость движения трамвая, а также линии тренда, используемая для выявления тенденций изменения скорости трамвая. Как видно из рисунка 4 зависимость скорости от human factors является полиномиальной

Таблица. Средняя скорость движения трамвая, км/ч

Номер участка дороги (рис. 2)	Результаты моделирования			
	Первый этап	Второй этап	Третий этап	Четвертый этап
2	25,32	25,32	23,41	20,16
3	22,99	20,16	23,32	19,53
4 (направление вправо)	24,32	17,26	11,20	10,04
4 (направление влево)	24,90	24,46	19,80	7,88

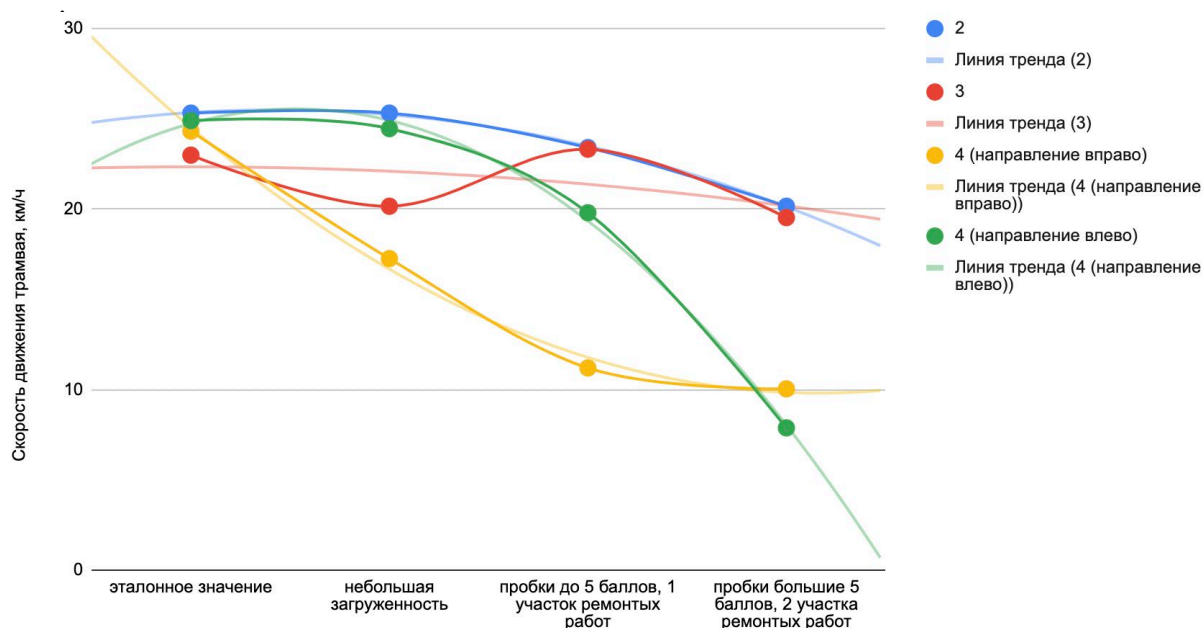


Рис. 4. Графики зависимости скорости на каждом участке дороги в зависимости от факторов, влияющих на скорость движения трамвая

Выполнен анализ полученных результатов и сравнение с данными, полученными при тестировании модели без влияния human factors. В результате тестирования видно, что при добавлении транспортных средств и одного участка ремонтных работ скорость трамвая снижается на 16%, а при увеличении числа ремонтных работ и значительных пробках – 69%. Также отмечено, что скорость трамвая почти не изменяется при движении по специальным выделенным полосам. Исходя из перечисленных этапов, можно говорить о том, что исследование проведено качественно, выдвинута гипотеза, формализован объект исследования, гипотеза подтверждена экспериментом. Работа имеет практическую значимость в области проектирования дорожного движения. Представленное исследование может быть использовано при проведении анализа проектируемых транспортных узлов для беспрепятственного движения беспилотного городского рельсового транспорта.

Литература

1. Krokos K.J., Baker D.P. Preface to the special section on classifying and understanding human error // Human Factors. – 2007. – V.49 (2). – Pp. 175–177.
2. Cacciabue P.C. Human error risk management methodology for safety audit of a large railway organisation. J. Wilson, B. Norris, S. Clarke, A. Mills (Eds.). Rail Human Factors: Supporting the Integrated Railway, Ashgate Publishing Limited, Cornwall. – 2005. DOI:10.1016/J.APERGO.2005.04.005.
3. Branton R. Investigations into the skills of train driving // Ergonomics. – 1997. – V.22. – Pp. 155–164.
4. Buck L. Errors in the perception of railway signals // Ergonomics. – 2006. – V.6. – Pp. 181–192.
5. Bourne A., Carey M. Integrating human factors into the development of railway systems People in Control: The Second International Conference on Human Interfaces in Control Rooms, Cockpits and Command Centres, Manchester, 19–21 June 2001. The Institution of Electrical Engineers, London. – Pp. 25–30.
6. Морозв И.И., Гасников А.В., Тарасов В.Н., Холодов Я.А., Холодов А.С. Численное исследование транспортных потоков на основе гидродинамической модели // Компьютерные исследования и моделирование. – 2011 – Т.3. – В.4. – С. 389–412.
7. Исследование Яндекса – Пробки в Москве 2013-2017. [Электронный ресурс]. — Режим доступа: https://yandex.ru/company/researches/2017/moscow_traffic_2017 (дата обращения: 30.11.2021).
8. Захаров Ю.И., Карнаух Е.С. Основные современные инструменты имитационного моделирования транспортных потоков // Вісник ПДАБА. – 2014. – №1(190). – С. 46–51.

УДК 620.179.16

РАЗРАБОТКА ТЕХНОЛОГИИ КОНТРОЛЯ ПАЯНЫХ СОЕДИНЕНИЙ ТЕПЛООБМЕННЫХ АППАРАТОВ

Малый В.В.¹ (аспирант), Костюхин А.С.¹ (аспирант)

Научный руководитель – канд. техн. наук Кинжагулов И.Ю.¹

¹Университет ИТМО

e-mail: banqvalera@mail.ru, noxx9999@yandex.ru, kinzhiki@mail.ru

В данной статье рассматриваются вопросы контроля качества паяных соединений теплообменных аппаратов. Описаны дефекты, возникающие при изготовлении теплообменных аппаратов. Проанализированы существующие решения, выявлены достоинства и недостатки, изложены предпочтительные решения для дальнейшей работы.

Ключевые слова: неразрушающий контроль, ультразвуковой контроль, контроль паяных соединений, теплообменные аппараты, фазированная антенная решетка.

Теплообменные аппараты (ТА) представляют собой геометрически сложную конструкцию с множеством различных внутренних полостей, что, в свою очередь обуславливает сложность технологии пайки, связанную с нанесением припоя и обеспечением стабильности величины гарантированного зазора между сопрягаемыми элементами. Наиболее остро стоит вопрос качества паяных соединений (ПС), так как из-за особенностей режимов работы ТА, а также из-за высоких нагрузок при эксплуатации, дефект может привести к выходу дорогостоящего ТА из строя, который по стоимости разработки и производства составляет 20 – 40% стоимости изделия.

Объект контроля (ОК) имеет сложную оребренную внутреннюю структуру. Паяная оребренная конструкция ТА представляют собой две пластины из разных металлов с фрезерованными на одной из них канавками. Ребра одной пластины соединяются с другой посредством припоя. Внутренняя структура ОК представлена на рисунке 1.

Основными дефектами ПС являются «непропай» и «неспай», а их наличие, в большинстве случаев, приводит к нарушениям теплообменных режимов в аппарате и как следствие к выходу изделия из строя.

Согласно ГОСТ 17325-79 [1], рассматриваемые в статье дефекты имеют определения:

1. Непропай – дефект паяного соединения, проявляющийся в частичном или полном незаполнении паяльного зазора припоем.
2. Неспай – дефект паяного соединения, проявляющийся в отсутствии сцепления паяемого материала с материалом паяного слоя.

Зачастую, качество ПС определяется гидравлическими испытаниями на прочность, пневматическими испытаниями на герметичность, а также огневыми испытаниями. Поскольку данные виды испытаний являются разрушающими, невозможно обеспечить целостность ТА в процессе его изготовления. Таким образом, остро стоит вопрос о создании технологий контроля неразрушающего контроля (НК) ПС ТА.

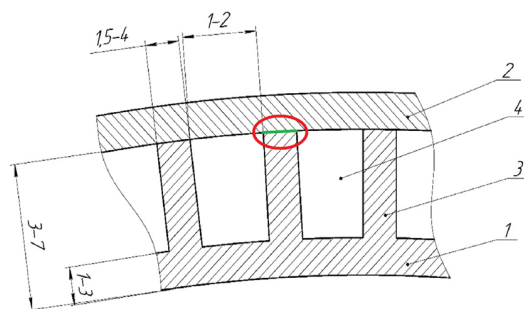


Рис. 1. Внутренняя структура ПС ТА:
1 – внутренняя стенка, 2 – внешняя стенка, 3 – ребро, 4 – канавка

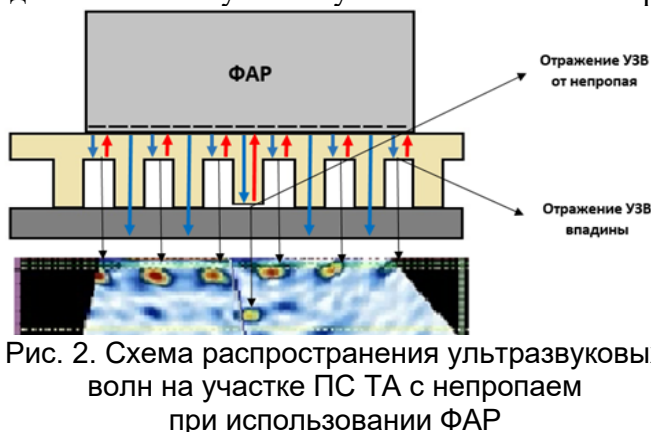
В настоящее время, существуют методики НК, позволяющие выявлять дефекты типа «непропай» и «частичный непропай» [2, 3]. Проанализировав данные работы, было отмечено, что ультразвуковой эхо-импульсный метод является наиболее перспективным методом контроля качества ПС ТА. Также, проведя анализ текущих достижений в области неразрушающего контроля паяных соединений, был сделан вывод, что на данный момент существуют методы и средства контроля, позволяющие обнаруживать дефекты типа «непропай» и «частичный непропай». Однако, данные методики не позволяют гарантированно выявлять дефект типа «неспай» в ПС ТА, что существенно увеличивает риски отказа всего изделия при воздействии эксплуатационных нагрузок. Предполагается, что для обеспечения выявления дефектов типа «неспай» необходимо использовать комбинирование эхо-импульсного и теневого метода ультразвукового контроля.

Ультразвуковой эхо-метод основан на возбуждении упругих колебаний в объеме контролируемого изделия и регистрации отраженного от дефекта сигнала. Информативными параметрами при применении эхо-импульсного метода являются амплитуда и время прихода эхо-сигнала. В бездефектном ПС ультразвуковой сигнал проходит через границу раздела двух акустически разнородных сред, что обуславливает возникновение акустического импеданса (отражения) на данной границе. Величина импеданса существенно зависит от сплошности диффузионной границы. Наличие в ПС дефекта типа «непропай» приведёт к незначительному увеличению акустического импеданса. При многократном прохождении диффузионного раздела сред ультразвуковой волной заметно большее её затухание при наличии «неспая», по сравнению с однородным соединением.

Ультразвуковой теневой метод основан на генерации ультразвуковых колебаний пьезоэлектрическим преобразователем, работающим в режиме излучения, и приеме прошедшего через объект контроля сигнала приемным преобразователем. Признаком обнаружения дефекта, при использовании данного метода, является ослабление прошедшего сигнала, вызванного экранированием части ультразвукового пучка несплошностью. В ПС, содержащим дефекты типа «непропай» и «неспай», амплитуда прошедшего сигнала уменьшается, по сравнению с прошедшим сигналом на бездефектном ПС.

Для упрощения интерпретации полученных результатов и визуализации контролируемого сечения ТА, предлагается использовать дефектоскоп с фазированными антенными решетками (ФАР), реализующим ультразвуковой эхо-метод.

Главное преимущество технологии ФАР перед одноэлементным пьезоэлектрическим преобразователем (ПЭП) заключается в том, что управление лучом и его фокусировка осуществляется с помощью одного преобразователя содержащим множество пьезоэлементов. Последовательное возбуждение пьезоэлементов образующих активную апертуру ФАР позволяет задавать фокальные законы для контролируемого сечения объекта, тем самым добиваясь наилучшей чувствительности на определенной глубине.



Секторное сканирование, используется для картографирования объектов, что значительно упрощает контроль объектов со сложной геометрией, таких как ТА. Большая, чем у одноэлементных ультразвуковых преобразователей апертура позволяет при установке ФАР на поверхность изделия регистрировать на экране дефектоскопа сигналы, соответствующие внутренней структуре объекта контроля под датчиком. Типовой S-скан с ПС с непропаем представлен на рисунке 2.

Реализация схем сканирования эхо-импульсным и теневым методом ввиду сложности внутренней ребренной структуры ОК накладывает значительные ограничения на скорость и точность результатов ручного контроля. Проведение сплошного ручного контроля ТА крайне

трудоемко, что в свою очередь повышает вероятность ошибки оператора при проведении и анализе результатов сканирования.

Повышение скорости и достоверности результатов контроля возможно за счет автоматизации процесса сканирования. Проведение автоматизированных измерений с точки зрения механики заключается в перемещении преобразователя по поверхности объекта контроля с постоянной скоростью по заданной траектории. Поскольку, для ввода ультразвуковых волн в ОК используется иммерсионная жидкость, требуется соблюдать постоянный по величине зазор между ПЭП и поверхностью ОК.

Автоматизации процесса НК ПС ТА, позволят:

1. Минимизировать влияние человеческого фактора при проведении НК, так как квалификация оператора не будет влиять на процесс измерений.
2. Стабилизировать акустический контакт за счет использования иммерсионных ванн / струйных подач иммерсионной жидкости, а также постоянного зазора между ПЭП и ОК.
3. Повысить оперативность контроля за счет повышения скорости сканирования и протоколирования результатов контроля по удобной для интерпретации форме (В-скан, С-скан).

Таким образом, проведен анализ методов и средств ультразвуковой дефектоскопии, в части их применимости для выявления дефектов типа «неспай» и «непропай» в паяных соединениях орбренных конструкций теплообменных аппаратов. Для повышения вероятности обнаружения дефектов в паяных соединениях, принято решение использовать комбинированный эхо-импульсный и теневой методы. Для упрощения интерпретации и визуализации результатов сканирования предлагается использование технологии ФАР для реализации эхо-метода. Поскольку реализация схем контроля ТА требует значительных временных затрат при ручном сканировании, для повышения скорости и достоверности результатов предлагается автоматизировать процесс контроля.

Литература

1. ГОСТ 17325-79. Пайка и лужение. Основные термины и определения: утвержден и введен в действие постановлением Государственного комитета СССР по стандартам от 11 октября 1979 г. N 3914: дата введения 1981-01-01. М.: Издательство стандартов. – 1990. – 22 с.
2. Кинжагулов И.Ю. Модель термооптического возбуждения ультразвуковых волн в паяных тонкостенных изделиях // Изв. вузов. Приборостроение. – 2011. – Т.54. – №7. – С. 39–44,
3. Сергеев Д.С., Астрединова Н.В. Методика контроля качества паяных соединений камер ЖРД с применением метода лазерно-ультразвуковой диагностики // Сборник трудов II Всероссийский конгресс молодых ученых. – СПб: НИУ ИТМО. – 2013. – С. 67.

УДК 004.056

ОБЕСПЕЧЕНИЕ ИНФОРМАЦИОННОЙ БЕЗОПАСНОСТИ В МЕДИЦИНСКИХ ВЕБ-ПРИЛОЖЕНИЯХ

Меренков Д.Н.¹ (студент)

Научный руководитель – канд. техн. наук Коржук В.М.¹

¹Университет ИТМО

e-mail: daniil@itmo.ru

Статья посвящена обеспечению информационной безопасности медицинских веб-приложений. Рассмотрены статистические вопросы, касающиеся стоимости утечки медицинской информации, проанализированы основные векторы атак на медицинские веб-приложения, а также методы защиты; предложен систематический подход к эффективной и безопасной разработке медицинских веб-приложений.

Ключевые слова: информационная безопасность, медицина, веб-приложения, систематический принцип разработки, утечка данных, методы защиты.

Основная цель работы – обеспечить эффективную и безопасную разработку медицинских веб-приложений с использованием систематического подхода, что приведет к увеличению уровня информационной безопасности. Задачи работы включают следующие шаги:

1. Определить актуальность обеспечения информационной безопасности в медицинских веб-приложениях.
2. Определить основные векторы атак на медицинские веб-приложения.
3. Определить основные методы защиты от атак на медицинские веб-приложения.
4. Предложить принцип систематической разработки медицинских веб-приложений.

Особенность медицинских веб-приложений заключается в том, что медицинские данные обладает особой ценностью для злоумышленников, а быстрое развитие цифровизации в данной сфере, подталкивает бизнес к быстрой разработке и релизу, что способствует увеличению числа ошибок, в том числе, ошибок безопасности.

По данным IBM Corporation 11 лет подряд здравоохранение имеет самую высокую стоимость утечки информации в отрасли. Затраты на утечку данных в сфере здравоохранения увеличились со средней общей суммы в 7 млн долларов в 2020 году до 9 млн долларов в 2021 году (рис. 1), т.е. на 30% [1], в то время, как во всех остальных секторах, затраты только уменьшились.

Финансовые потери (млн. \$) в различных секторах в 2020 и 2021 годах

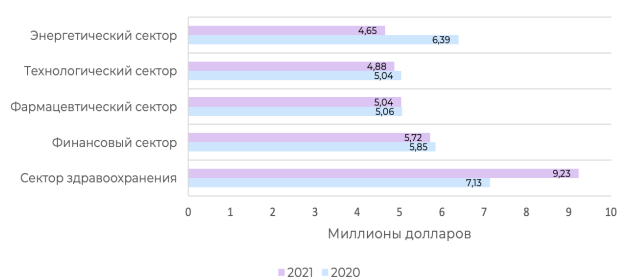


Рис. 1. Сравнение финансовых потерь в различных секторах в 2020 и 2021 гг.

Сравнительная диаграмма количества актов утечек и утекших данных по годам

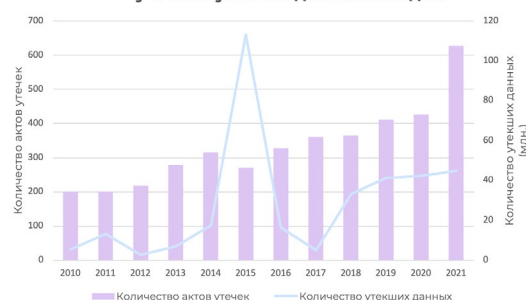


Рис. 2. Сравнительная диаграмма количества актов утечек и утекших данных по годам

Почти 45 миллионов медицинских записей были раскрыты или украдены в 2021 году, что делает его вторым худшим годом с точки зрения утечки медицинских записей (рис. 2).

Три четверти утечек за год (73,9%) были связаны со взломом или другими IT-инцидентами, остальные утечки произошли по причине человеческого фактора, либо физического доступа к данным [2].

Согласно отчетам Check Point и Fortified Health Security, число кибератак на медицинские учреждения выросло на 45% с ноября 2020 года [3].

Крупнейшей утечкой медицинских данных в 2021 году был инцидент с Florida Healthy Kids Corporation. Утечка данных произошла из-за того, что поставщик систем безопасности не применял патчи для исправления нескольких уязвимостей на веб-сайте организации в течение 7 лет. В результате утечки была раскрыта информация о здоровье 3,5 миллиона человек.

Данная статистика дает понимание того, что здравоохранение является наиболее целенаправленной и чувствительной областью для атак злоумышленников [2].

Приведем основные векторы атак на медицинские веб-приложения, часть из которых входит в OWASP ТОП 10 2021 (Open Web Application Security Project):

1. Использование уязвимостей контроля доступа (A01:2021-Broken Access Control).
2. Незащищённость конфиденциальных данных (A02:2021-Cryptographic Failures).
3. DoS/DDoS (Denial of Service/ Distributed Denial of Service).
4. Инъекции (A03:2021-Injection).
5. XSS (Cross-Site Scripting).
6. Использование уязвимостей хранения сессий (A07:2021-Identification and Authentication Failures).
7. CSRF (Cross-Site Request Forgery).
8. SSRF (A10:2021-Server-Side Request Forgery).
9. Brute Force (Атака методом перебора).

Наиболее распространенная угроза безопасности веб-приложений – это эксплуатация уязвимостей. Уязвимости веб-приложений возникают тогда, когда разработчики добавляют небезопасный код в веб-приложение. Это может происходить как на этапе разработки, так и на этапе доработки или исправления найденных ранее уязвимостей.

Рассмотрим существующие методы защиты [4]:

1. Мониторинг (Мониторинг является неотъемлемой частью безопасности веб-приложений, так как первым шагом злоумышленника при подготовке к атаке является сканирование, что можно отследить на раннем этапе).
2. WAF (Web Application Firewall) - совокупность мониторов и фильтров, предназначенных для обнаружения и блокирования сетевых атак на веб-приложение. WAF автоматически позволяет экранировать входные и выходные данные. Для защиты тщательно проверяются данные, полученные от пользователя, используются чёрные или белые списки последовательностей символов, экранируется текст запроса:
 - 2.1. Сигнатурный анализ (сигнатурный метод в своей реализации использует словарь вредоносного трафика для сравнения).
 - 2.2. Репутационный фильтр IP-адресов.
 - 2.3. Сканер уязвимостей.
 - 2.4. Защита от DoS атак.
3. Регулярные обновления безопасности.
4. Использование DLP и UBA (по данным InfoWatch в медицинской сфере в 2021 году оказалось около 30% внутренних нарушителей [5]. Использование DLP и UBA позволяет защититься от них за счет того, что нетипичное поведение пользователя будет расценено как аномалия, и об этом немедленно будет оповещена служба информационной безопасности).
5. Database firewall (DBF) / Database Activity Monitoring (DAM) – это аппаратно-программные комплексы для мониторинга, аудита и контроля доступа к информации в базах данных и защиты от целевых атак на них.

Принцип систематической разработки, представленный на рисунке 3, положительно раскрывает себя в условиях быстрого спроса на разработку. Конкурентный рынок разработки веб-приложений создает сложную и рискованную культуру быстрого выпуска веб-приложений, не готовых к "production", что приводит к небрежности в различных фундаментальных аспектах проектирования.

Рассмотрим основные шаги в систематической разработке.

Первый шаг – это понимание идеи: этот шаг инициирует идеи о развитии продукта.

На этом этапе разработчик, программист и владелец веб-приложения обсуждают функциональность и идеи о том, каким будет веб-приложение после разработки и какие функции требуются владельцу.

Вторым шагом идет определение требований. После обсуждения и анализа идеи веб-приложения необходимо определить, какие процессы требуют внимания при разработке и проектировании приложения.

Далее необходимо классифицировать потенциальные угрозы - этот шаг отвечает за предотвращение явных эксплойтов в системе за счет определения основных угроз и проработки методов их закрытия.

Основной шаг - дизайн и код. Это этап, на котором разработчики разрабатывают систематическое приложение, связывая определенные требования между собой.

На следующем шаге происходит настройка и первичная эксплуатация средств защиты - на данном этапе настраиваются средства защиты, описанные ранее, а также сопоставляется работа бизнес-логики приложения с используемыми средствами защиты. Процесс идет параллельно этапу дизайна и кода.

Далее идет тестирование, данный шаг не нуждается в подробном описании. После тестирования происходит проверка удовлетворения потребностей. Если потребности не удовлетворены, возвращаемся к пересмотру определения требований или дизайн и коду.

Когда все потребности удовлетворены, начинается процесс эксплуатации программного продукта и средств защиты.

Финальным шагом или результатом деятельности предыдущих этапов является упрощение разработки с учетом обеспечения ИБ.

В целом, после описания концепции и рабочего процесса становится ясно, что описанный подход является первым начальным шагом к обеспечению эффективной и безопасной разработки медицинских веб-приложений. Стоит принять во внимание, что предлагаемый подход на данном этапе отображает только концептуальное моделирование идеи. В дальнейшей работе планируется формализовать предложенный подход, оценить его слабые стороны и дать не только качественную, но и количественную оценку его эффективности.

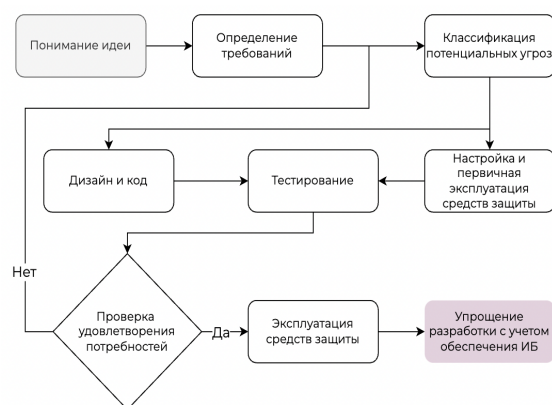


Рис. 3. Принцип систематической разработки

Литература

1. Cost of a Data Breach Report 2021. [Электронный ресурс]. – Режим доступа: <https://www.ibm.com/security/data-breach> (дата обращения: 28.02.2022).
2. Largest Healthcare Data Breaches of 2021. [Электронный ресурс]. – Режим доступа: <https://www.hipaajournal.com/largest-healthcare-data-breaches-of-2021/> (дата обращения: 28.02.2022).
3. Healthcare Accounts for 79% of All Reported Breaches, Attacks Rise 45%. [Электронный ресурс]. – Режим доступа: <https://healthitsecurity.com/news/healthcare-accounts-for-79-of-all-reported-breaches-attacks-rise-45> (дата обращения: 28.02.2022).

4. Информационная безопасность web-приложений. Современные решения. [Электронный ресурс]. – Режим доступа: <https://gardatech.ru/articles/smi/informatsionnaya-bezopasnost-web-prilozheniy-sovremennye-resheniya/> (дата обращения: 28.02.2022).
5. Медицина: утечки информации по вине сотрудников. [Электронный ресурс]. – Режим доступа: <https://www.infowatch.ru/analytics/daydzhesty-i-obzory/meditsina-utechki-informatsii-po-vine-sotrudnikov> (дата обращения: 28.02.2022).

УДК 004.93.11

ПОИСК АНОМАЛЬНЫХ ПОСЛЕДОВАТЕЛЬНОСТЕЙ ПИКСЕЛЕЙ В РАМКАХ АНАЛИЗА СНИМКОВ КОЖНЫХ ПОРАЖЕНИЙ

Милантьев С.А.¹ (аспирант), Литвинов П.А.¹ (аспирант)

Научный руководитель – д-р техн. наук, профессор Бессмертный И.А.¹

¹Университет ИТМО

e-mail: geerkus@gmail.com, p.litvinov7@gmail.com

Целью исследования является проверка гипотезы о возможности применения методов извлечения признаков для определения на фото болезни Боуэна и злокачественной меланомы. В ходе исследования гипотеза была проверена на датасете ISIC. Было выявлено, что извлечение признаков из цветов дает точность выше, чем извлечение признаков из градаций серого. При этом, наибольшая точность была достигнута при следующей последовательности операций: Preprocessing → RGB → HOG → SVM, и составила $72,5 \pm 2\%$.

Ключевые слова: компьютерное зрение, HOG, сверточные сети, болезнь Боуэна, злокачественная меланома.

Онкология традиционно лидирует в антирейтинге по смертности [1]. По данным Единого государственного реестра ЗАГС за 2020 год от онкологических заболеваний умерло 289,9 тысяч человек. Специфика онкологических заболеваний в том, что заболевание тем легче вылечить, чем раньше оно было диагностировано. Именно ранняя, качественная и доступная диагностика дает возможность существенно снизить смертность от онкологических заболеваний. Одним из основополагающих методов является дерматоскопия. В отличие от биопсии дерматоскопия позволяет провести специалисту анализ без повреждения кожного покрова. Применимость методов машинного обучения в данном случае значительно снизит ошибку классификации и уменьшит влияние человеческого фактора на результат. Использование и улучшение классических подходов для анализа кожных поражений является приоритетным направлением исследований.

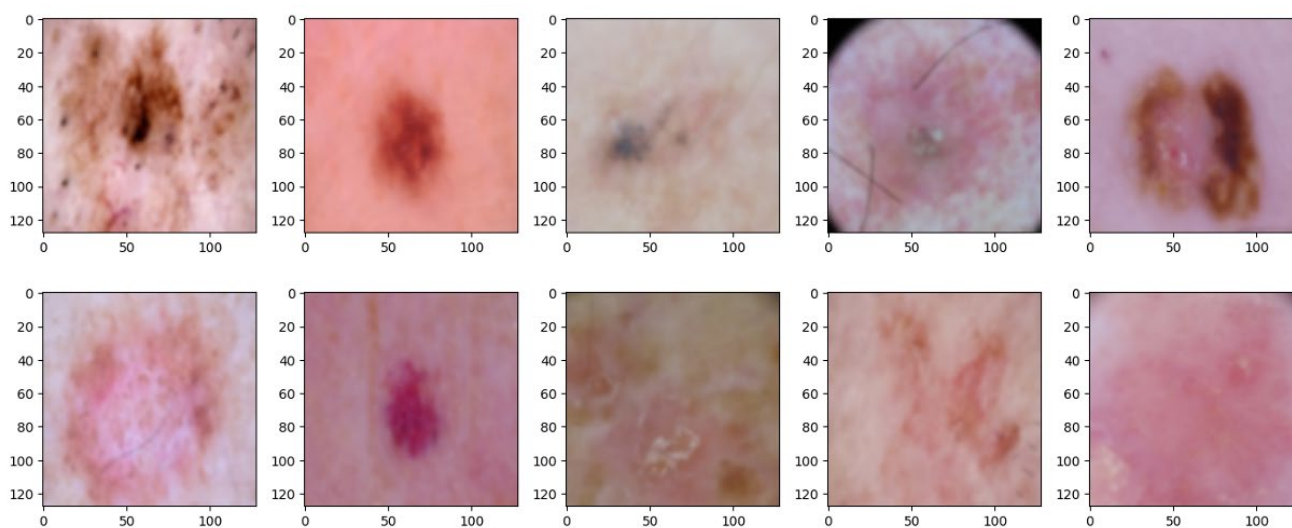


Рис. 1. Примеры исследуемых изображений

В рамках данного исследования предлагается использовать комбинацию метода гистограммы направленных градиентов и классической модели по методу опорных векторов для поиска аномальных последовательностей пикселей в рамках анализа снимков кожных поражений. Для исследования была использована подвыборка датасета ISIC в 3000 изображений, содержащая снимки болезни Боуэна и злокачественной меланомы (рис. 1).

Было предложено применить следующую последовательность операций: Preprocessing → RGB → HOG → SVM. Детальнее рассмотрим каждый этап конвейера.

Перед расчетом гистограммы направленных градиентов необходимо произвести предобработку. Предобработка позволяет не только снизить размерность, но и удалить артефакты, уменьшающие конечную точность. Первым шагом идет снижение размерности, критически важно правильно выбрать (или подобрать) размерность. В нашем случае экспериментально был установлен оптимальный размер изображения 128x128 пикселей. Стоит отметить, что размерность от 90x90 до 160x160 пикселей также показывала приемлемые результаты. Следующим шагом препроцессинга было размытие по Гауссу. В рамках данного исследования использовалось размытие с ядром 5x5 и 3x3.

К получившимся после Гауссовского сглаживания изображениям применялись цветовые преобразования в YCbCr, RGB, SoG. В ходе исследования последующие шаги (HOG и SVM) применялись ко всем цветовым схемам. Результаты сравнения точности предложенного метода на разных цветовых схемах представлены в конце настоящей статьи.

Для поиска признаков на изображениях был применен метод гистограммы направленных градиентов. Метод сводится к подсчету количества направлений градиента в локальных областях изображения. В нашем случае в каждом блоке находится по четыре ячейки. В свою очередь, в каждой ячейке содержится по 4 градиентных вектора, по которым строится гистограмма направлений. Нормализация ячеек внутри блока производится с применением L2-нормы.

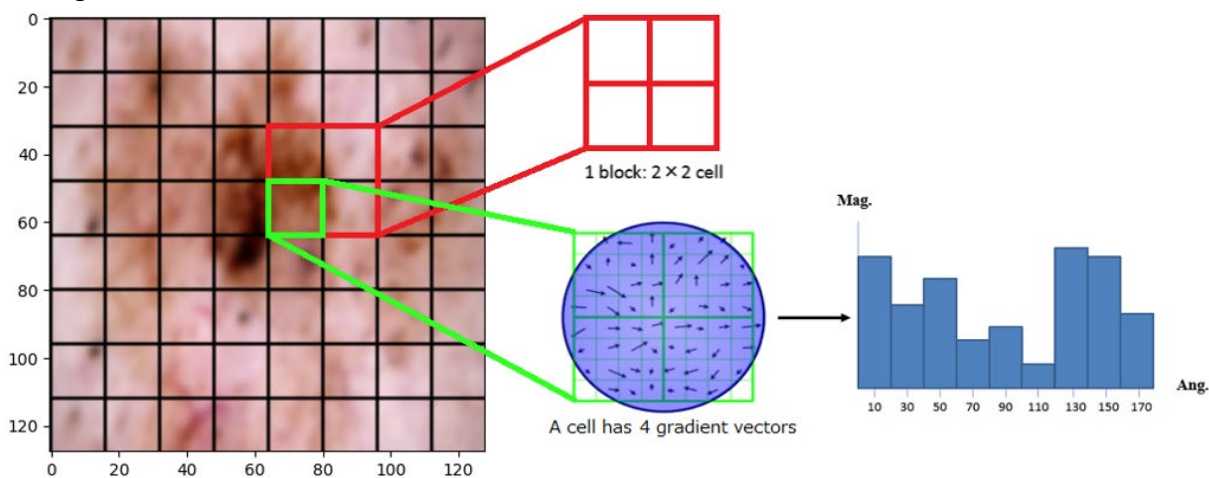


Рис. 2. Принцип работы HOG на тестовых данных

Вышеприведенный алгоритм применялся к каждому цветовому каналу (рис. 2). Технически, алгоритм состоит из этапов [2]:

1. Вычисление градиента.
2. Построение гистограммы направлений градиента.
3. Нормализация внутри блока.

Так, при размере изображения в 128 x 128 пикселей, и при размере ячейки в $2 \times 2 = 4$ ячейки в одном блоке, получаем 32 x 32 блока при единичном страйде. Принимая 4 градиентных вектора на ячейку получаем $4 \times 4 = 16$ признаков в одном блоке. Значит, в одном изображении содержится $32 \times 32 \times 16 = 16\ 384$ признаков или 49 152 с учетом 3 каналов. Пример работы HOG на тестовых данных с цветовой схемой RGB представлен на рисунке 3. Стоит отметить специфику предметной области, что в каждом цветовом канале содержится разная информация. Таким образом, применяется транзитная архитектура решения (рис. 4).

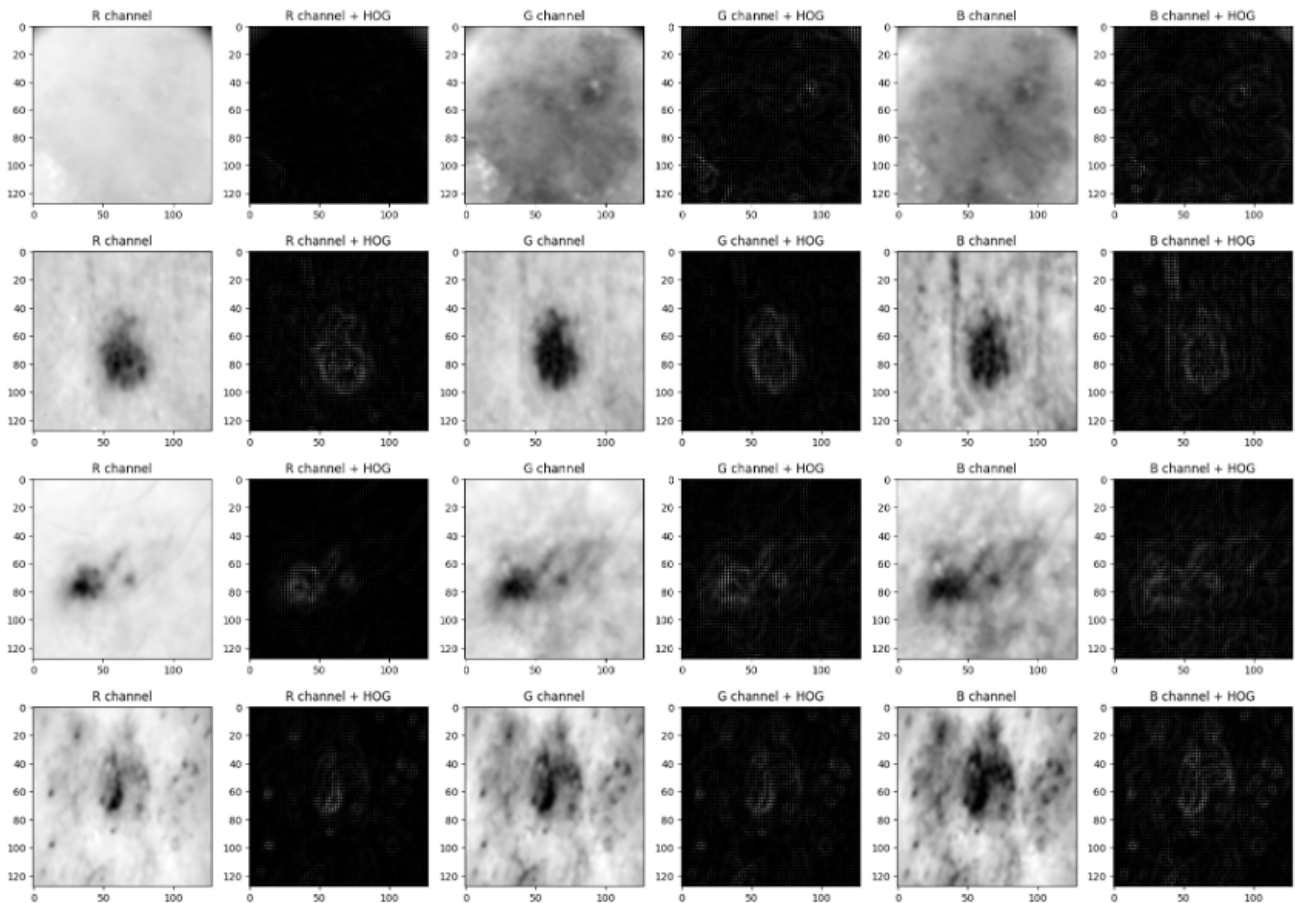


Рис. 3. Результаты работы HOG на тестовых данных

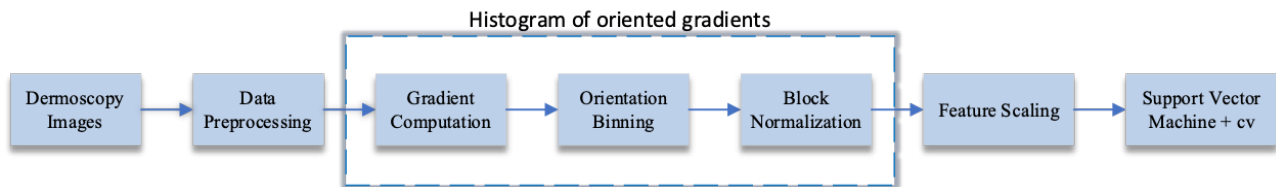


Рис. 4. Транзитная архитектура решения

Для дальнейшего развития исследования было решено применить метод опорных векторов и классическую SVM модель, его реализующую. Такое решение было принято в связи с тем, что решаемая задача фактически является задачей бинарной классификации. Генерализованная архитектура нейросетевой модели [3] с добавлением дополнительных признаков изображения представлена на рисунке 5.

Рассмотрим закрашенную часть модели. В этой модели используется сверточная нейронная сеть и метаданные о пациенте, проходящие через два Dense слоя (256 юнитов в каждом). Полносвязный слой сверточной нейронной сети объединяется с вектором результата по нейросети метаданных и передаются в объединяющий dense слой (1024 юнита). Аналогичным образом, через несколько Dense-слоев, мы внедряем признаки из изображения.

Детальнее рассмотрим количество юнитов в слоях. Предварительно обработанные метаданные поступают в двухслойную нейронную сеть с 256 нейронами в каждой. Каждый слой содержит нормализацию по батчам, функцию активации ReLU (для нормализации входных данных, трансформирующие их в шкалу с нулевым математическим ожиданием и единичной дисперсией) и dropout с низкой вероятностью для глобальной регуляризации обучения. Полносвязный слой сверточной нейронной сети объединяется с вектором результата по нейросети метаданных, и передаются в объединяющий dense-слой с 1024 нейронами. Затем результат передается в полносвязный классификационный слой.

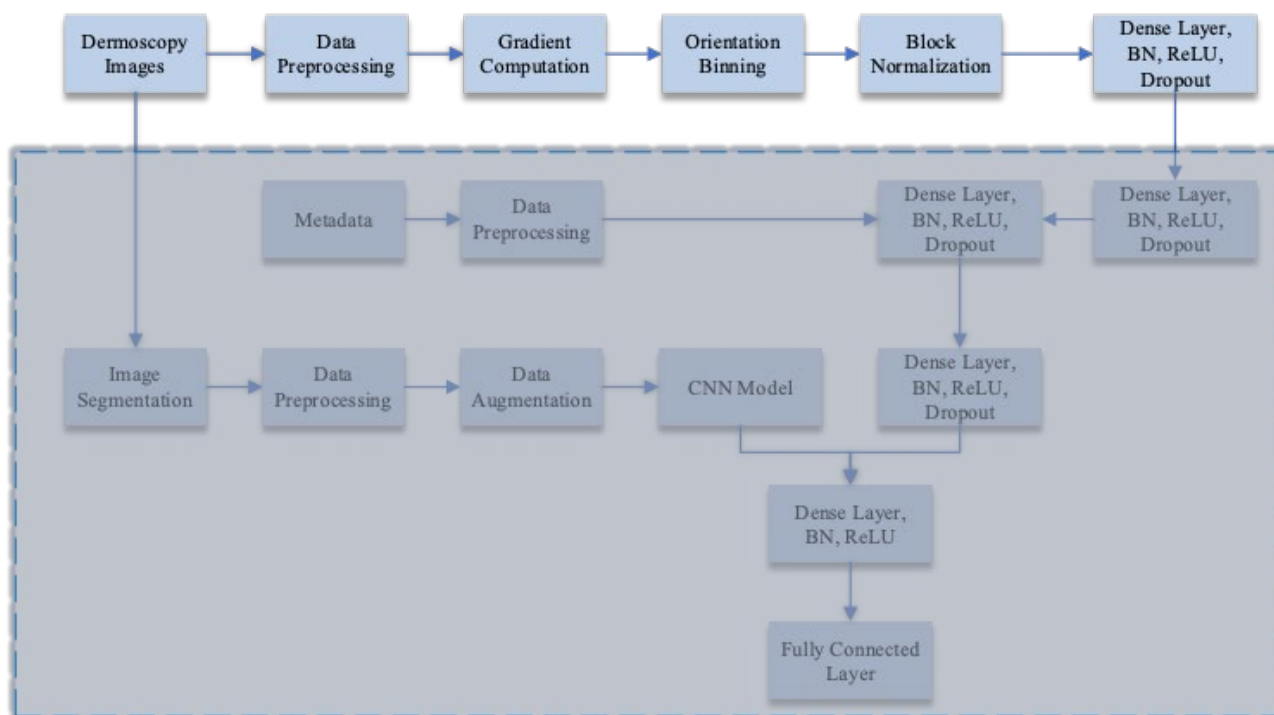


Рис. 5. Целевая архитектура решения

Полученные в ходе исследования результаты были собраны и консолидированы в виде таблицы. Результаты исследования представлены в таблице.

Таблица. Результаты исследования

Experiment	Class	Accuracy	Precision	Recall	F1-Score
Baseline		$50 \pm 3,5\%$	$0,5 \pm 0,035$	$0,5 \pm 0,035$	$0,5 \pm 0,035$
Shades of Gray + HOG + SVM		64%	0,64	0,64	0,64
YCbCr + HOG + SVM		$70 \pm 2\%$	$0,74 \pm 0,01$	$0,7 \pm 0,03$	$0,69 \pm 0,03$
	MEL	-	$0,64 \pm 0,02$	$0,86 \pm 0,03$	$0,77 \pm 0,01$
RGB + HOG + SVM	SCC	-	$0,83 \pm 0,02$	$0,57 \pm 0,04$	$0,68 \pm 0,02$
		$72,5 \pm 2\%$	$0,76 \pm 0,01$	$0,73 \pm 0,02$	$0,73 \pm 0,02$
RGB + HOG + SVM	MEL	-	$0,68 \pm 0,01$	$0,88 \pm 0,02$	$0,77 \pm 0,01$
	SCC	-	$0,84 \pm 0,01$	$0,58 \pm 0,03$	$0,69 \pm 0,02$

Результаты исследования показали, что при извлечении признаков RGB и YCbCr получаются схожие результаты, однако, RGB показывает чуть более высокую точность. Было выявлено, что работы с метаданными пациентов, а также цветовые признаки кожных поражений имели наибольший импакт среди других. Здесь стоит отметить, что в случае кожных поражений цвет содержит ценную информацию, поэтому сразу была выдвинута гипотеза, что при извлечении признаков из градаций серого результат будет сильно хуже, чем при извлечении признаков из RGB.

Литература

1. Всемирная организация здравоохранения. Раковые поражения. [Электронный ресурс]. – Режим доступа: <https://www.who.int/ru/news-room/fact-sheets/detail/cancer> (дата обращения: 03.03.2022).
2. Xu P., Huang L. & Song Y. An optimal method based on HOG-SVM for fault detection // Multimedia Tools and Applications – 2022. – Vol.81. – Pp. 6995–7010.
3. Milantev S., Bessmertnyi I., et al. Skin Lesion Analysis Using Ensemble of CNN with Dermoscopic Images and Metadata // CEUR Workshop Proceedings. – 2021. – Vol.2893.

УДК 004.056

ОЦЕНКА УСТОЙЧИВОСТИ ТЕХНИЧЕСКОГО ЗРЕНИЯ В БЕСПИЛОТНЫХ ТРАНСПОРТНЫХ СРЕДСТВАХ

Мыськив И.И.¹ (аспирант)

Научный руководитель – канд. техн. наук, доцент Заколдаев Д.А.¹

¹Университет ИТМО

e-mail: Myskiv.i@itmo.ru

Работа выполнена в рамках темы НИР №620164 «Методы искусственного интеллекта для киберфизических систем».

В данной статье рассматривается концепция беспилотного транспортного средства и этап восприятия окружающей среды беспилотным транспортным средством, описаны основные проблемы и уязвимости систем восприятия с применением составительных примеров в глубоких нейронных сетях, исследованы основные методы оценки устойчивости глубоких нейронных сетей в системах восприятия против составительных атак.

Ключевые слова: машинное обучение, беспилотные транспортные средства, искусственный интеллект, глубокие нейронные сети, оценка устойчивости, составительные примеры, восприятие.

В последнее десятилетие исследования автоматизированных систем вождения или беспилотных транспортных средств (БТС) показали потенциал интеграции повышения безопасности вождения, удобства передвижения и повышения энергоэффективности на дорогах общего пользования, привлекая при этом большой интерес как со стороны научных кругов, так и бизнеса.

БТС объединяет в себе пять групп технологий:

- датчики и исполнительные механизмы: устройства, обладающие различными возможностями, такими как распознавание и обнаружение объектов, приведение в действие и т.д.;
- искусственный интеллект (ИИ): алгоритмы, которые позволяют вычислительным устройствам (ВУ) выполнять задачи, обычно связанные с окружающей средой;
- машинное обучение (МО): алгоритмы (включая различные нейронные сети), которые позволяют ВУ принимать решение и повышают их способность прогнозировать события или ситуации;
- облачные вычисления: решения, обеспечивающие доступ к общим наборам ресурсов или сервисам, таким как серверы и приложения;
- коммуникации и/или сети: радио-технологии, протоколы связи и беспроводная связь позволяющие осуществлять обмен данными между различными объектами или ВУ (связи V2V, V2X, V2N и т. д.).

Все эти технологии позволяют использовать систему автоматизированного управления БТС, которая взаимодействует с окружающей средой и предпринимает действия, которые позволяют добраться до нужного места назначения [1].

Упрощенная архитектура управления БТС на основе данных технологий представлена на рисунок 1.

Управление БТС описывается в трех основных этапах:

1. Восприятие окружающей среды с помощью датчиков, сенсоров, камер и т.д.

2. Прогнозирование действий на основе алгоритмов искусственного интеллекта, нейронных сетей, машинного обучения.
3. Выполнение заданных действий на основе спрогнозированных ранее результатов.

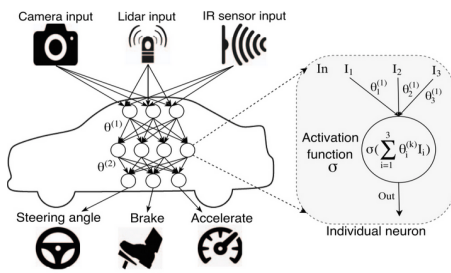


Рис. 1. Упрощенная схема управления БТС

Восприятие БТС или техническое зрение имеет решающее значение для БТС, чтобы принимать решения, планировать и работать в реальных условиях с помощью многочисленных функций и операций, от отображения сетки занятости, распознавания дорожных знаков, обнаружения объектов и т. д.

Некоторыми примерами областей восприятия, включая БТС, являются обнаружение препятствий [2, 3], распознавание объектов [4, 5], семантическая классификация мест [6, 7] и др.

Несмотря на достигнутый значительный прогресс, отрасль находится на начальной стадии этой транспортной революции, и ей предстоит преодолеть множество препятствий, выходящих за рамки технических проблем, связанных с оценкой и стандартами, которые необходимо установить, прежде чем БТС смогут быть коммерциализированы.

Цель. Одним из ключевых проблем безопасности в алгоритмах восприятия с помощью МО было замечено [1], что глубокие нейронные сети (ГНС), в том числе хорошо обученные, оптимизированные для задач восприятия, нестабильны по отношению к так называемым враждебным возмущениям (или состязательным атакам). Такие враждебные возмущения представляют собой (минимальные) изменения входного изображения, часто незаметные для человеческого глаза, которые приводят к неправильной классификации изображения в нейронной сети. Примеры включают не только искусственно созданные случайные возмущения, но и модификации изображений с камеры [5], которые соответствуют изменению размера, обрезке или изменению условий освещения. Они могут быть разработаны без доступа к обучающему набору и могут передаваться в том смысле, что пример, неправильно классифицированный одной сетью, также неправильно классифицируется сетью с другой архитектурой, даже если он обучен на разных данных.

На рисунке 2 показаны противоречивые искажения изображений автомобилей, которые ошибочно классифицируются как птица, лягушка, самолет или лошадь высококвалифицированной современной сетью распознавания.



Рис. 2. Ошибки в классификации

Во всех этих случаях безопасность решения может быть сведена к обеспечению правильного поведения компонента машинного обучения. Однако, методологии обеспечения безопасности и проверки для машинного обучения мало изучены.

Для оценки устойчивости используют методы:

1. Состязательное обучение ГНС. Состязательное обучение с учетом неправильной классификации, чтобы включить в процесс обучения.

Методы состязательного обучения считаются одной из самых сильных принципиальных защит от состязательных атак. Состязательное обучение подвергает модель включением состязательных примеров во время обучения, чтобы получить некоторый уровень иммунитета против них и повысить устойчивость.

К примеру, в работе [4] улучшили состязательное обучение, изменив основанные на генеративно-состязательных сетях (GAN) изменяя представления исходных изображений.

Среди других вариантов состязательных обучающих моделей есть работа [5], где на каждой итерации обучения модель проверяется на надежность с использованием выпуклой релаксации, а дополнительные примеры вычисляются в соответствии с этой релаксацией для целей обучения. В работе [6] также предложили метод планирования отсева для повышения эффективности состязательного обучения с помощью одноэтапных методов.

2. Преобразования входных данных ГНС. Вместо того, чтобы сосредоточиться на устойчивости модели к «состязательным» входам, методы, основанные на преобразовании, направлены на очистку входных данных, чтобы сделать их благоприятными для целевой модели. Например, было изучено сжатие входных данных на основе JPEG для устранения конфликтных возмущений в изображениях [7, 8].

Как правило, преобразование входных данных обеспечивает преимущество, что его можно легко использовать в сочетании с другими защитными механизмами, например, с моделями, обученными по состязательным примерам. В некоторых случаях различные входные преобразования также объединяются для повышения их коллективной эффективности.

В работе [9] представлен один из первых примеров преобразования входных данных с использованием GANs. Их метод, Defense-GAN, изучает распределение чистых изображений. Для вывода он вычисляет выходной сигнал, близкий к входному изображению, которое не содержит потенциального враждебного возмущения.

Защита, основанная на снижении шума, предложена в работе [10], которая избирательно снижает шум в областях изображения с повышенным вниманием, чтобы восстановить правильную метку.

В работе [10] предложена генеративная очистка ансамбля с петлей обратной связи для очистки изображения от конкурирующих шаблонов. Данный метод также опирается на внешние генеративные модули для устранения помех в изображениях противников.

3. Поиск состязательных примеров ГНС. Вместо того, чтобы активно создавать надежную модель на этапе обучения, существуют также методы, которые предоставляют дополнительные механизмы и модули для предварительно обученных моделей, чтобы защитить их от атак противника. В основном эти методы ограничиваются обнаружением наличия враждебных возмущений во входных данных во время вывода.

В работе [11, 12] предложили механизм условно-классовой реконструкции изображений для обнаружения состязательных примеров во время тестирования. Авторы также представили атаку для преодоления этой защиты, продемонстрировав лучшую надежность CapsNet по сравнению со свёрточными нейронными сетями (CNNs) для их атаки. Что еще более важно, их атака демонстрирует большее визуальное сходство между примерами состязательности и категорией целевых объектов для CapsNet. В данном случае это демонстрирует большую согласованность восприятия между представлением CapsNet и зрительной системой человека по сравнению с CNNs.

В работе [13] сосредоточились на классификаторах изображений и подходе метода «BlackBox» для проверки безопасности глубоких нейронных сетей, который не требует знаний об алгоритме. Алгоритм использует методы обнаружения объектов, такие как SIFT (Масштабно-инвариантное преобразование объектов), для извлечения объектов из изображения. Эти объекты преобразуются в изменяемое распределение значимости, где пикселям, влияющим на композицию изображения по отношению к зрительной системе человека, присваивается высокая вероятность. Затем формируется создание состязательных примеров как пошаговую стохастическую игру для двух игроков, где цель первого игрока состоит в том, чтобы минимизировать расстояние до состязательного примера, манипулируя функциями, а второй игрок может быть кооперативным, состязательным или случайным. Используя поиск по дереву Монте-Карло, постепенно исследуется пространство состояний игры для поиска альтернативных состязательных примеров.

Для снижения рисков безопасности ГНС и оценки устойчивости алгоритмов против состязательных атак могут применяться такие методы как:

- состязательное обучение ГНС;

- преобразования входных данных для ГНС;
- поиск состязательных примеров в ГНС;
- разграничение физических и сетевых доступов до компонентов применимых в ГНС.

Алгоритмы МО и ГНС лежат в основе современных систем БТС. Однако, они сталкиваются с рядом угроз и уязвимостей.

Некоторые методы оценки устойчивости специфичны для МО и ГНС, но другие являются стандартными техническими и организационными средствами контроля для снижения рисков общих атак. Важно применять все типы контроля, поскольку в системах БТС, помимо уязвимостей, специфичных для МО и ГНС, существуют также уязвимости общего типа, которые также могут быть использованы злоумышленниками.

Литература

1. Premebida C., Ambrus R., Marton Z.-C. Intelligent Robotic Perception Systems // Applications of Mobile Robots. 2018. DOI: 10.5772/intechopen.79742.
2. Russell S.J., Norvig P. Artificial Intelligence: A Modern Approach. 2nd ed. Upper Saddle River, New Jersey: Prentice Hall. – 2003. – Pp. 1123.
3. Cadena C., Carlone L., Carrillo H., Latif Y., Scaramuzza D., Neira J., Reid I., Leonard J.J. Past, present, and future of simultaneous localization and mapping: Toward the robust-perception age // IEEE Transactions on Robotics. – 2016. – №32(6). – Pp.1309–1332.
4. Gowal S., Qin C., Huang P.-S., Cemgil T., Dvijotham K., Mann T., Kohli P. Achieving robustness in the wild via adversarial mixing with disentangled representations // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. – 2020. – Pp. 1211–1220.
5. Vivek B. and Babu R.V. Single-step adversarial training with dropout scheduling // 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). IEEE. – 2020. – Pp. 947–956.
6. Liu Z., Liu Q., Liu T., Xu N., Lin X., Wang Y., Wen W. Feature distillation: Dnn-oriented jpeg compression against adversarial examples // 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). IEEE. – 2019. – Pp. 860–868.
7. Das N., Shanbhogue M., Chen S.-T., Hohman F., Chen L., Kounavis M.E., Chau D.H. Keeping the bad guys out: Protecting and vaccinating deep learning with jpeg compression // Computer Science. – 2017. ArXiv abs/1705.02900.
8. Guo C., Rana M., Cisse M., Van Der Maaten L. Countering adversarial images using input transformations // ICLR. – 2018. – Pp. 123–154. ArXiv, abs/1711.00117.
9. Samangouei P., Kabkab M., Chellappa R. Defense-gan: Protecting classifiers against adversarial attacks using generative models // ICLR. – 2018. – Pp. 12–24. ArXiv abs/1805.06605.
10. Gupta P. and Rahtu E. Ciidefence: Defeating adversarial attacks by fusing class-specific image inpainting and image denoising // Proceedings of the IEEE/CVF International Conference on Computer Vision. – 2019. – Pp. 6708–6717.
11. Yuan J. and He Z. Ensemble generative cleaning with feedback loops for defending adversarial attacks // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. – 2020. – Pp. 581–590.
12. Qin Y., Frosst N., Sabour S., Raffel C., Cottrell G., Hinton G. Detecting and diagnosing adversarial images with class-conditional capsule reconstructions // ICLR. – 2020. ArXiv abs/1907.02957.
13. Wicker M., Huang X., Kwiatkowska M. Feature-Guided Black-Box Safety Testing of Deep Neural Networks // TACAS 21 October 2017. – Pp. 408–426.

УДК 65.011.56

ИСПОЛЬЗОВАНИЕ СИСТЕМНОЙ ИНЖЕНЕРИИ И ИМИТАЦИОННОГО МОДЕЛИРОВАНИЯ ДЛЯ СКЛАДСКИХ СИСТЕМ

Никитина А.О.¹ (студент)

Научный руководитель – Киприянов К.В.¹

¹Университет ИТМО

e-mail: anastasianuk@mail.ru

В данной работе рассматривается использование системной инженерии в качестве построения диаграмм SysML, для возможности создания имитационной модели AnyLogic склада и материальных потоков между другими помещениями на предприятии.

Ключевые слова: системная инженерия, SysML, AnyLogic, цифровой двойник, склад, имитационное моделирование, моделирование производственных процессов.

Назначение системной инженерии заключается в руководстве разработкой сложных, комплексных систем. Система определяется как совокупность взаимосвязанных компонентов, которые работают совместно для достижения общей цели [1]. Предметно-ориентированный язык моделирования систем SysML подходит для реализации задач системной инженерии, так как наиболее точно описывает происходящие в нашей работе процессы и взаимодействия [2]. Разработка диаграммы использования, диаграмм видов деятельности, диаграммы требований и диаграммы блоков упрощает начало создания модели склада и складских операций [3]. Так же для создания модели существует необходимость в разработке диаграммы бизнес-процессов (рис. 1).

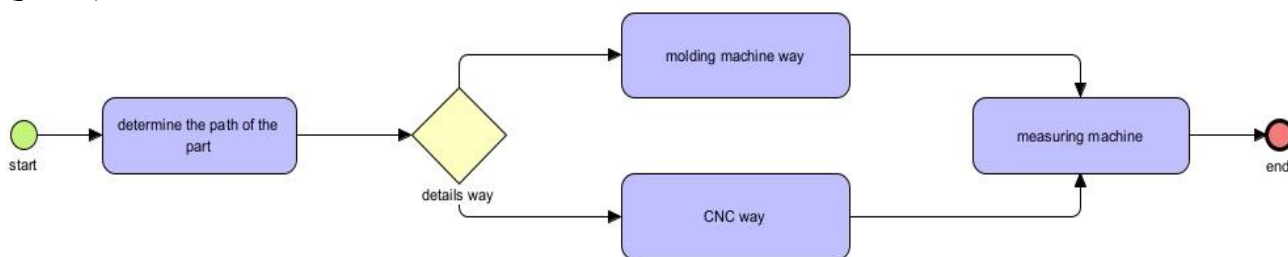


Рис. 1. Бизнес-процесс «Работа лаборатории»

Имитационное моделирование

Метод имитационного моделирования является одним из самых мощных и наиболее эффективных методов исследования процессов и производственных систем. Имитационная модель должна отражать большое число параметров, логику и закономерности поведения моделируемого объекта [4].

Проведя исследования и анализ с различными системами, среда моделирования AnyLogic оказалась самой подходящей для моделирования производственных процессов. Данная среда разработки не ограничивает только одной парадигмой моделирования, позволяя использовать различные уровни абстрагирования, различные стили и концепции, строить модели в рамках той или иной парадигмы и смешивать их при создании одной и той же модели, использовать ранее разработанные модули, собранные в библиотеки, дополнять и строить свои собственные библиотеки модулей. Благодаря встроенным возможностям анимации и наглядной

визуализации результатов, в процессе работы модели позволяют понять суть процессов, происходящих в моделируемой системе, упростить отладку модели. Среда моделирования AnyLogic поддерживает проектирование, разработку, документирование модели, выполнение компьютерных экспериментов с моделью, включая различные виды анализа.

Создание имитационной модели в среде AnyLogic начнем с подготовки графической разметки пространства и создания простого процесса, который имитирует поступление ящиков с деталями на склад [5]. Для создания диаграммы процесса нам понадобятся блоки Библиотеки моделирования процессов:

1. Source (sourceDetails) – для создания агентов, используется для обозначения начала диаграммы процесса.
2. Rack store (rackDetails) – помещает агента в ячейку заданного стеллажа или зоны хранения.
3. Delay (waiting) – задерживает агентов на заданный период времени.
4. Rack pick (pickDetails) – извлекает агента из ячейки стеллажа или зоны хранения и перемещает его в заданный узел сети.
5. Sink – для уничтожения агентов, используется для обозначения завершения процесса.

Соединив блоки, получится диаграмма процесса, которую мы назвали «source details» (рис. 2). В ней создаются агенты, перемещаются на складской стеллаж в заданную ячейку, и после установленной задержки извлекаются и перемещаются в заданный узел сети.

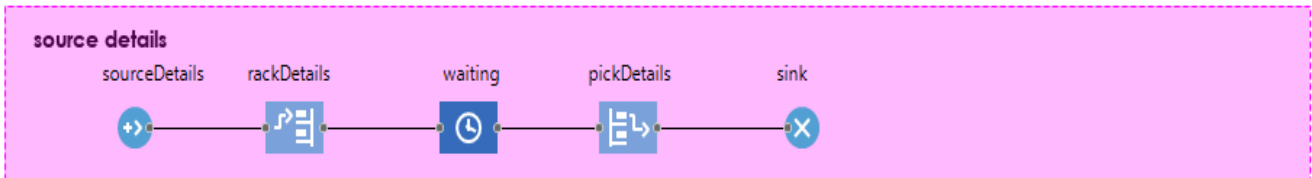


Рис. 2. Диаграмма процесса «source details»

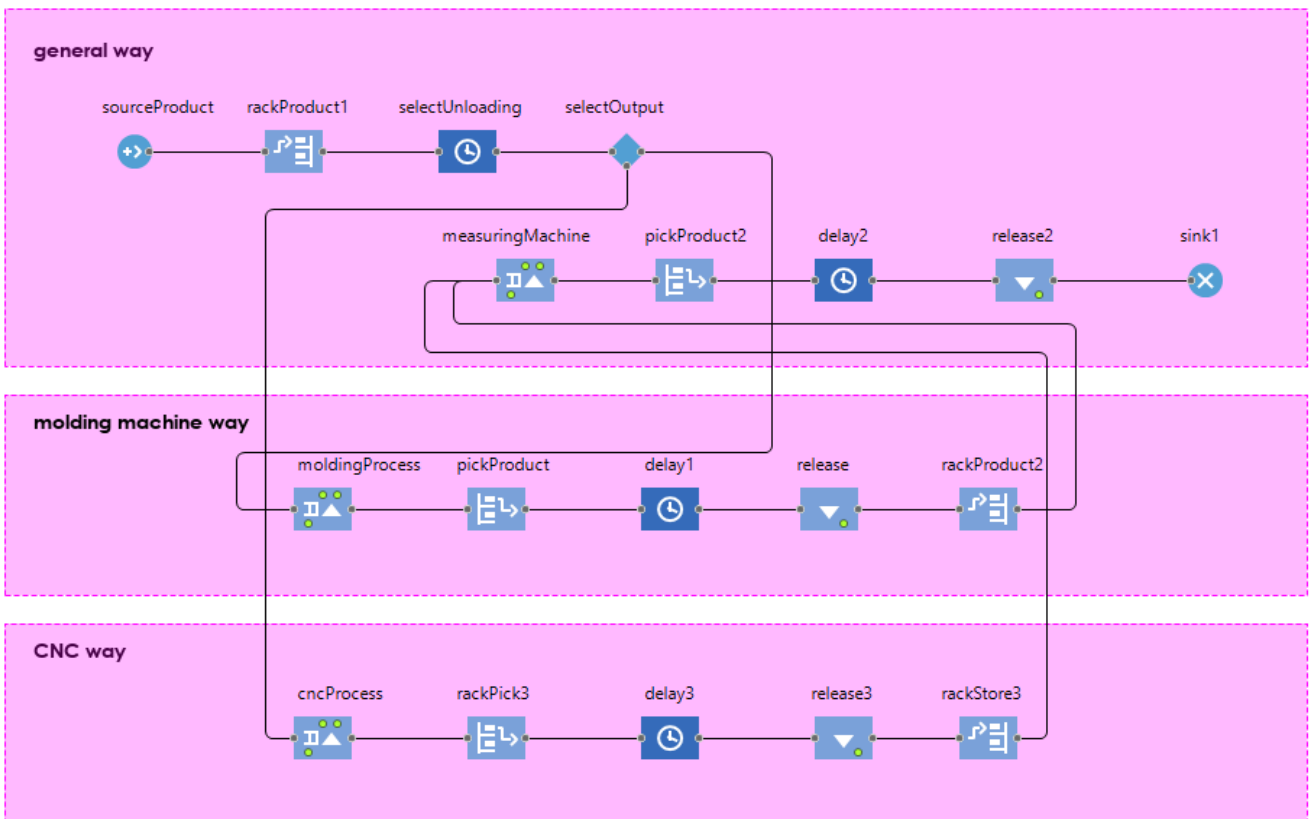


Рис. 3. Диаграмма процессов «molding machine way» и «CNC way»

Далее создана диаграмма процесса (рис. 3), в которой извлеченных агентов со стеллажей будут направлять на путь одного из двух процессов лаборатории: «molding machine – measuring machine» или «CNC lab – measuring machine». Для создания этого процесса нам понадобятся те же блоки Библиотеки моделирования процессов, а также:

1. Seize (moldingProcess, cncProcess, measuringMachine) – захватывает для агента заданное количество ресурсов.
2. Release – освобождает заданное количество ресурсов, ранее захваченных агентом с помощью блока Seize.
3. Select Output – блок направляет входящих агентов в один из двух выходных портов в зависимости от выполнения заданного (детерминистического или заданного с помощью вероятностей) условия.

Диаграммы процессов готовы. Далее необходимо графически отобразить пути, по которым будут двигаться агенты (рис. 4).

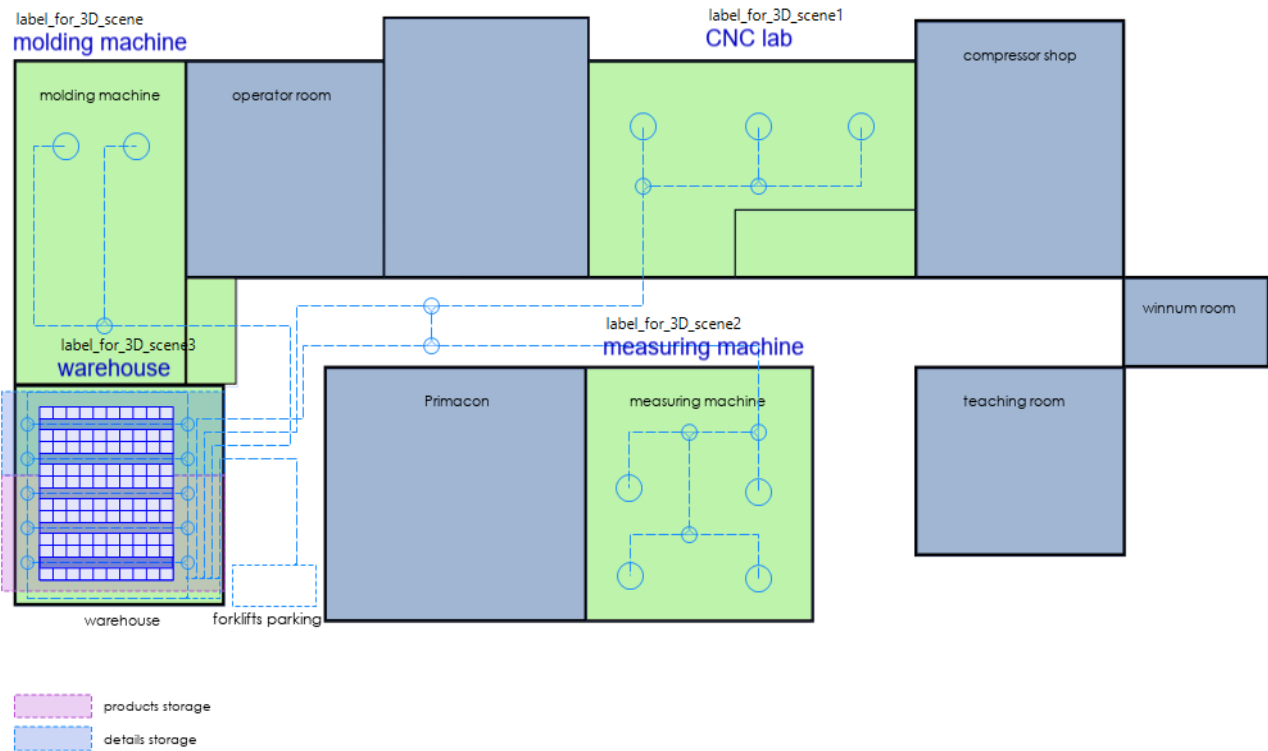


Рис. 4. Логистическая схема лаборатории



Рис. 5. Resources

Создаем следующие ресурсы (рис. 5):

1. Racksystemdetails – система стеллажей для деталей.
2. Racksystemproduct – система стеллажей для изделий.
3. Forklifts – погрузчики.
4. Molding – литьевая машина.
5. CNC – станок ЧПУ.
6. Measuring – измерительная машина.

Далее, создаем агентов, и добавляем каждому агенту трехмерную анимацию (рис. 6).

Создав имитационную модель материальных потоков выполнения складских задач нам необходимо импортировать ее для дальнейших взаимодействий.

Функционал AnyLogic позволяет бесплатно импортировать созданную модель в Cloud (рис. 7).



Рис. 6. Agents

Рис. 7. Симуляция модели в Cloud

Построение SysML-диаграмм помогло подойти к этапу создания имитационной модели работы склада. Созданная имитационная модель склада лаборатории в среде разработки AnyLogic отражает работу предприятия и позволяет отслеживать материальные потоки между другими помещениями на предприятии.

Литература

1. Системная инженерия. Введение в SysML. [Электронный ресурс]. – Режим доступа: <http://www.cs.vsu.ru/~svv/se/lec6.pdf> (дата обращения: 20.01.2022).
2. Business process diagram. [Электронный ресурс]. – Режим доступа: <https://www.ibm.com/docs/kk/elm/6.0.2?topic=processes-business-process-diagram-overview> (дата обращения: 18.01.2022).
3. Visual Paradigm. [Электронный ресурс]. – Режим доступа: <https://www.visual-paradigm.com/> (дата обращения: 18.01.2022).
4. Киселёва М.В. Имитационное моделирование систем в среде AnyLogic: учебно-методическое пособие, Екатеринбург: УГТУ – УПИ. – 2009. – 88 с.
5. Блог AnyLogic. [Электронный ресурс]. Режим доступа: <https://www.anylogic.ru/blog> (дата обращения: 27.01.2022).

УДК 778.64

ПРИМЕНЕНИЕ АДДИТИВНЫХ ТЕХНОЛОГИЙ В ПРОТОТИПИРОВАНИИ И ТОПОЛОГИЧЕСКОЙ ОПТИМИЗАЦИИ В РОБОТОТЕХНИКЕ

Павлов В.А.¹ (студент), Мордвинцева Ю.А.¹ (студент), Кофонов Р.В.¹ (студент)

¹Университет ИТМО

e-mail: viktor.pavlov.1998@list.ru, ulia190799@gmail.com, 89995264987@mail.ru

Статья рассматривает проблему выбора технологии 3D-печати для топологически оптимизированных деталей в робототехнике, на основании проведенных экспериментов на испытательной машине Shimadzu AG-50kNXD и актуальных данных испытания из русских и иностранных источников.

Ключевые слова: аддитивные технологии, аддитивные технологии в робототехнике, 3D-печать, FDM, SLS, SLM.

Под топологически оптимизированными деталями подразумевается деталь, которая обладает требуемыми механическими свойствами при минимальной массе. Примером такой модели может служить сустав робота (рис. 1).

Для моделирования топологически оптимизированных деталей применяют CAE программы (Ansys Mechanical) и CAE модули, встроенные в CAD программах (Solidworks Simulation, Fusion 360 Simulation, Autodesk Inventor).

Для изготовления деталей такой сложной формы в серийном производстве рационально применять аддитивные технологии, то есть технологии, которые позволяют послойно синтезировать деталь.

В рамках статьи рассматриваются основные типы 3D-печати – FDM, SLA, SLS, SLM. FDM – технология послойного наплавления материала, ключевые параметры от которых зависят механические свойства: толщина стенки в мм; скорость перемещения по осям x, y, z в мм/с; шаблон заполнения на рисунке 2, процент заполнения, ориентация детали в рабочем пространстве.

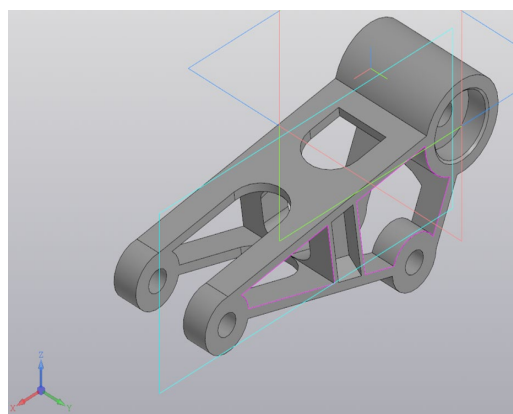


Рис. 1. Топологический оптимизированный сустав робота паука

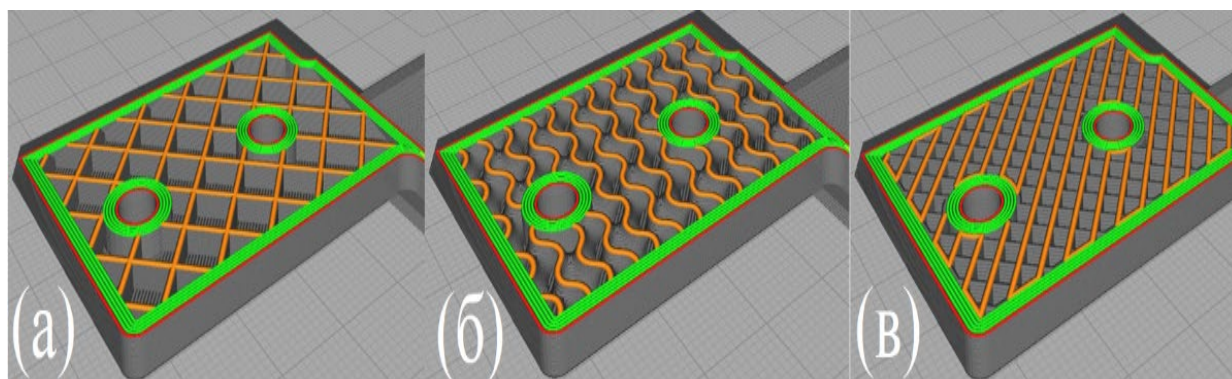


Рис. 2. Наиболее часто используемые шаблоны заполнения модели:
а) сетка; б) героид; в) зиг – заг

На основании экспериментов, проведенных на испытательной машине Shimadzu AG-50kNXD. Были получены данные, приведенные в таблице 1. Очевидно, что от ориентации детали зависят механические характеристики. В таком случае можно сделать вывод о том, что при использовании FDM технологии у синтезированной детали в явном виде присутствует анизотропия свойств. Средняя точность (поле допуска линейных размеров) печати составила ± 50 мкм, для печати был использован принтер Flying Bear Ghost 5 с закрытой камерой и материал PETG.

Таблица 1. Данные испытаний FDM лопаток ГОСТ ISO 37–2013

Вертикальная ориентация, Предел прочности на изгиб (МПа)	Горизонтальная ориентация, предел прочности на изгиб (МПа)
78,8	95,5
72,5	92,6
74,2	96,1
74,6	94,4
73,7	93,9
74,9	94,5
77,1	95,1
75,2	97,2
76,5	96,3
77,0	94,2
74,9	95,5
75,2	93,8
76,5	96,4
Среднее значение: 75,5	Среднее значение: 95,0

SLA технология предполагает использование для синтеза конструкции ректопластичные полимерные материалы, полимеризация которых идет за счет УФ – излучения. Согласно данным, полученным в результате эксперимента в таблице 2, можно сделать вывод о том, что детали, синтезированные SLA технологией, имеют низкую анизотропию свойств в сравнении с FDM. Средняя точность печати данной технологии ± 30 мкм, для печати образцов был использован принтер Anycubic Photon Monop и смола Anycubic.

Таблица 2. Данные испытаний SLA лопаток ГОСТ ISO 37–2013

Предел прочности при изгибе (МПа)
115,3
112,4
105,9
109,1
111,7
Среднее значение: 110,9



Рис. 3. Топологически оптимизированная конструкция куба для проведения испытаний на сжатие

SLS технология предполагает синтез конструкции за счет селективного лазерного спекания, данная технология находится уже в другом ценовом сегменте – размер рабочей области не превышает 750x1000x1000 мм. Для печати данной технологии могут использоваться наполненные композитные материалы до 40%. В качестве наполнителя используют рубленое волокно, а матрица представляет собой термопластичный полимерный материал. Данная технология позволяет печатать более сложные топологически оптимизированные детали, пример такой конфигураций детали приведен на рисунке 3. Средняя точность данной технологий составляет ± 20 мкм. Предел прочности при растяжении не превышает 50 МПа [1].

SLM технология предполагает синтез конструкции на подложке за счет селективного лазерного спекания, в данном случае используется уже металлический порошок: сплавы

алюминия легированные магнием и кремнием, никелевые сплавы (IN625, IN718), титановые сплавы, кобальтовые сплавы легированные молибденом и хромом, нержавеющие и инструментальные стали, например 4X5MФС (AISI 1.2344). Предел прочности для детали полученной технологией SLM составляет 660 МПа для стали AISI 316L (аналог 03X17H14M3) [2].

Полученные и собранные данные могут быть использованы для разработки макетов, прототипов и полноценных топологически оптимизированных деталей роботов.

Литература

1. Jyothish Kumar L., Pulak M. Pandey, David Ian Wimpenny 3D Printing and Additive Manufacturing Technologies, Springer Nature Singapore Pte Ltd. – 2019. – №1. – Pp. 217–222.
2. Грач С.М., Погорелко Н.А., Яшнов В.А. Расчет лучевых траекторий и поляризации радиоволн декаметрового диапазона для условий воздействия на ионосферу мощным радиоизлучением // Вестник Нижегородского университета им. Н.И. Лобачевского. – 2012. – №5(1). – С. 43–50.

УДК 004.4

РАЗРАБОТКА TELEGRAM-БОТА «ОПРЕДЕЛЕНИЕ УРОВНЯ ГОТОВНОСТИ ТЕХНОЛОГИИ»

Павлов В.А.¹ (студент), Полицинский А.С.¹ (студент), Макаренко З.В.¹ (студент), Сунцова Д.И.¹ (студент), Бахолдин П.П.¹ (студент)

Научные руководители – канд. физ.–мат. наук, доцент Кремлев А.С.¹, канд. физ.–мат. наук, доцент Маргун А.А.¹

¹Университет ИТМО

e-mail: viktor.pavlov.1998@list.ru

Работа выполнена в рамках темы НИР №620164 «Методы искусственного интеллекта для киберфизических систем».

В статье представлен Telegram-бот, который может помочь пользователю определить уровень готовности технологии (УГТ), основываясь на ГОСТ Р 57194.1-2016. Рассмотрены и проанализированы аналогичные варианты автоматизации оценки УГТ. Представлен алгоритм работы бота, его исполнение и внешний вид интерфейса. Для оценки корректности работы бота взят проект, для которого УГТ определен вручную, после чего это значение сопоставлено с полученным заключением от бота. В результате бот подтвердил оценку, сделанную экспертами.

Ключевые слова: уровень готовности технологий (УГТ), второе мнение, Technology Readiness Level (TRL), автоматизация оценивания готовности, развитие проекта, динамика развития, технологии, инновации.

Оценка уровня готовности технологии в наше время играет немалую роль для развития новых перспектив в различных областях, помогая оценить экономические риски будущего проекта, а также его актуальность и новизну. При помощи различных методик оценки УГТ, которые рассматриваются в [1–3], разработчик может контролировать процесс создания технологии, а также претендовать на различные виды финансирования, в том числе и от государства [4].

В зависимости от методики и области оценивания, шкалы УГТ могут различаться, но в общем случае она имеет вид, представленный в таблице 1 [5].

Для автоматизации и ускорения оценки уровня готовности разрабатываются различные калькуляторы УГТ [6–11], которых существует большое множество, но они имеют ряд недостатков, в частности, основной проблемой является то, что они написаны под иностранные стандарты и не переведены на русский язык.

Таблица 1. Шкала уровней готовности технологии

Наименование уровня	Описание уровня
УГТ 1	Сформулирована базовая идея. Выявлены и опубликованы фундаментальные принципы
УГТ 2	Сформулирована техническая концепция и/или предполагаемые применения возможных концепций для перспективных объектов
УГТ 3	Даны аналитические и экспериментальные подтверждения по важнейшим функциональным возможностям и/или характеристикам выбранной концепции
УГТ 4	Компоненты и/или макеты проверены в лабораторных условиях
УГТ 5	Компоненты и/или макеты подсистем верифицированы в условиях, близких к реальным

Наименование уровня	Описание уровня
УГТ 6	Модель или прототип системы/подсистемы продемонстрированы в условиях, близких к реальным
УГТ 7	Прототип системы прошел демонстрацию в эксплуатационных условиях
УГТ 8	Создана штатная система и освидетельствована (квалифицирована) посредством испытаний и демонстраций. Технология проверена на работоспособность в своей конечной форме и в ожидаемых условиях эксплуатации в составе технической системы (комплекса)
УГТ 9	Продемонстрирована работа реальной системы в условиях реальной эксплуатации

Например, известен калькулятор, реализованный в программе Microsoft Excel, разработанный AFRL (Air Force Research Laboratory – исследовательская лаборатория ВВС США), который представляет собой два листа электронной таблицы Excel: главный лист – «TRL Calculator» и итоговый лист – «Summary». Его интерфейс представлен на рисунке 1. Пользователь отвечает на вопросы, заполняя анкету, после чего калькулятор выдает присвоенный уровень готовности и дает наглядную картину в процентном соотношении выполненных условий для перехода на следующий уровень [12].

Хотя данный калькулятор многофункционален и позволяет достаточно точно определить УГТ, он имеет ряд недостатков, основным из которых является его недостаточная оптимизация: при работе в данном калькуляторе наблюдались частые сбои приложения. Также существенными минусами являются довольно сложный интерфейс, потребность в экспертизе для его использования и отсутствие поддержки русского языка.

Criar Novo Produto

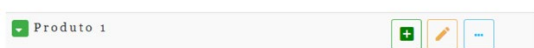


Рис. 2. Интерфейс онлайн-калькулятора IMATEC

проект, в котором отмечает наличие конкретных документов или этапов.

Без заполнения нужного количества положительных ответов калькулятор не даст перейти на следующий уровень готовности. Хотя данный калькулятор достаточно прост, но также не имеет перевода на русский язык и ориентирован на зарубежные стандарты.

Из-за низкой распространенности калькуляторов УГТ часто прибегают к оценке экспертами, что подразумевает значительный объем работы, т.к., например, стандарт [13] предлагает опросник из 274 показателей, для которых необходимо в процентном соотношении указать выполнение каждого пункта. Это вносит в оценку готовности технологии субъективную составляющую, ведь само оценивание будет зависеть от знаний и аналитических способностей эксперта.

Для того, чтобы нивелировать ряд рассмотренных проблем, сделать оценку УГТ доступной и удобной, в данной работе предлагается калькулятор УГТ, реализованный в виде бота в распространенном мессенджере Telegram.

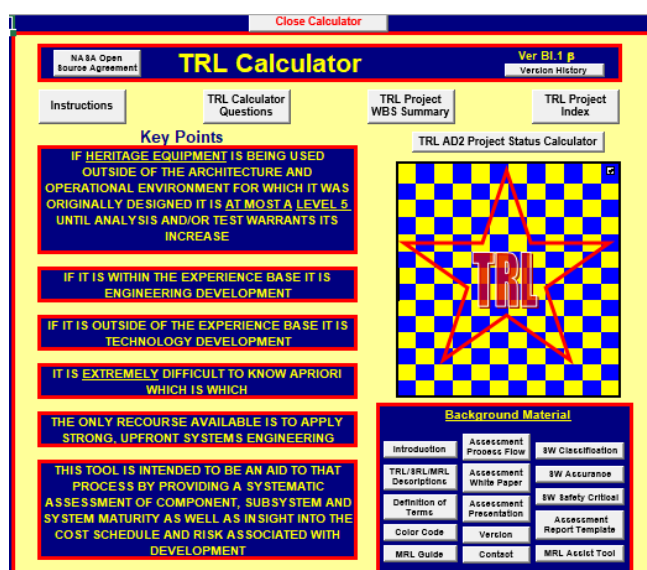


Рис. 1. Стартовое окно калькулятора УГТ, разработанного AFRL

Также известен калькулятор, который можно использовать онлайн через веб-сайт, от разработчиков IMATEC [9]. Его интерфейс представлен на рисунке 2. Оценка УГТ происходит следующим образом: пользователь создает новый

использовать онлайн через веб-сайт, от разработчиков IMATEC [9]. Его интерфейс представлен на рисунке 2. Оценка УГТ происходит следующим образом: пользователь создает новый

Описания калькулятора УГТ, реализованного в виде Telegram-бота

На рисунках 3 – 5 представлен алгоритм бота.

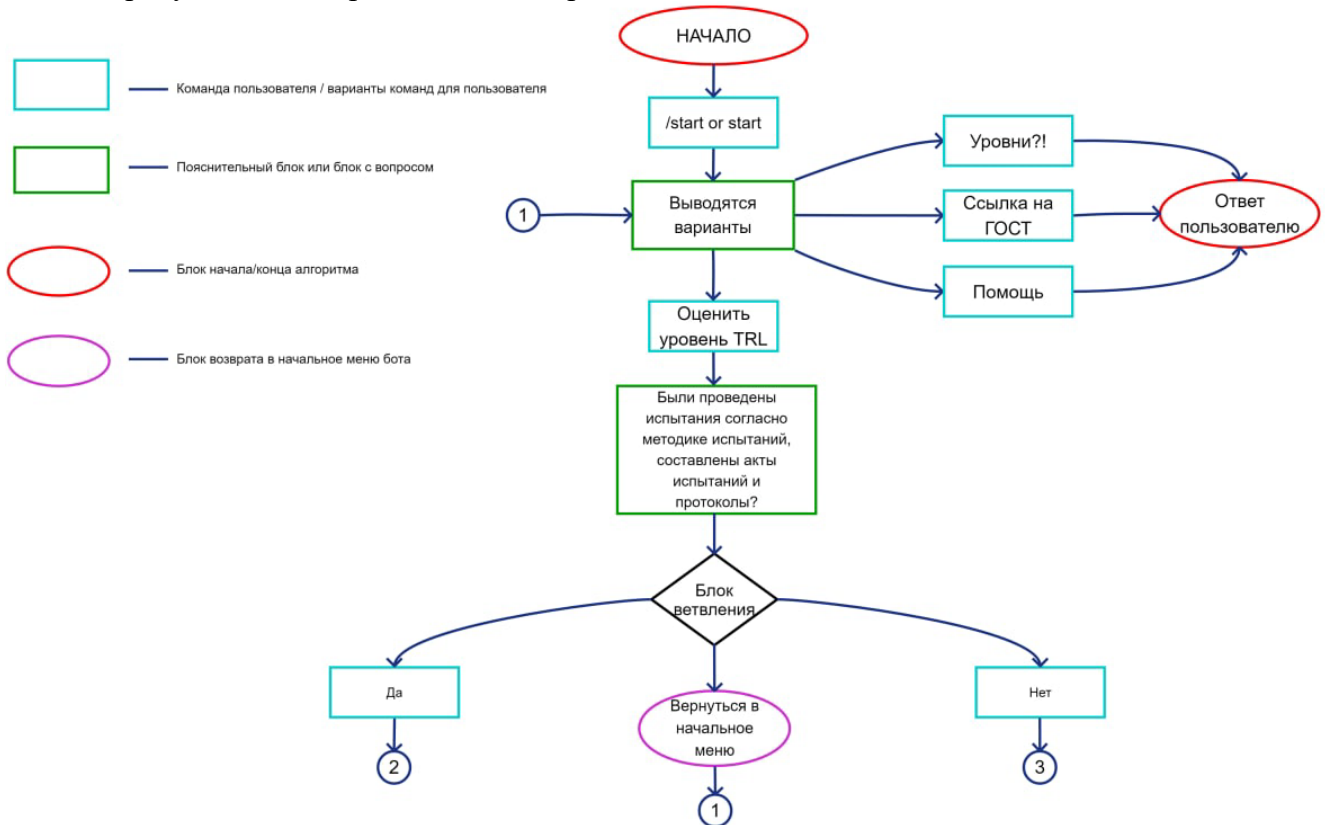


Рис. 3. Главная ветвь алгоритм УГТ-бота

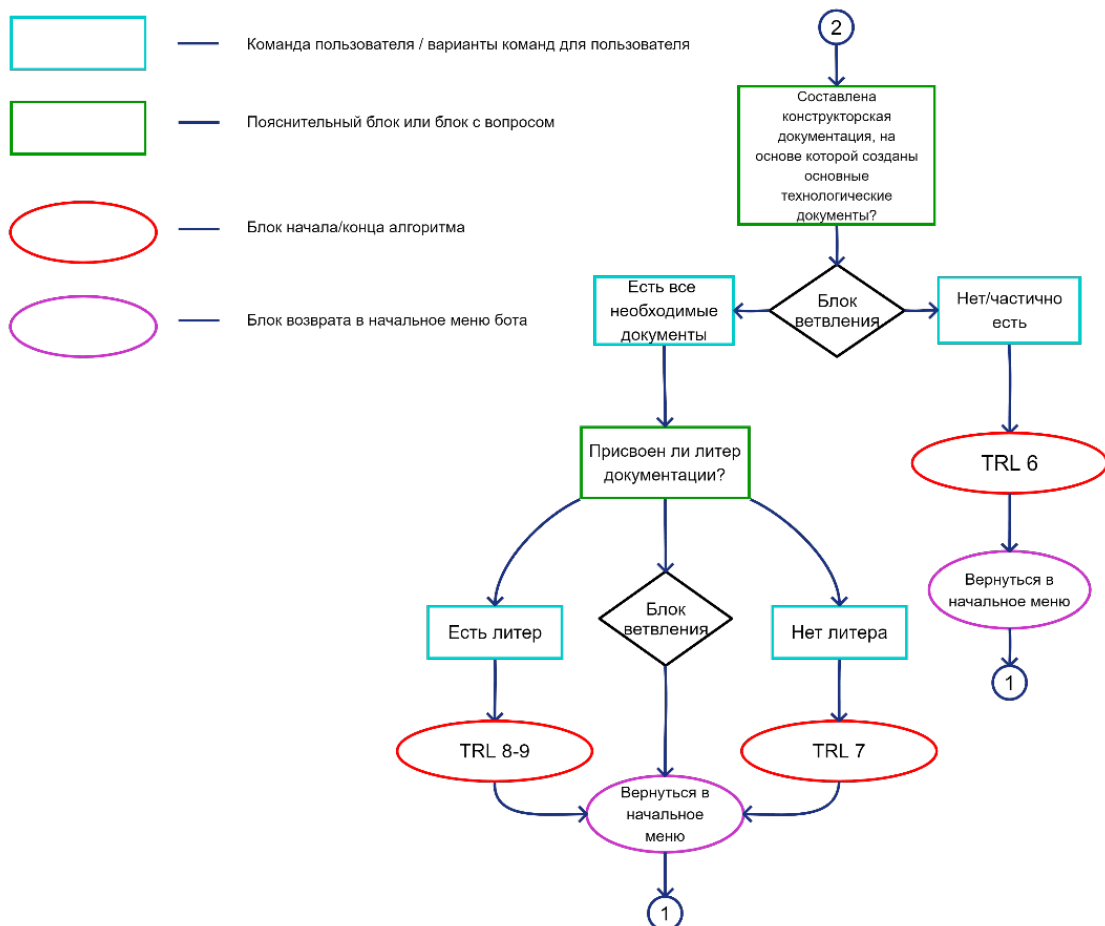


Рис. 4. Вторая ветвь алгоритма УГТ-бота

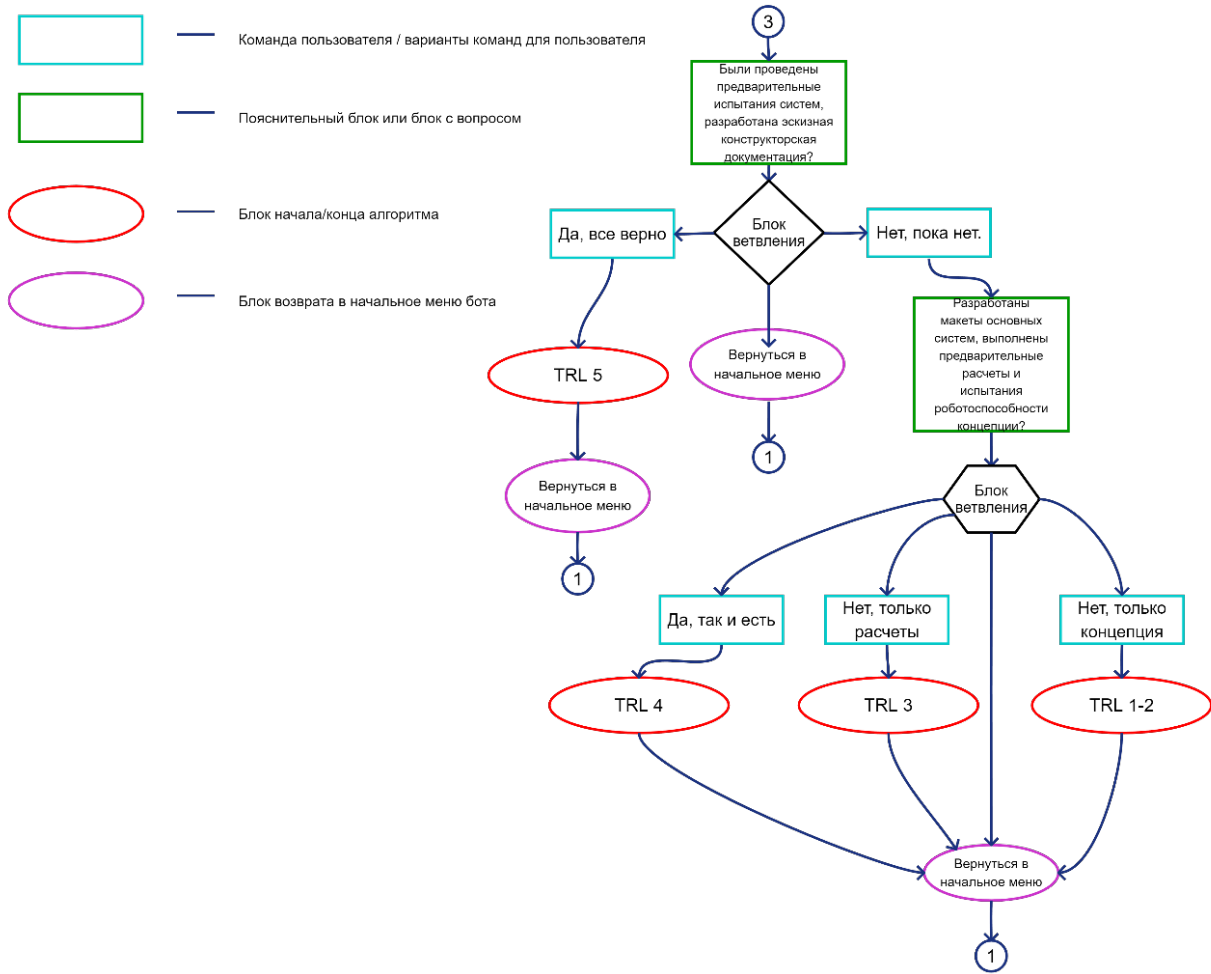


Рис. 5. Третья ветвь УГТ-бота

Пользователю через меню, представленное на рисунке 6, предлагается:

1. Провести оценку УГТ.
2. Получить информацию о том, какие бывают уровни готовности.
3. Получить ссылку на отечественный ГОСТ.

Получить помощь и инструкцию по использованию калькулятора.

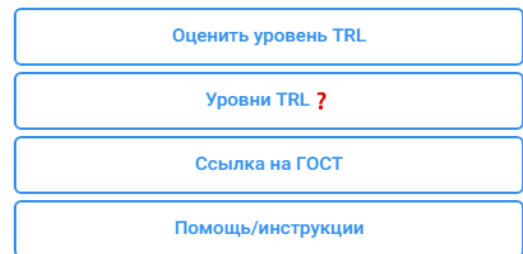


Рис. 6. Стартовое меню калькулятора

Сама оценка УГТ реализована в виде опросника, в котором пользователю предлагается обозначить наличие или отсутствие нужных пунктов для присвоения уровня готовности. Если на текущий вопрос поступает отрицательный ответ, то уровень готовности снижается. Если для текущего уровня готовности все ответы положительные, то происходит переход на следующий уровень, где пользователю предстоит снова подтвердить наличие или недостаток пунктов, которые нужны для перехода на этот уровень. Таким образом, в результате прохождения опроса бот сообщает, какой уровень готовности можно присвоить разрабатываемой технологии. На рисунке 7 представлен процесс прохождения опроса. Ключевые показатели, на основании которых бот принимает решение, соответствуют рекомендациям стандарта (табл. 2) [14].

Оценка УГТ играет большую роль в реализации инновационных проектов, позволяя разработчику и заказчику контролировать продвижение проекта, а также оценивать возникновение различных экономических рисков.

Представленный калькулятор УГТ, реализованный в виде Telegram-бота, направлен на автоматизацию и ускорение оценки УГТ, доступность использования и поддержку отечественного стандарта. Он также не требует скачивания отдельного приложения и не имеет проблем с оптимизацией.

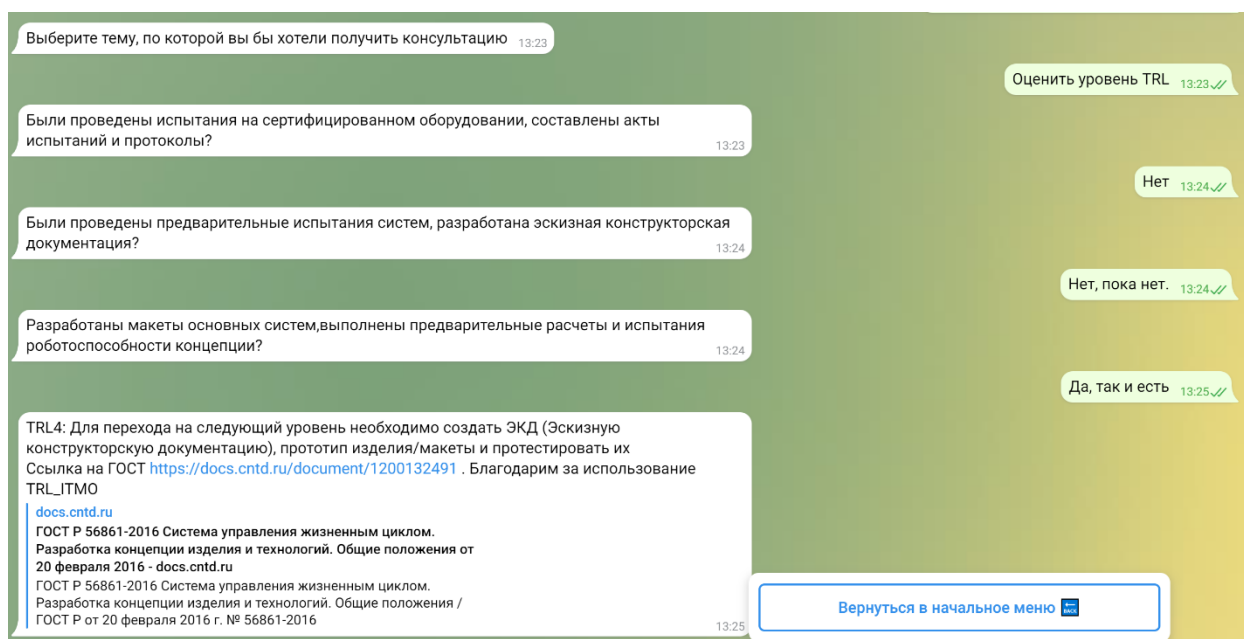


Рис. 7. Прохождение опроса пользователем

Таблица 2. Соответствие показателей для принятия решения ботом стандарту [14] на примере уровня УГТ 4

Уровень УГТ	Ключевые показатели для принятия решения ботом	Описание и примерный состав работ согласно стандарту [14]
4	<ul style="list-style-type: none"> разработаны макеты основных систем выполнены предварительные расчеты испытана работоспособность концепции 	<ul style="list-style-type: none"> продемонстрирована работоспособность на подробных макетах применены трехмерные модели с высоким масштабом и точностью моделирования подробно описана конструкция устройства на основе новой технологии

На данном этапе калькулятор не может претендовать на высокую объективность и является исключительно источником второго мнения. Но в перспективе планируется увеличить функционал калькулятора, чтобы можно было оценить не только уровень готовности непосредственно технологии, но также подключить уровень готовности производства и оценку экономических рисков. Кроме того, стоит задача усовершенствования самого опросника УГТ для увеличения объективности оценки.

Литература

1. Комаров А.В., Пихтарь А.Н., Гриневский И.В., Комаров К.А., Голицын Л.В. Концептуальная модель оценки технологической готовности научно-технологического проекта и его потенциала на ранних стадиях разработки // Экономика науки. – 2021. – №7(2). – С. 111–134.
2. Сартори А.В., Сушков П.В., Манцевич Н.М. Принципы бережливого управления исследованиями и разработками на основе методологии уровней готовности инновационного проекта // Экономика науки. – 2019. – Т.6. – №1–2. – С. 22–34.
3. Сартори А.В., Гареев А.Р., Ильина Н.А., Манцевич Н.М. Применение подхода уровней готовности для различных предметных направлений в бережливом НИОКР // Экономика науки. – 2020. – Т.6. – №1–2. – С. 118–134.
4. Хаматханова А.М. Готовность к промышленному внедрению как индикатор выбора приоритетных технологических направлений // Экономика науки. – 2016. – Т.2. – №1. – С. 23–34.
5. Гранич В.Ю., Дутов А.В., Мирошкин В.Л., Сыпало К.И. Об уровнях готовности технологий и применении Калькулятора УГТ для их оценивания // Экономика науки. – 2020. – Т.6. – №1–2. – С. 6–10.

6. DAU Tools. [Электронный ресурс]. – Режим доступа: <https://www.dau.edu/cop/stm/lists/tools/allitems.aspx> (дата обращения: 02.02.2022).
7. Дмитренко И.П. Вариант калькулятора уровней готовности технологий повышенной точности // Актуальные проблемы гуманитарных и естественных наук. – 2016. – №12-1. – С. 103–107.
8. Дмитренко И.П. Вариант калькулятора уровней готовности технологий повышенной точности на базе открытого программного продукта // Актуальные проблемы гуманитарных и естественных наук. – 2017. – №1-1. – С. 64–66.
9. Xavier Jr A., Veloso A., Souza J., Kaled Cás P., Cappelletti C. AEB Online Calculator for Assessing Technology Maturity: IMATEC // Journal of Aerospace Technology and Management. – 2020. – №12. DOI:10.5028/jatm.v12.1098ю.
10. Altunok T., Cakmak T. A technology readiness levels (TRLs) calculator software for systems engineering and technology management tool // Advances in Engineering Soft-ware. – 2010. – №41. – Pp. 769–778.
11. Жебель В.В. и др. Программное средство для комплексной оценки технологической готовности инновационных научно-технологических проектов // Экономика науки. – 2018. – Т.4. – №1. – С. 58–68.
12. Calculator TRL AFRL Version 2.2. [Электронный ресурс]. – Режим доступа: <http://p0.storage.canalblog.com/06/32/447324/60486844.xls> (дата обращения: 02.02.2022).
13. ГОСТ Р 58048-2017. Трансфер технологий. Методические указания по оценке уровня зрелости технологий. – Введен 01.06.2018. – М.: Стандартинформ. – 2018. – 41 с.
14. ГОСТ Р 56861-2016. Система управления жизненным циклом. Разработка концепции изделия и технологий. Общие положения. – Введен 01.10.2016. – М.: Стандартинформ. – 2019. – 15 с.

УДК 620.179.16; 004.94

МОДЕЛЬ РАЗДЕЛЬНО-СОВМЕЩЕННОГО ОПТИКО-АКУСТИЧЕСКОГО ПРЕОБРАЗОВАТЕЛЯ

Сергеев А.С.¹ (студент), Алифанова И.Е.¹ (аспирант)Научный руководитель – канд. техн. наук Кинжагулов И.Ю.¹¹Университет ИТМО

e-mail: sherbairina@mail.ru, kinzhiki@mail.ru

Работа выполнена в рамках темы НИР №620164 «Методы искусственного интеллекта для киберфизических систем».

Для возбуждения акустических колебаний применяется раздельно-совмещенный оптико-акустический преобразователь и лазерно-ультразвуковой дефектоскоп. Конструкция раздельно-совмещенного оптико-акустического преобразователя должна обеспечивать точность измерений времени достижения продольной подповерхностной волной приемника акустических колебаний. Для анализа регистрируемых акустических сигналов и выделения из них сигнала продольной подповерхностной волны в данной работе предложена и разработана конечно-элементная модель акустического тракта раздельно-совмещенного оптико-акустического преобразователя.

Ключевые слова: остаточные напряжения, акустический метод, неразрушающий контроль, продольная подповерхностная волна, компьютерное моделирование.

В настоящее время ультразвуковые методы контроля занимают одно из ведущих мест не только в дефектоскопии изделий, но и в оценке структуры и прочностных характеристик материалов, напряженно-деформированного состояния изделий. Одним из решений задачи повышения точности измерения скорости распространения ультразвуковых волн является применение метода, основанного на термооптическом возбуждении ультразвуковых колебаний [1]. Для реализации метода используется лазерно-ультразвуковой дефектоскоп с оптико-акустическим преобразователем (ОАП).

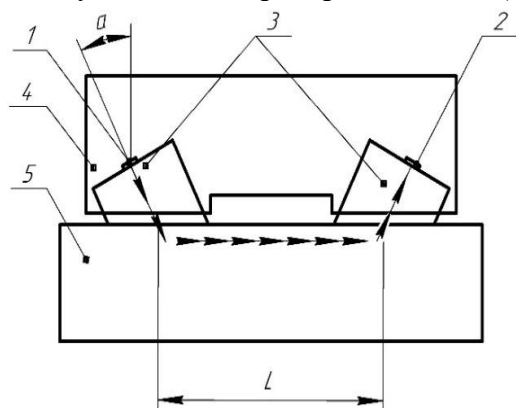


Рис. 1. Конструктивная схема оптико-акустического преобразователя (стрелками изображен путь ультразвуковых колебаний):
1 – генератор акустических колебаний;
2 – приемник акустических колебаний;
3 – преломляющая призма; 4 – корпус;
5 – объект контроля

ОАП предназначен для преобразования лазерных импульсов в акустические, передачи их в исследуемую среду и регистрации отраженных и рассеянных акустических сигналов [1] (рис. 1). В общем корпусе ОАП 4 объединены генератор и приемник ультразвуковых колебаний. Генератор 1 за счет термоакустического эффекта посредством лазерного нагрева генерирует свехкороткий (порядка десятков наносекунд) акустический импульс. С помощью преломляющей призмы 3 акустические колебания вводятся в объект контроля 5. Угол наклона призмы α соответствует минимальному углу падения продольной волны, при котором преломленная продольная волна будет распространяться по границе раздела сред, или первому критическому [2],

что является обязательным условием распространения продольной подповерхностной волны. Продольная подповерхностная волна проходит в объекте контроля 5 путь L , равный базе ОАП (рис. 1), и попадает на чувствительный элемент приемника 2.

Конструкция ОАП должна обеспечивать точность измерений времени достижения продольной подповерхностной волной приемника акустических колебаний [3].

Целью работы является разработка конечно-элементной модели акустического тракта ОАП, которая позволит:

- анализировать регистрируемые с помощью ОАП акустические сигналы;
- выделить из всех регистрируемых сигналов сигнал продольной подповерхностной волны.

Конечно-элементная модель акустического тракта ОАП была разработана в программном комплексе COMSOL Multiphysics с помощью интерфейса Elastic Waves, Time Explicit, предназначенного для моделирования упругих волн во временной области с помощью явного решателя на основе разрывного метода Галеркина [4].

Геометрия расчетной области представлена на рисунке 2. При разработке конечно-элементной модели введены следующие допущения:

- плоская геометрия расчетной области;
- генератор акустических колебаний заменен зондирующим импульсом акустических колебаний;
- объект контроля имеет бесконечную длину.

Генератор ОАП 1 описывается функцией $F(t)$:

$$F(t) = \sin \sin (2\pi ft) \cdot \exp \left(-\frac{t}{T_0} \right)^2,$$

где f – частота ультразвуковых колебаний;

T_0 – период колебаний;

t – время распространения колебаний.

Функция $F(t)$ описывает короткий радиоимпульс, представляющий синусоидальные колебания с колоколообразной огибающей. Расчетная область 3, 4 описывается уравнением Ламе, позволяющим определить характеристики распространения произвольного упругого возмущения в изотропном твердом теле [5]:

$$\rho \frac{\partial^2 u}{\partial t^2} = (\lambda + 2\mu) \text{grad div } u + \mu \Delta u,$$

где ρ – плотность среды;

$u = (u_x, u_y)$ – вектор смещения частиц;

λ и μ – константы, характеризующие упругость среды (постоянные Ламе),

Δ – оператор Лапласа.

Все компоненты смещения представляются дифференциальными уравнениями, имеющими вид классического волнового уравнения:

$$\frac{\partial^2 u}{\partial t^2} = c^2 \frac{\partial^2 u}{\partial x^2},$$

где c – скорость распространения ультразвуковых колебаний в среде.

Для уменьшения количества отраженных сигналов программный комплекс позволяет ввести в модель идеально-поглощающие слои 5. Таким образом, вводится допущение, что объект контроля имеет бесконечную длину.

Рассмотрим результаты моделирования (рис. 3) для толщины объекта контроля 0,02 м. Интенсивность цвета отражает распределение полей перемещений $u(x, y, t)$ в каждой точке расчетной области модели.

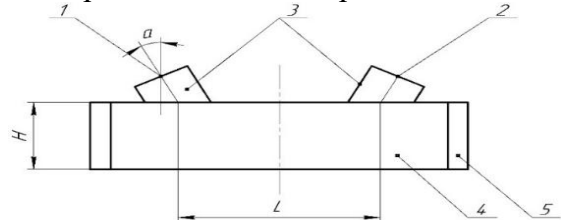


Рис. 2. Геометрия расчетной области оптико-акустического преобразователя: 1 – генератор акустических колебаний, 2 – приемник акустических колебаний, 3 – преломляющая призма, 4 – объект контроля, 5 – идеально-поглощающие слои

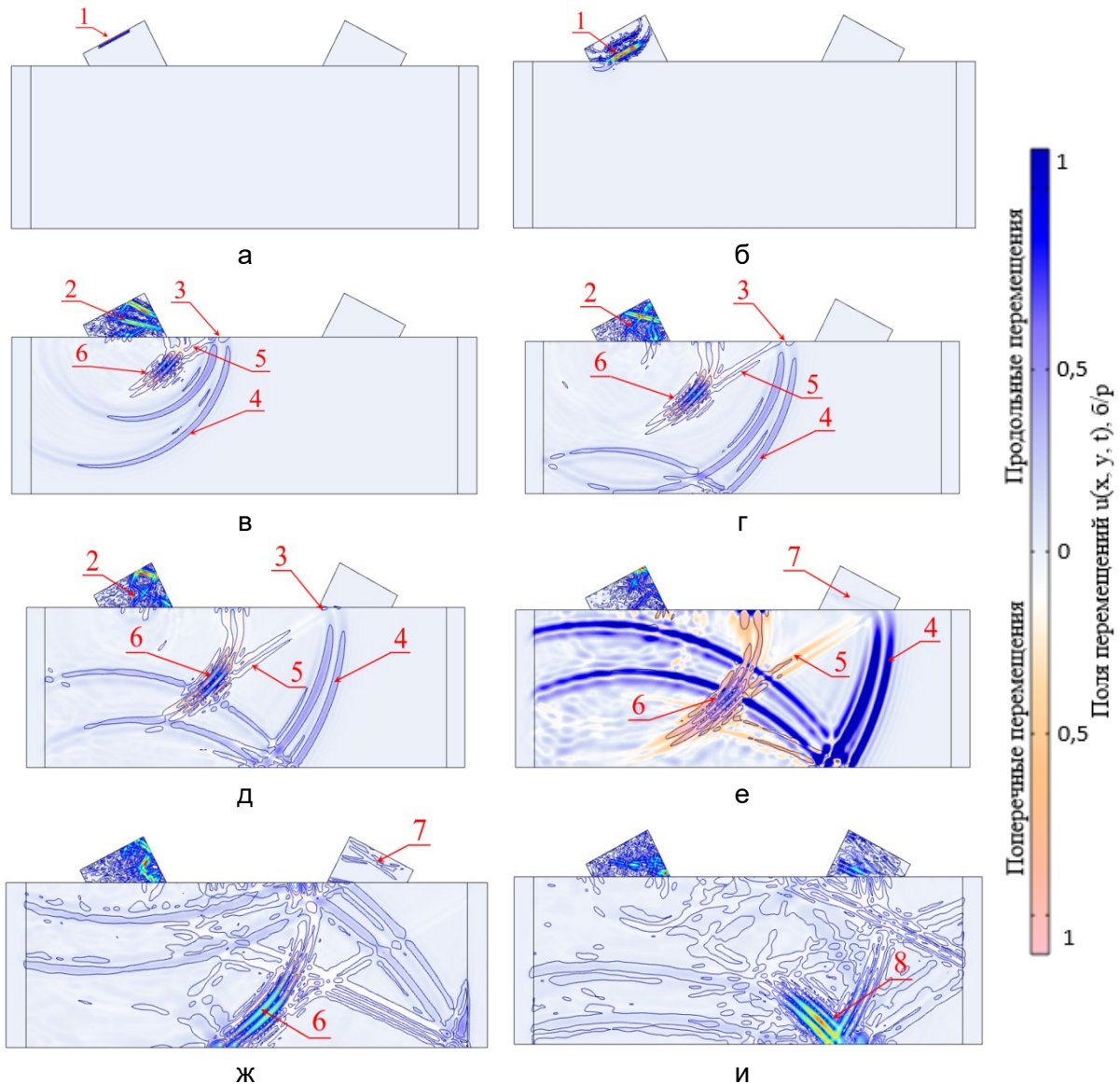


Рис. 3. Распределение полей перемещений $u(x, y, t)$ в моменты времени: а) $2 \cdot 10^{-7}$ с; б) $2 \cdot 10^{-6}$ с; в) $4,5 \cdot 10^{-6}$ с; г) $6 \cdot 10^{-6}$ с; д) $7 \cdot 10^{-6}$ с; е) $8 \cdot 10^{-6}$ с; ж) $10 \cdot 10^{-6}$ с; и) $13 \cdot 10^{-6}$ с; 1 – продольная волна в призме генератора; 2 – многократно переотражаемые волны в призме генератора, 3 – поверхностная волна; 4 – фронт продольной подповерхностной волны; 5 – поперечная волна, связанная с распространением продольной подповерхностной волны; 6 – поперечная волна, связанная с трансформацией продольной волны; 7 – продольная волна в призме приемника; 8 – поперечная волна, отраженная от донной поверхности объекта контроля

В начальный момент времени генерируется продольная УЗВ, которая распространяется в призме генератора 1 (рис. 3, а). На границе раздела призма-объект контроля наблюдается явление преломления УЗВ 1 (рис. 3, б). Частично колебания отражаются от границы раздела призма-объект контроля, затем многократно отражаются от границ призмы и т. д. 2 (рис. 3, в).

По границе раздела со скоростью продольной волны распространяется неоднородная поверхностная волна с углом ввода 90° 3 (рис. 3, в–д). Продольная волна 1 (рис. 3, а–б), падающая под первым критическим углом, и неизбежно сопутствующие ей продольные волны, падающие с углами, отличными от первого критического, в материале объекта контроля трансформируются в поле продольной подповерхностной волны 4 (рис. 3, в–е) [2].

Поле продольной подповерхностной волны влечет за собой распространение боковой поперечной волны, отходящей в нижней среде от границы под третьим [2] критическим углом 5 (рис. 3, в–е).

В результате трансформации продольной волны в объекте контроля также распространяется поперечная волна с углом ввода, равным третьему критическому углу θ (рис. 3, в–ж).

Продольная подповерхностная волна распространяется со скоростью несколько большей, чем скорость продольной волны [2] и первая достигает границы раздела объект контроля–приемник 4 (рис. 3, д). Следом за ней границы раздела достигает поверхностная волна 3 (рис. 3, д).

На границе раздела вновь наблюдаются явления преломления и трансформации, и в призме приемника распространяется только продольная волна 7 (рис. 3, е–ж) значительно меньшей амплитуды (контраст рис. 3, е) намеренно увеличен, чтобы волна 7 была различима. В некоторый момент времени продольная волна 7 достигает приемник акустических колебаний (рис. 3, ж).

Поперечная волна достигает донной поверхности объекта контроля 6 (рис. 3, ж) и отражается от нее 8. (рис. 3, и). При определенных соотношения базы L ОАП и толщины объекта контроля H, отраженная поперечная волна также может быть зафиксирована приемником ОАП.

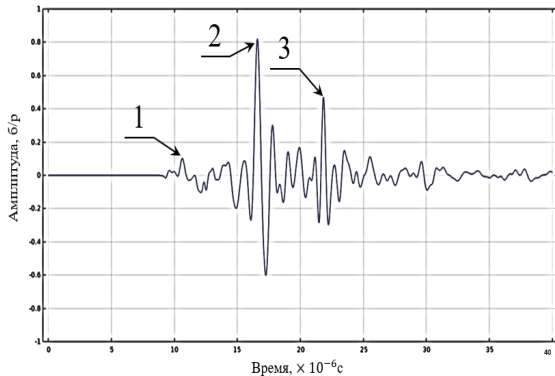


Рис. 4. А-скан, полученный по результатам моделирования на объекте контроля толщиной 0,02 м: 1 – продольная подповерхностная волна; 2 – поверхностная волна; 3 – поперечная волна, отраженная от донной поверхности

На рисунке 4 приведен пример А-скана (изменение амплитуды колебаний частиц твердого тела на приемнике ОАП по времени), полученного по результатам моделирования для объекта контроля толщиной 0,02 м.

Выполним эксперимент с использованием прямоугольных образцов со следующими параметрами: габаритные размеры – 0,07 × 0,07 м; толщины – 0,01; 0,02; 0,03; 0,04 м; материал – сталь 40Х13. Измерения проведены с помощью ОАП ПЛУ-6Н-02 (ООО «ЛИНКС 2000», Россия) (рис. 5) и лазерно-ультразвукового дефектоскопа УДЛ-2М (ООО «ЛИНКС 2000», Россия). В составе УДЛ-2М использован лазер с диодной накачкой (длительность импульса – 70–80 нс, энергия в импульсе – 100 мкДж, частота повторения импульсов – 1000 Гц).

Поочередно на каждый образец был установлен ОАП ПЛУ-6Н-02 (рис. 5) с помощью которого проведено 20 измерений времени прихода продольной подповерхностной волны на приемник. С помощью лазерно-ультразвукового дефектоскопа УДЛ-2М сформирован А-скан (рис. 6).

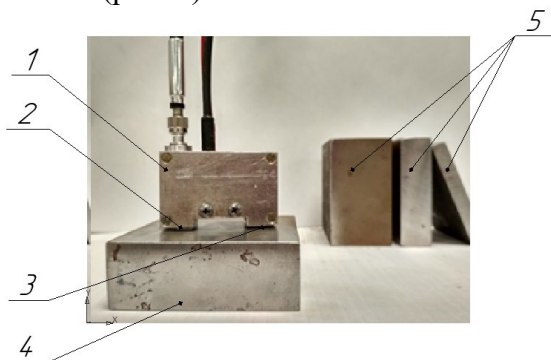


Рис. 5. Общий вид оптико-акустического преобразователя ПЛУ-6Н-02 и образцов: 1 – корпус; 2 – призма генератора; 3 – призма приемника; 4 – образец толщиной 0,03 м; 5 – образцы толщиной 0,01, 0,02 и 0,04 м

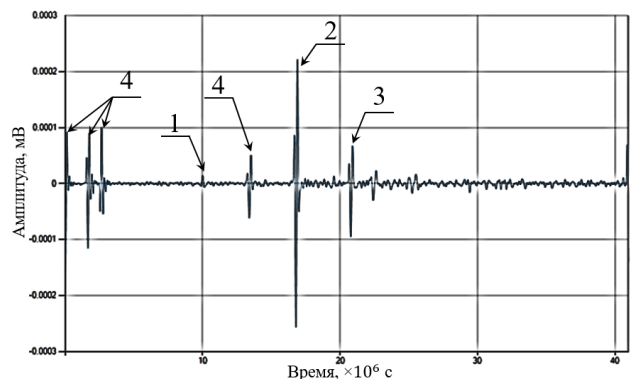


Рис. 6. А-скан, полученный в результате эксперимента при толщине образца 0,02 м: 1 – продольная подповерхностная волна; 2 – поверхностная волна; 3 – поперечная волна, отраженная от донной поверхности; 4 – собственный шум оптико-акустического преобразователя

А-сканы, полученные в результате моделирования и при проведении эксперимента (рис. 4, 6), имеют отличия, обусловленные следующими причинами:

1. Вычислительные средства не позволяют уменьшить шаг моделирования по времени, поэтому длительность сигналов, полученных при моделировании, больше, чем при проведении эксперимента.
2. Конечно-элементная модель ОАП имеет идеальную конструкцию (генератор и приемник акустических колебаний изолированы друг от друга, соединительные элементы не оказывают влияния и др.) в отличие от реального ОАП, поэтому на рисунке 6 наблюдаются собственные шумы ОАП 4 (рис. 6), которые отсутствуют на рисунке 4.
3. Справочные исходные данные для моделирования не совсем соответствуют реальным свойствам материалов преломляющей призмы ОАП и образцов, поэтому время регистрации акустических сигналов при моделировании и при проведении эксперимента отличается.

Разработанная конечно-элементная модель позволила проанализировать регистрируемые оптико-акустическим преобразователем акустические сигналы и выделить сигнал продольной подповерхностной волны. В дальнейшем модель может быть использована при проектировании новых оптико-акустических преобразователей.

Литература

1. Карабутов А.А. Лазерно-ультразвуковой дефектоскоп. Патент РФ RU2381496C1. Заявка 2008121227/28, 2008.05.29, опубл. 2010.02.10.
2. Разыграев Н.П. Физика, терминология и технология в ультразвуковой дефектоскопии головными волнами // Дефектоскопия. – 2020. – №9. – С. 3–19.
3. Alifanova I.E., Vychenok V.A., Fedorov A.V., Беркутов И.В., Khoshev A.E. Methodology for assessing the uncertainty of measurements of mechanical stresses by the ultrasonic method with the help of an optical-acoustic separate-combined transducer // Journal of Physics: Conference Series. – 2021. – V.2127. – №.1. – Pp. 1–7.
4. Comsol. DG-FEM: новая технология для масштабных расчетов распространения акустических и упругих волн в COMSOL Multiphysics®. [Электронный ресурс]. Режим доступа: <https://www.comsol.ru/video/dg-fem-new-technology-for-acoustic-and-elastic-wave-modeling-on-large-scales-in-comsol-webinar-ru>, свободный. Яз. рус. (дата обращения: 19.01.2022).
5. Тихонов А.Н., Самарский А.А. Уравнения математической физики: Учеб. пособие. М.: Наука. – 1977. – 736 с.

УДК 004.67

РАЗРАБОТКА ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ ДЛЯ ОБРАБОТКИ, ВИЗУАЛИЗАЦИИ И ХРАНЕНИЯ РЕЗУЛЬТАТОВ УЛЬТРАЗВУКОВОГО КОНТРОЛЯ

Сергеев А.С.¹ (студент)

Научный руководитель – д-р техн. наук Федоров А.В.¹

¹*Университет ИТМО*

e-mail: alex.sergeev115@gmail.com

Работа выполнена в рамках темы НИР №620164 «Методы искусственного интеллекта для киберфизических систем»

Работа посвящена разработке программного обеспечения для УЗК контроля, позволяющая производить точную визуализацию данных, полученных с дефектоскопа и поиск дефектов в выбранной пользователем области.

Ключевые слова: ультразвуковой контроль, ПО, программное обеспечение, В скан.

В рамках данной работы было разработано программное обеспечение (ПО) для ультразвукового контроля [1–4]. Данное ПО обеспечивает такие возможности, как:

- подключение дефектоскопа для ручного контроля;
- отрисовка А и В сканов;
- выставление поисковых уровней для подсвечивания дефектов на В скане;
- сохранение и загрузка данных.

Для данного ПО было разработано три модуля, а именно:

- обработки данных;
- визуализация данных;
- хранения данных.

Модуль обработки данных. Данный модуль отвечает за преобразования данных и поиск дефектов. Под преобразование данных подразумевается реализация таких алгоритмов как:

- усиление сигнала;
- фильтр;
- поиск дефектов.

Для получения информации о наличии дефекта и его амплитуду необходимо выделить область на В скане после чего нужно нажать на появившуюся кнопку сохранить измерение. После того как была нажата кнопка «сохранить измерение» запускается алгоритм поиска дефекта выбранной области. Сначала определяются границы поисковой области, после определения границ запускается цикл от начала до конца поисковой области внутри цикла проверится условие, где сравниваются значение поискового уровня с данными на выбранном отрезке. После завершения работы цикла формируется массив с значениями, амплитуда которых превышает пороговый уровень. После этого полученный массив проходит еще один цикл по нахождению максимального значения амплитуды. Данное значение будет выведено пользователю.

Визуализация данных. Данный метод отвечает за отображение А и В скан. А скан представляет из себя графическое отображение формы сигнала в реальном времени

распространяющейся звуковой волны от преобразователя до получения отражения от нижней части исследуемого образца. Данные с дефектоскопа поступают в модуль обработки данных откуда передаются в модуль визуализации. Полученные данные сортируются, после чего визуализируются в виде А скан. Актуальный А скан (рис. 1) отображается в интерфейсе пользователя.

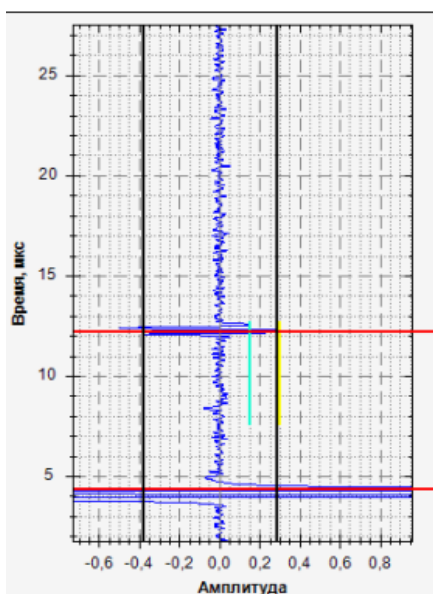


Рис. 1. Изображение А-скан

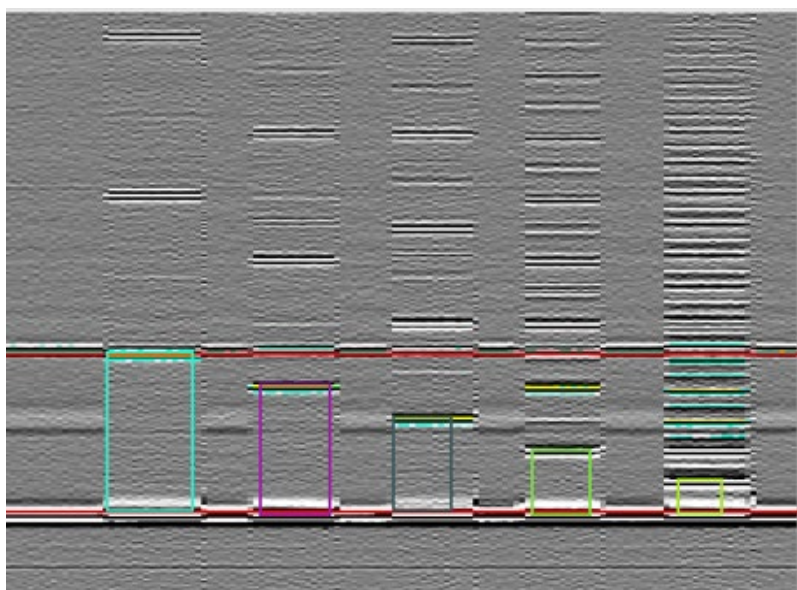


Рис. 2. Изображение В-скан

Для формирования В скан необходимо произвести первичную обработку данных. Полученные данные разделяются на действительную и мнимую часть с помощью быстрого преобразования Фурье, после чего для выбранного диапазона спектра формируется фильтр. Сформированный фильтр применяется к мнимой и действительной части. Для полученного результата применяется обратное преобразование Фурье. После обратного преобразования данные отправляются в функции для отрисовки В скан. После запуска функции данные проходят интерполяцию. Интерполированные данные в цикле проходят через ряд условий, где им присваивается определенная градация серого, полученный цвет записывается в массив цветов. На основе массива цветов строится В скан. Полученные данные в режиме реального времени отображаются на В скан рисунок 2.

Таким образом, было разработано программное обеспечение для ручного УЗК контроля. Данное ПО позволяет обеспечить быструю и достоверную визуализацию данных, а также позволяет произвести быстрый анализ на наличие дефектов в интересующей пользователя области.

Литература

1. Ультразвуковой неразрушающий контроль. [Электронный ресурс]. – Режим доступа: <http://earchive.tpu.ru/bitstream/11683/27405/1/TPU158607.pdf> (дата обращения: 22.06.2021).
2. Совершенствование алгоритмов обработки акустического сигнала в целях повышения эффективности оценки качества компонентов АЭС. [Электронный ресурс]. – Режим доступа: <https://core.ac.uk/download/pdf/53092854.pdf> (дата обращения: 10.01.2022).
3. Автоматизированный ультразвуковой контроль металлургической продукции в условиях массового производства на примере разработок ИЭС им. Е.О. Патона. [Электронный ресурс]. – Режим доступа: <http://www.panatest.ru/static?al=article-auzk> (дата обращения: 03.01.2021).
4. Ультразвуковой акустический микроскоп. [Электронный ресурс]. – Режим доступа: <https://avanteh.ru/cat/micro/akusticheskaja-mikroskopija/article/ultrazvukovoiakusticheskii-mikroskop.html> (дата обращения: 15.06.2021).

УДК 004.021

МЕТОД ТЕСТИРОВАНИЯ БЕЗОПАСНОСТИ КОМПЬЮТЕРНЫХ СИСТЕМ НА ОСНОВЕ КОНЦЕПЦИИ ЦИФРОВЫХ ДВОЙНИКОВ

Татаров Д.А.¹ (студент)

Научный руководитель – канд. техн. наук, доцент Менищikov А.А.¹

¹Университет ИТМО

e-mail: datatarov@itmo.ru, menshikov@itmo.ru

В работе был проведен поиск и анализ существующих работ по построению и применению цифровых двойников для обеспечения защищенности киберфизических систем. Были выявлены достоинства и недостатки предложенных решений. Предложено решение по построению цифрового двойника киберфизической системы, которое нивелирует недостатки существующих решений.

Ключевые слова: цифровой двойник, тестирование на проникновение, эмуляция, виртуализация, SDN, AutomationML.

Множество компаний обязаны проходить тестирование на проникновение и анализ защищенности, однако, тестирование может включать в себя нагрузочное тестирование, а также работу тяжелых сетевых сканнеров, которые могут нарушить работоспособность системы. Также при эксплуатации некоторых уязвимостей может произойти перезагрузка целевого устройства или отказ в обслуживании. Если дело касается финансовых организаций, то даже минуты простоя могут принести многомиллионные убытки. Поэтому администраторы сетей стараются не допускать аудиторов и тестировщиков к информационной системе. Кроме того, при построении новой системы, невозможно учесть все риски, не имея готовой системы, которую можно протестировать.

Для решения проблемы было предложено строить тестовые системы на основании технологии цифровых двойников.

Цифровой двойник — виртуальное отображение существующей системы, обновляющееся в реальном времени, дублирующее ключевые характеристики исходной системы и способное воспроизводить состояния системы при различных входных данных.

Существующие решения

В настоящее время многие производители микроконтроллеров и сетевого оборудования уже имеют готовые проприетарные решения по виртуализации своего оборудования. Однако, эти решения, зачастую не предполагают использования в кластерах и сетях. При использовании решений сразу нескольких производителей могут возникнуть проблемы с объединением их в одну систему.

Примеры существующих решений для виртуализации сетевого оборудования, серверов и программируемых логических контроллеров (ПЛК):

1. VMware NSX.
2. Cisco Enterprise Network Functions Virtualization.
3. SolarWinds Virtualization Manager.
4. Siemens SIMATIC.
5. ABB Ability.
6. Mininet.

Научные работы в данной области

Во время проведения обзора научных работ в данной предметной области было найдено три работы [1–3], суть которых заключается в построении цифровых двойников

киберфизических систем, для последующего обеспечения безопасности исходных систем. Данные работы сопровождаются репозиториями с исходным кодом приложений для построения систем цифровых двойников на основе реальных систем.

Также были найдены две научные работы, которые не имеют тестовых реализаций технологии цифровых двойников в открытом доступе.

Первая работа описывает методологию построения цифровых двойников киберфизических систем на примере электростанции [4].

Вторая работа описывает методологию построения процессов синхронизации реальной системы с технологией цифрового двойника [5].

Все существующие работы в данной предметной области сфокусированы на создание цифровых двойников киберфизических систем, состоящих в основном из низкоуровневых элементов, таких как микроконтроллеры, элементы систем IOT (датчики, передатчики), простейшее сетевое оборудование (коммутаторы, концентраторы).

Все решения построены на базе фреймворка для создания SDN Mininet и парсера спецификаций AutomationML.

Описанные в научных работах решения реализованы на базе приведенных выше фреймворков, являются так называемыми Proof Of Concept, т.е. они реализованы в минимальном исполнении, достаточном для демонстрации возможности реализации.

Предложенное решение

1. Средство виртуализации компонентов.

Чтобы устранить имеющиеся проблемы, было предложено использовать проприетарное решение для виртуализации и создания SDN VMWARE NSX.

VMware NSX — это семейство программных продуктов для виртуальных сетей и безопасности, созданное на основе VMware vCloud Networking and Security (vCNS) и платформы сетевой виртуализации (NVP) Nicira.

VMware vCNS виртуализирует сетевые уровни модели OSI 4-7 сети, а NVP от Nicira виртуализирует сетевую структуру на уровнях 2 и 3.

VMware NSX обладает всеми преимуществами Mininet, позволяет виртуализировать те же компоненты, что и Mininet, однако, имеет возможность также виртуализировать OS Windows. Виртуализация продуктов семейства Windows необходима, так как подавляющее число предприятий имеет в основе своей информационной инфраструктуры систему Active Directory.

Стоит заметить, что VMWARE NSX также лишен недостатков изоляции процессов и ядра.

Из минусов можно отметить тот факт, что данное ПО является проприетарным и распространяется по лицензии.

2. Язык описания спецификаций.

Для описания спецификаций не было найдено решения, поддерживающего как низкоуровневые компоненты, так и высокоуровневые, поэтому было принято решение о разработке собственного языка описания.

3. Жизненный цикл цифрового двойника.

На рисунке приведен жизненный цикл цифрового двойника, демонстрирующий этапы переноса реальной системы в виртуальную, а также непрерывного обновления этой виртуальной системы.

4. Агрегация и сбор данных.

Агрегация данных необходима как для построения изначальной системы цифрового двойника, так и для отслеживания изменений.

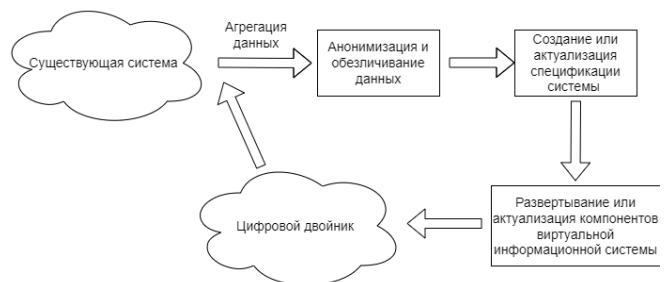


Рисунок. Жизненный цикл цифрового двойника

Агрегацию данных можно условно разделить на пассивные и активные источники. Пассивные источники – информация, получаемая без непосредственного взаимодействия с

целевой системой. Зачастую, информация из этих источников представляет собой исчерпывающие знания о системе (white box). Активные источники – информация, которая может быть получена с нулевыми начальными знаниями о целевой системе (black box).

Примеры пассивных источников:

- конфигурационные файлы;
- лог-файлы сетевых подключений;
- сетевой трафик;
- исходный код или бинарные файлы.

Примеры активных источников:

- сетевое сканирование;
- профилирование приложений прикладного уровня;
- Fuzzing приложений прикладного уровня.

В существующих работах упоминаются только некоторые из приведенных пассивных источников, при этом пренебрегаются активные.

Применение цифровых двойников способно облегчить для компаний прохождения тестирований на проникновение, на основании цифровых двойников возможно проводить отработку большего количества сценариев атак для тренировки отдела SOC, а также для тестирования компонентов обнаружения и предотвращения атак. Основное преимущество заключается в том, что цифровой двойник запускается в отдельной среде выполнения и никак не влияет на безопасность и производительность основной системы.

Литература

1. Eckhart M., Ekelhart A. Towards security-aware virtual environments for digital twins // Proceedings of the 4th ACM Workshop on Cyber-Physical System Security. – 2018. – Pp. 61–72.
2. Eckhart M., Ekelhart A. A Specification-based state replication approach for digital twins // Proceedings of the 2018 Workshop on Cyber-Physical Systems Security and Privacy. – 2018. – Pp. 36–47.
3. Dietz M., Vielberth M., Pernul G. Integrating digital twin security simulations in the security operations center // Proceedings of the 15th International Conference on Availability, Reliability and Security. – 2020. – Pp. 1–9.
4. Atalay M., Angin P. A Digital Twins approach to smart grid security testing and standardization // IEEE 2020 IEEE International Workshop on Metrology for Industry 4.0 & IoT (MetroInd4.0&IoT). – 2020. – Pp. 435–440.
5. Gehrman C.; Gunnarsson M. A Digital twin based industrial automation and control system security architecture // IEEE Transactions on Industrial Informatics. – 2019. – Pp. 669–680.

УДК 629.05

АЛГОРИТМЫ СОВМЕСТНОЙ НАВИГАЦИИ И КАРТОГРАФИРОВАНИЯ, ПОСТРОЕННЫЕ НА ОСНОВЕ ПРОЦЕДУР ЧАСТИЧНОГО АНАЛИТИЧЕСКОГО ИНТЕГРИРОВАНИЯ

Титов Р.У.¹ (аспирант), Моторин А.В.^{1,2}

Научный руководитель – чл.-кор. РАН Степанов О.А.^{1,2}

¹Университет ИТМО

²АО «Концерн «ЦНИИ «Электронприбор»

e-mail: roman100895@rambler.ru, motorin.a@mail.ru

*Работа выполнена при поддержке гранта РФФИ проект № 18-19-00627
(<https://rscf.ru/project/18-19-00627/>).*

Описывается общая постановка задачи SLAM с позиций байесовского подхода. Рассматриваются варианты её решения с использованием линеаризации и процедур частичного аналитического интегрирования и сравнивается их вычислительная сложность.

Ключевые слова: алгоритмы SLAM, байесовский подход, частичное аналитическое интегрирование.

В настоящее время всё большую актуальность приобретают алгоритмы решения задачи совместной навигации и картографирования, в англоязычной литературе именуемые, как алгоритмы SLAM (Simultaneous Localization And Mapping). Их распространение связано с необходимостью одновременного решения задач локализации подвижного объекта и построения карты неизвестной ему среды в условиях отсутствия связи с системами глобального позиционирования [1–3]. Задача SLAM может быть описана в общем виде с позиций байесовского подхода [1, 4]. Цель работы – пояснить работу двух возможных байесовских алгоритмов в задаче SLAM и сопоставить их вычислительную сложность.

Общая постановка задачи SLAM и алгоритмы ее решения. Будем считать, что решается задача SLAM на плоскости в предположении, что имеется L ориентиров, представленных вектором $M = [m^1, \dots, m^L]^T$. Эта задача сводится к уточнению координат подвижного объекта $X_i = [x_1, x_2, \dots, x_i]^T$ и координат ориентиров m^l в моменты времени i по данным о приращениях координат $U_i = [u_1, u_2, \dots, u_i]^T$, полученных, например, от автономной навигационной системы, и измерениям $Z_i = [z_1, z_2, \dots, z_i]^T$ до ориентиров, например, измерениям дальностей и/или пеленгов.

С байесовских позиций оптимальная оценка искомого вектора $[X_i, M]^T$ в общем случае представляется в следующем виде:

$$[\hat{X}_i, \hat{M}]^T = \iint [X_i, M]^T f(X_i, M/Z_i, U_i) dX_i dM, \quad (1)$$

где $f(X_i, M/Z_i, U_i)$ – апостериорная функция плотности распределения вероятности (ф.п.р.в.).

Таким образом, решение задачи SLAM сводится к вычислению многомерного интеграла (1) при известной статистической взаимосвязи оцениваемых параметров объекта и карты $[X_i, M]^T$ и связанных с ними измерений Z_i и известных значениях U_i . Существует ряд алгоритмов получения такой оценки [4]. Одним из вариантов является аналитическое вычисление интеграла (1), что возможно при гауссовской ф.п.р.в. $f(X_i, M/Z_i, U_i)$. Рассматриваемая задача является нелинейной, что приводит к негауссовской ф.п.р.в. Полагая для простоты, что уравнения, описывающие изменения координат объекта, линейные, и осуществляя линеаризацию измерений относительно координат объекта и координат

ориентиров, можно получить выражения для рекуррентного вычисления параметров гауссовской аппроксимации $f(X_i, M/Z_i, U_i)$, которые могут быть найдены с помощью обобщённого фильтра Калмана (ФК), в англоязычной литературе именуемого как Extended Kalman Filter (EKF). Размерность вектора состояния $[x_i, M]^T$ в ФК в таком случае составит $[2 + 2L]$, где L – количество ориентиров, а размерность матрицы ковариаций составит $[2 + 2L] \times [2 + 2L]$. При увеличении числа ориентиров имеет место квадратичный рост увеличения размерности матриц, используемых в ФК, и, как следствие, существенное возрастание вычислительной сложности алгоритма.

Другой обсуждаемый здесь вариант построения алгоритма опирается на линеаризацию измерений только относительно подвектора M и решение нелинейной задачи оценивания подвектора X_i с использованием метода Монте-Карло и идеи аналитического интегрирования по части переменных. Алгоритм интегрирования по части переменных применительно к задачам нелинейной фильтрации предложен достаточно давно [5–7], а применительно к задачам SLAM в [8] (в англоязычной литературе именуется как FAST SLAM). Апостериорная ф.п.р.в. $f(X_i, M/Z_i, U_i)$ может быть представлена следующим образом

$$f(X_i, M/Z_i, U_i) = f(X_i/Z_i, U_i)f(M/X_i, Z_i, U_i), \quad (2)$$

где с учетом предположения о независимости ошибок измерения плотность $f(M/X_i, Z_i, U_i)$ представляется как

$$f(M/X_i, Z_i, U_i) = \prod_{l=1}^L f(m^l/X_i, Z_i, U_i). \quad (3)$$

Для понимания сути обсуждаемого алгоритма целесообразно рассмотреть задачи картографирования и локализации отдельно. При известных координатах робота нетрудно заметить, что задача SLAM редуцируется к задаче картографирования. Координаты ориентиров оцениваются независимо друг от друга и для реализации рекуррентной процедуры оценивания могут быть использованы L обобщенных двумерных ФК. При известных же координатах ориентиров задача SLAM превращается в нелинейную задачу оценивания координат по измерениям дальностей и пеленгов до точечных ориентиров (задача навигации).

Решение задачи SLAM основано на аппроксимации апостериорной ф.п.р.в. $f(x_i/Z_i, U_i)$ координат объекта с использованием методов Монте Карло путем представления ее в виде:

$$f(x_i/Z_i, U_i) \approx \sum_{j=1}^K \mu_i^j \delta(x_i - \chi_i^j), \quad (4)$$

где χ_i^j – элементы выборки объема K , соответствующей ф.п.р.в. прогноза $f(x_i/U_i)$,
 μ_i^j – весовые коэффициенты.

Элементы выборки χ_i^j могут быть получены из априорной выборки χ_0^j , соответствующей априорному гауссовскому распределению с использованием модели динамики объекта (рис. 1). Каждую χ_i^j можно рассматривать как гипотезу о местоположении объекта в момент времени i , а последовательности $[\chi_1^j, \chi_2^j, \dots, \chi_i^j]$ – гипотезы о пути робота.

Весовые коэффициенты μ_i^j вычисляются согласно соотношениям правила Байеса:

$$\mu_i^j = \frac{\mu_{i-1}^j f(z_i/\chi_i^j, U_i, Z_{i-1})}{\sum_{j=1}^K \mu_{i-1}^j f(z_i/\chi_i^j, U_i, Z_{i-1})}, \quad (5)$$

где функция правдоподобия $f(z_i/\chi_i^j, U_i, Z_{i-1})$ вычисляется согласно

$$f(z_i/\chi_i^j, U_i, Z_{i-1}) = \frac{1}{\sqrt{2\pi|\Lambda_i|}} \exp\left[-\frac{1}{2}(z - \hat{z}_i^j)^T \Lambda_i^{-1}(z - \hat{z}_i^j)\right], \quad (6)$$

где \hat{z}_i^j - прогноз измерений,

Λ_i - матрица ковариаций измерений.

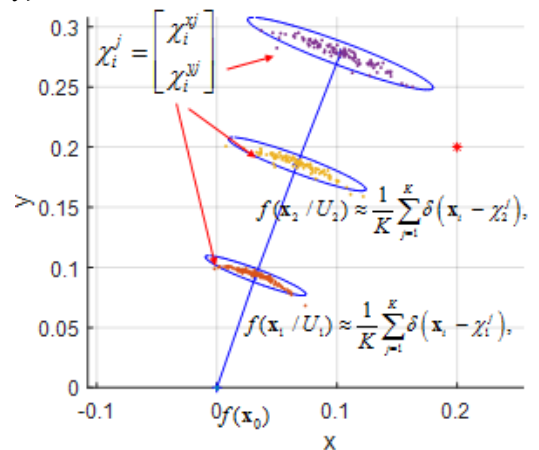


Рис. 1. Иллюстрация получения выборки χ_i^j

Интегрирование по части переменных реализуется при вычислении показателя экспоненты (6) с учетом линеаризации измерений относительно координат ориентиров. Весовые коэффициенты определяют вероятность нахождения объекта в окрестности координат χ_i^j при измерениях Z_i . Для повышения эффективности метода Монте-Карло, кроме того, периодически выполняется процедура так называемой последовательной перевыборки [6, 7], что приводит к устранению маловероятных χ_i^j .

При решении задачи SLAM с использованием описанного алгоритма его можно рассматривать как многоальтернативную задачу картографирования со множеством гипотез о пути объекта, аналогично тому, как эти гипотезы возникают в задаче навигации. Однако, в случае SLAM наличие K альтернативных гипотез о пути приведет также к K гипотезам о координатах ориентиров карты, что, принимая во внимание необходимость оценивания координат ориентиров, приведет к реализации $K \cdot L$ обобщенных ФК размерности 2 (для рассматриваемой задачи на плоскости). Можно отметить, что вычислительная сложность такого алгоритма растет линейно с увеличением числа ориентиров, в отличие от обобщенного ФК, получаемого при линеаризации относительно всего вектора $[x_i, M]^T$. Более того, из-за реализации процедур последовательной перевыборки, часть обобщенных ФК из банка будут идентичны. Используя этот факт можно сократить размерность решения задачи до $K \cdot \log(L)$ [4, 8].

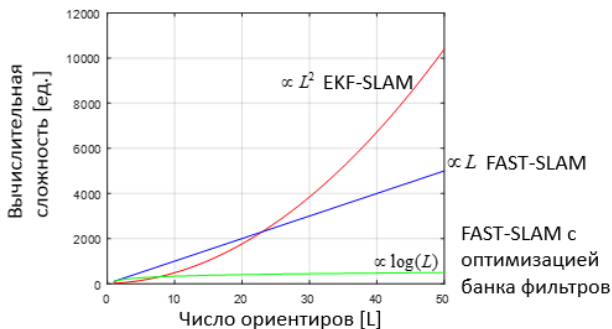


Рис. 2. Рост вычислительной сложности в зависимости от числа ориентиров L

Заметим, что помимо вычислительной сложности в дальнейшем предполагается оценить точность, достигаемую с использованием каждого из этих алгоритмов.

Приведена общая байесовская постановка задачи SLAM и описаны два возможных алгоритма ее решения. Один алгоритм представляет собой обобщенный фильтр Калмана для всех оцениваемых переменных, а другой основан на использовании метода Монте-Карло совместно с привлечением процедур аналитического интегрирования по части переменных. Проведено сопоставление алгоритмов с точки зрения их вычислительной сложности. Отмечена возможность заметного сокращения вычислительной сложности при использовании обобщенного ФК при большом числе ориентиров.

Литература

1. Cadena C. et al. Past, Present, and Future of Simultaneous Localization and Mapping: Toward the Robust-Perception Age // IEEE Transactions on Robotics. – 2016. – Vol.32. – №6. – Pp. 1309–1332.
2. Fairfield N., Kantor G., Wettergreen D. Towards particle filter SLAM with three dimensional evidence grids in a flooded subterranean environment // Proceedings 2006 IEEE International Conference on Robotics and Automation, ICRA 2006. – Pp. 3575–3580.
3. Barkby S., Williams S., Pizarro O., Jakuba M. An efficient approach to bathymetric SLAM // 2009 IEEE/RSJ International Conference on Intelligent Robots and Systems. – 2009. – Pp. 219–224.
4. Montemerlo M., Thrun S., Koller D., Wegbreit B. FastSLAM: A Factored Solution to the Simultaneous Localization and Mapping Problem // Eighteenth National Conference on Artificial Intelligence. – 2002. – Pp. 593–598.

5. Степанов О.А. Применение теории нелинейной фильтрации при решении задач обработки навигационной информации // 3rd ed. Санкт-Петербург: ГНЦ РФ ЦНИИ “Электроприбор”. – 2003. – С. 370.
6. Степанов О.А., Торопов А.Б. Применение последовательных методов Монте-Карло с использованием процедур аналитического интегрирования при обработке навигационной информации // Материалы XII Всероссийского совещания по проблемам управления ВСПУ-2014. – 2014. – Pp. 3324–3337.
7. Doucet A.N. De Freitas Gordon N. Sequential Monte Carlo Methods // Practice. New York, NY: Springer-Verlag. – 2001. – 582 с.
8. Murphy K.P. Bayesian map learning in dynamic environments // Proceedings of the 12th International Conference on Neural Information Processing Systems. – 1999. – Pp. 1015–1021.

УДК 004.056.055

АНАЛИЗ СТОЙКОСТИ КРИПТОГРАФИЧЕСКИХ СХЕМ, ОСНОВАННЫХ НА ИЗОГЕНИЯХ НЕСУПЕРСИНГУЛЯРНЫХ ЭЛЛИПТИЧЕСКИХ КРИВЫХ

Хуцаева А.Ф.¹ (студент)

Научный руководитель – Давыдов В.В.¹

¹Университет ИТМО

e-mail: gladdude1@gmail.com

Работа выполнена в рамках темы НИР №620164 «Методы искусственного интеллекта для киберфизических систем».

Работа частично финансировалась Федеральной программой академического лидерства Приоритет 2030.

В работе проведен анализ криптографической стойкости схем, основанных на изогениях несуперсингулярных эллиптических кривых. Исследованы алгоритмы построения таких кривых, один из которых является наиболее эффективным по времени генерации. Рассмотрена атака скрытого сдвига, позволяющая провести взлом криптосистем, использующих несуперсингулярные кривые, за субэкспоненциальное время.

Ключевые слова: Несуперсингулярные эллиптические кривые, изогении, криптографическая стойкость, группа классов идеалов, полиномы Вебера.

В настоящее время эллиптические кривые широко используются в криптосистемах с открытым ключом, стойкость которых основывается на сложности решения задачи дискретного логарифмирования. Однако, данная математическая задача может быть решена за полиномиальное время с появлением полноценного квантового компьютера. Таким образом, представляется актуальным изучение математических задач, устойчивых к квантовым атакам. В связи с этим, особое внимание на сегодняшний день уделяется разделу постквантовой криптографии – уже завершены три тура конкурса NIST (Национальный институт стандартов и технологий) на стандартизацию постквантовых алгоритмов. Одним из разделов постквантовой криптографии является криптография на основе изогений эллиптических кривых – относительно «молодая» область, в которой безопасность зависит от проблемы нахождения верного пути в графе изогений. Важным фактором при построении криптографических схем, основанных на изогениях, является выбор эллиптических кривых. В данной работе проведён анализ стойкости таких схем при использовании несуперсингулярных кривых.

В работе рассматривались кривые над конечным полем с характеристикой, отличной от 2 и 3, в форме Вейерштрасса (1):

$$y^2 = x^3 + Ax + B. \quad (1)$$

Точки эллиптической кривой образуют аддитивную группу, где нейтральным элементом группы является точка на бесконечности. Также одним из важных параметров кривой является j -инвариант, имеющий постоянное значение независимо от линейных преобразований.

Изогенией между двумя эллиптическими кривыми E_1 и E_2 , определенной над полем K , называется рациональное отображение (2):

$$\varphi: E_1 \rightarrow E_2, \quad (2)$$

переводящее бесконечно удаленную точку одной кривой на другую [1] (3):

$$\varphi(P_\infty) = P_\infty. \quad (3)$$

В группе точек эллиптической кривой задается эндоморфизм $\phi: P \rightarrow nP, n \in Z$. В свою очередь эндоморфизм группы образует кольцо эндоморфизмов $End(E)$ эллиптической кривой.

Согласно теореме Тейта, эллиптические кривые E_1 и E_2 , определенные над конечным полем, являются изогенными тогда и только тогда, когда они имеют одинаковое количество точек, то есть $\#E_1 = \#E_2$.

Для подсчета количества изогений степени l используется дискриминант Фробениуса (4):

$$D_\pi = t^2 - 4q, \quad (4)$$

отношение $\left(\frac{D_\pi}{l}\right)$ – символ Кронекера, в зависимости от значения символа определяется количество изогений:

1. $\left(\frac{D_\pi}{l}\right) = -1$, то изогений степени l не существует.
2. $\left(\frac{D_\pi}{l}\right) = 1$, то существует две изогении степени l .
3. $\left(\frac{D_\pi}{l}\right) = 0$, то существует одна или $l+1$ изогений.

Данное отношение позволяет строить изогенные графы определенной формы. Так, изогенный граф – это (мульти)-граф, узлы которого являются j -инвариантами изогенных кривых, а ребра – изогениями между ними. Граф, состоящий из простого числа эллиптических кривых, соединенных изогениями степеней, удовлетворяющих $\left(\frac{D_\pi}{l}\right) = 1$, называется изогенной звездой. На основе этой звезды была предложена криптографическая система с открытым ключом [2].

Задача поиска изогений сводится к нахождению пути на графе изогенных эллиптических кривых, сложность которой зависит от числа классов идеалов дискриминанта мнимого квадратичного порядка. Число классов идеалов должно быть простым, что затрудняет поиск секретного пути для злоумышленника. Если оно является составным, то граф изогенных кривых может быть разбит на подциклы небольшой длины, что значительно снижает сложность вычислительной задачи и, как следствие, безопасность криптосистемы.

Стойкость алгоритмов шифрования, основанных на изогениях несуперсингулярных эллиптических кривых, основана на задаче поиска пути изогений в графе между эллиптическими кривыми. Для взлома криптосистем необходимо найти определенную изогению (точный путь). Поиск ведется до нахождения единственно точного пути, считается что данная задача является экспоненциальной.

Однако, в 2011 Чайлдс, Джао и Сухарев [3] предоставили работу, в которой приведен квантовый алгоритм нахождения изогений несуперсингулярных эллиптических кривых за субэкспоненциальное время.

В работе построение изогений легко сводится к проблеме скрытого сдвига с использованием группового действия.

Пусть E_1, E_2 – изогенные несуперсингулярные эллиптические кривые с эндоморфизмом кольца O_d , $Cl(O_d)$ – группа классов идеалов, $Ell_{q,n}(O_d)$ – множество из n эллиптических кривых над полем $GF(q)$, функции f_1, f_2 определены как (5):

$$f_1, f_2: Cl(O_d) \rightarrow Ell_{q,n}(O_d), \quad (5)$$

Функцию f_c можно получить с помощью идеала $[b]$ и j -инварианта E_c согласно с формулой (6):

$$f_c([b]) = [b] * j(E_c). \quad (6)$$

В них «скрыт» $[s] \in Cl(O_d)$, где $[s]$ – класс идеалов, и выполняется следующее (7):

$$[s] * j(E_1) = j(E_2). \quad (7)$$

Иными словами, с помощью классов идеалов осуществляется переход между разными классами эллиптических кривых.

В частности, при равенстве (8) получается, что $[s] \in f_1, f_2$:

$$f_1([b]) = f_0([b][s]). \quad (8)$$

Для решения данной задачи были использованы квантовые алгоритмы Куперберга и Регева.

Иными словами, предложенная атака напрямую зависит от кольца эндоморфизмов, у несуперсингулярных кривых оно коммутативно. Именно по этой причине на сегодняшний день в рамках постквантовой криптографии особый интерес представляют суперсингулярные эллиптические кривые, у которых кольцо эндоморфизмов некоммутативно. Однако, всё ещё есть попытки адаптировать несуперсингулярные кривые к требованиям постквантовой криптографии [4], но исследования всё ещё не подтверждают стойкости данных схем.

Тем не менее для классической криптографии криптосистемы на изогениях несуперсингулярных кривых остаются актуальными.

Так, для атаки на данные криптосистемы (с использованием точек) необходимо решить две задачи: поиска эквивалентного пути и задачи дискретного логарифмирования в группе рациональных точек эллиптической кривой. Временная сложность оценивается величиной $\max[\min[O(\#U), O(m^{\frac{n}{2}+1} * n)], O(\sqrt{\#E_0(F_p)})]$ операций, где U – изогенная звезда, а под $E_0(F_p)$ понимается изоморфизм, m – количество шагов изогенией одной степени выбранного пути, n – количество различных степеней изогений.

В то время как сложность взлома криптосистем на основе j -инварианта оценивается в $\min[O(\#U), O(m^{\frac{n}{2}+1} * n)]$ операций. Очевидно, что сложность экспоненциальная (в рамках классической криптографии).

Одним из важных условий при построении данных криптосистем является следующее: для существования изогенной звезды большой простой длины достаточно, чтобы число классов идеалов было большим простым числом.

Для удовлетворения данного условия при построении кривых необходимо выбирать кривые с как можно большими значениями дискриминантов квадратичных полей. В работе были проанализированы алгоритмы построения эллиптических кривых, а именно полиномы Вебера и Гильберта. Анализ показал эффективность использования полиномов Вебера, при увеличении основного параметра – дискриминанта квадратичного поля увеличивается и время генерации полиномов. В таблице представлено время, затрачиваемое на построение полиномов, где D – дискриминант мнимого квадратичного порядка, $h(D)$ – число классов идеалов.

Таблица. Время, затрачиваемое на построение полиномов

D	h(D)	Полином Вебера	Полином Гильберта
		Время, секунды	Время, секунды
228	4	0.08	0.73
3.092	26	0.32	110.26
6.744	44	0.95	699.08
9.924	52	1.97	1,365.29

Можно заметить, что при увеличении значений дискриминанта полиномы Вебера генерируются значительно быстрее.

В данной работе была проанализирована стойкость криптосистем на основе изогений несуперсингулярных эллиптических кривых. Анализ показал уязвимость таких систем к атаке скрытого сдвига, что говорит о невозможности использования несуперсингулярных кривых в задачах поиска изогений между эллиптическими кривыми. Слабость систем, основанных на таких кривых, заключается в коммутативности кольца эндоморфизмов, благодаря которому легко найти засекреченный путь на изогенном графе. Однако, в рамках классической криптографии рассматриваемые криптосистемы являются стойкими благодаря большому простому числу классов идеалов. Дополнительно были проанализированы алгоритмы построения несуперсингулярных кривых, где самым эффективным из рассмотренных является построение полиномов Вебера.

Литература

1. Ростовцев А.Г. Эллиптические кривые в криптографии. Теория и вычислительные алгоритмы. – М.: НПО «Профессионал». – 2005. – 485 с.
2. Rostovtsev A., Stolbunov A. Public-key cryptosystem based on isogenies. [Электронный ресурс]. – Режим доступа: <http://eprint.iacr.org/2006/145> (дата обращения: 09.01.2022).
3. Childs A., Jao D., Soukharev V. Constructing elliptic curve isogenies in quantum subexponential time. [Электронный ресурс]. – Режим доступа: <https://arxiv.org/abs/1012.4019> (дата обращения: 12.01.2022).
4. De Feo L., Kieffer J., Smith B. Towards practical key exchange from ordinary isogeny graphs // *Advances in Cryptology*. – 2018. – С. 365–394.

УДК 534.143; 534.134; 535.312.2; 537.63

РАЗРАБОТКА МЕТОДА ПОВЫШЕНИЯ ЭФФЕКТИВНОСТИ МАГНИТООПТИЧЕСКОГО ДАТЧИКА ГРАДИЕНТА МАГНИТНОГО ПОЛЯ С ТОРСИОННЫМ ПОДВЕСОМ ЧУВСТВИТЕЛЬНОГО ЭЛЕМЕНТА ЗА СЧЕТ УМЕНЬШЕНИЯ ВРЕМЕНИ УСПОКОЕНИЯ

Чернов Р.И.¹ (студент), Тихомиров А.В.¹ (студент), Еремук В.В.¹

Научный руководитель – д-р техн. наук, доцент Гришенцев А.Ю.¹

¹*Университет ИТМО*

e-mail: aeijo@mail.ru

Работа выполнена в рамках темы НИР №620164 «Методы искусственного интеллекта для киберфизических систем».

В работе рассмотрено применение жидкостного, воздушного и магнитоиндукционного успокоителя в конструкции магнитооптического датчика градиента магнитного поля с торсионным подвесом чувствительного элемента. Результаты математического моделирования показали, что наиболее целесообразным методом повышения эффективности датчика является применение магнитоиндукционного успокоителя.

Ключевые слова: датчик градиента магнитного поля, математическое моделирование, электромагнитное поле, переходные процессы, магнитоиндукционный успокоитель.

Перспективным методом измерения магнитного поля является измерение с помощью датчиков градиента магнитного поля с подвесом чувствительного элемента [1–2]. Одной из проблем, препятствующих процессу измерений магнитного поля с помощью датчиков с подвесом, является инерционность чувствительного элемента датчика. Инерционность чувствительного элемента датчика обуславливает неприемлемую продолжительность затухающих колебаний, возникших под воздействием внешнего магнитного поля. В работе рассматривается использование в конструкции датчика различных типов успокоителей, обеспечивающих уменьшение времени успокоения чувствительного элемента датчика. Целью исследования является анализ и моделирование датчика градиента магнитного поля с торсионным подвесом чувствительного элемента и поиск конструктивных решений. Объектом исследования является датчик градиента магнитного поля с торсионным подвесом чувствительного элемента. Предметом исследования являются физические процессы, происходящие в датчике при воздействии внешнего магнитного поля.

Принцип действия магнитооптического датчика градиента магнитного поля с торсионным подвесом чувствительного элемента заключается в измерении угла поворота индикаторного магнита при воздействии внешнего магнитного поля. Структурная блок-схема датчика представлена на рисунке 1.

Одним из компонентов датчика является лазер, луч которого направлен на индикаторный магнит. Индикаторный магнит имеет зеркальную поверхность, отражаясь от которой лазерный луч попадает в систему зеркал. Система зеркал обеспечивает увеличение хода луча при сохранении компактности датчика. После многократного отражения в системе зеркал лазерный луч попадает на фоточувствительную линейку, представленную линейной ПЗС-матрицей.

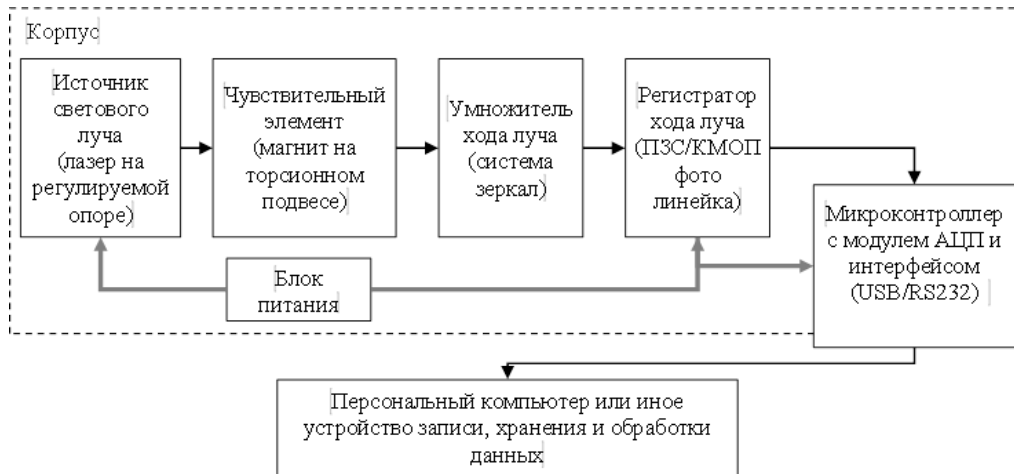


Рис. 1. Блок-схема магнитооптического датчика градиента магнитного поля с торсионным подвесом чувствительного элемента

В ходе исследований была разработана математическая модель чувствительного элемента датчика. Поведение чувствительного элемента датчика описано совокупностью моментов сил, действующих на систему торсионного подвеса чувствительного элемента (1).

$$M_{и} + M_{с} + M_{п} = M_{м}, \quad (1)$$

где $M_{и}$ – момент силы инерции, действующий на магнит,

$M_{с}$ – момент силы сопротивления, включающий в себя сопротивление воздуха и силы трения,

$M_{п}$ – момент противодействующей силы нити,

$M_{м}$ – момент силы воздействия внешнего магнитного поля на постоянный магнит.

Получено дифференциальное уравнение второго порядка, описывающее колебания чувствительного элемента датчика под действием магнитного поля (2).

$$J \frac{d^2\varphi}{dt^2} + P \frac{d\varphi}{dt} + kG\varphi = [p_m \times B], \quad (2)$$

где J – момент инерции магнита,

φ – угол поворота магнита,

t – время,

$\frac{d^2\varphi}{dt^2}$ – вектор углового ускорения,

P – коэффициент успокоения системы, безразмерная величина,

$\frac{d\varphi}{dt}$ – вектор угловой скорости,

G – коэффициент кручения нити (модуль кручения нити),

k – единичный вектор момента силы противодействия нити, направление которого выбирается в зависимости от направления закручивания нити,

p_m – магнитный момент,

B – индукция внешнего магнитного поля.

Выражение для расчёта момента инерции цилиндра относительно поперечной оси симметрии (3) известно из литературы [3].

$$J = m \left(\frac{R^2}{4} + \frac{h^2}{12} \right), \quad (3)$$

где h – высота магнита.

При расчете и проектировании электромеханических приборов коэффициент успокоения (4) обычно рассчитывают с помощью эмпирических формул [4].

$$P \approx \left(\frac{0.25}{\delta} + 0.153 \right) ab\beta \cdot 10^{-10} \text{ Н} \cdot \text{м} \cdot \text{с/рад}, \quad (4)$$

где δ – зазор между крылом и стенками камеры (мм),

a и b – размеры крыла (мм),

β – расстояние от оси вращения до центра крыла (мм).

Коэффициент кручения нити зависит от параметров нити, представляет собой момент противодействующей силы нити на единичный угол её поворота и равен (5).

$$G = \frac{\pi N r^4}{2l}, \quad (5)$$

где N – противодействующий модуль кручения нити,
 r – радиус сечения нити,
 l – длина нити.

Момент силы воздействия внешнего магнитного поля на постоянный магнит определяется в соответствии с взаимной ориентацией внешнего возмущающего магнитного поля и магнитного поля магнита. Внешнее магнитное поле представим в виде гармонических Фурье сумм трехкомпонентного вектора (6). А магнит заменим эквивалентным ему единичным контуром с током.

$$\mathbf{B} = \{B_x, B_y, B_z\}. \quad (6)$$

В программной среде Wolfram Mathematica была разработана программа, реализующая математическую модель системы торсионного подвеса чувствительного элемента датчика и решения дифференциального уравнения (2).

В ходе моделирования получены графики затухающих колебаний индикаторного магнита под воздействием внешнего магнитного поля при различных коэффициентах успокоения. Графики представлены на рисунках 2 и 3.

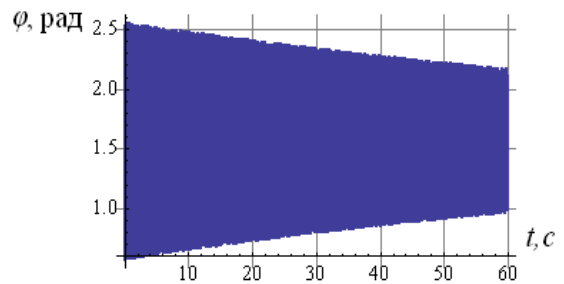


Рис. 2. Затухание крутильных колебаний магнита (в воздухе) как отклик на изменение внешнего магнитного поля по функции Хевисайда

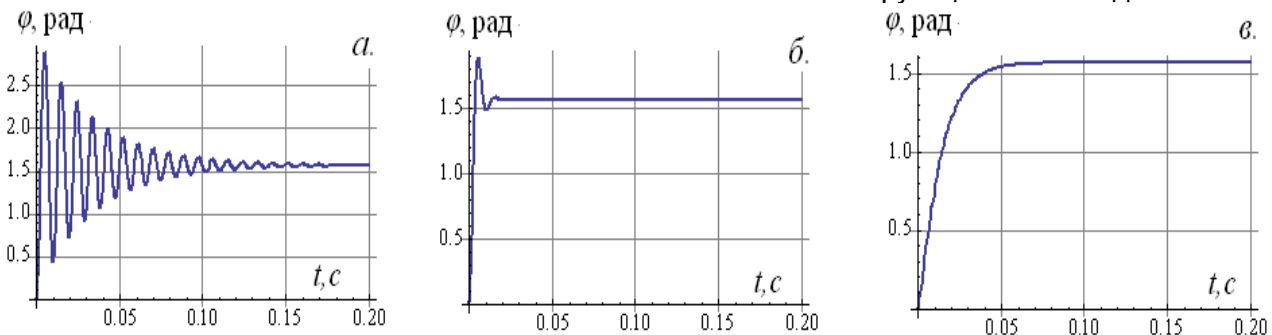


Рис. 3. Затухание колебаний магнита на торсионном подвесе при различных коэффициентах успокоения: а) коэффициент успокоения 10^{-6} Н · м · с/рад; б) коэффициент успокоения 10^{-5} Н · м · с/рад; в) коэффициент успокоения 10^{-4} Н · м · с/рад

Результаты моделирования показали, что приемлемое время затухания колебаний достигается при значениях коэффициента успокоения из диапазона, соответствующего рисункам 3.б и 3.в. Для достижения такого коэффициента успокоения при использовании воздушного успокоителя, представленного на рисунке 4, необходимо существенно увеличить площадь его крыльев.

Использование жидкостного успокоителя представляется нецелесообразным ввиду зависимости вязкости жидкости от температуры, сложности расчета траектории лазерного луча с учетом преломления и из-за усложнения конструкции.

Таким образом, предпочтительно использование магнитоиндукционного успокоителя (рис. 5), дополнительное преимущество которого заключается в том, что при круговой прорези в медной трубке для лазерного луча возможно регулирование коэффициента успокоения.

Разработана математическая модель чувствительного элемента датчика градиента магнитного поля. Произведено моделирование магнитооптического датчика градиента магнитного поля с торсионным подвесом чувствительного элемента.

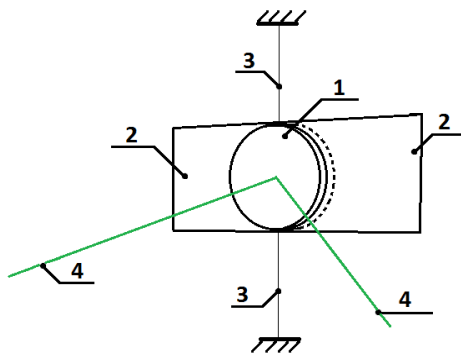


Рис. 4. Воздушный успокоитель:
 1 – магнит на торсионном подвесе;
 2 – крыло успокоителя;
 3 – нить торсионного подвеса;
 4 – лазерный луч

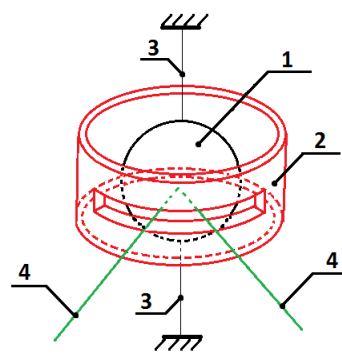


Рис. 5. Магнитоиндукционный успокоитель:
 1 – магнит на торсионном подвесе;
 2 – медная трубка успокоителя с прорезью;
 3 – нить торсионного подвеса;
 4 – лазерный луч

Произведен анализ применения жидкостного, воздушного и магнитоиндукционного успокоителя. Магнитоиндукционный успокоитель позволяет решить проблему инерционных колебаний индикаторного магнита, сохранив компактность физических размеров датчика и обеспечив возможность настройки коэффициента успокоения датчика. Использование других видов успокоителей сопряжено с усложнением конструкции датчика и увеличением его физических размеров. Таким образом, результаты моделирования и анализа показали, что для снижения инерционности чувствительного элемента датчика наиболее целесообразно использовать магнитоиндукционный успокоитель. Дальнейший план исследований включает в себя разработку опытного образца магнитооптического датчика градиента магнитного поля с торсионным подвесом чувствительного элемента, проведение натуральных экспериментов по измерению и детектированию магнитного поля с помощью датчика и дальнейшее совершенствование математической модели датчика.

Литература

1. Сергушин П.А., Мусалимов В.М. Моделирование динамики магнитоэлектростатического вариометра. Научно-технический вестник Санкт-Петербургского государственного университета информационных технологий, механики и оптики. – 2007. – №37. – С. 333–338.
2. Коробейников А.Г., Перечесова А.Д., Калапышина И.И. Исследование параметров полимерного подвеса чувствительного элемента в магнитометрическом приборе для измерения низкочастотных магнитных полей. Сборник статей XXV Всероссийской конференции и 16-ой Школы молодых ученых. Структура и динамика молекулярных систем. Национальный парк Марий Чодра, озеро Яльчик. – 2018. – С. 75–82.
3. Никитин Н.Н. Курс теоретической механики: Учебник для машиностроит. и приборостроит. спец. вузов. – 5-е изд., перераб. и доп. М.: Высшая школа. – 1990. – 607 с.
4. Преображенский А.А. Аналоговые электроизмерительные приборы. Учебное пособие для вузов / Дмитриев Ф.С., Киселева Е.А., Лебедева Г.П. и др. – М.: Высшая школа. – 1979. – 352 с.

УДК 681.516.33

ИССЛЕДОВАНИЕ АЛГОРИТМОВ УПРАВЛЕНИЯ ДВИЖЕНИЕМ ЧЕТВЕРОНОГИМИ РОБОТАМИ

Шамраев А.Д.¹ (студент)

Научный руководитель – д-р техн. наук, профессор Колюбин С.А.¹

¹Университет ИТМО

e-mail: adshamraev@itmo.ru, s.kolyubin@itmo.ru

Работа выполнена в рамках темы НИР №620164 «Методы искусственного интеллекта для киберфизических систем».

В данной работе был сделан обзор алгоритмов управления движением четвероногими роботами. Алгоритмы делятся на два направления: оптимальные регуляторы и регуляторы с обучением с подкреплением (RL).

Ключевые слова: четвероногие роботы, оптимизация, MPC, WBIC, RPC, Reinforcement Learning, SLIP, оптимальное управление, управление на основе обучения с подкреплением.

Шагающие роботы представляют особый интерес среди мобильных роботов. В отличие от колесных роботов шагающие роботы обладают более высокой проходимостью, но и имеют гораздо более сложное управление.

Один из типов шагающих роботов, четвероногие роботы, более распространен и упор в обзоре будет сделан на этот тип роботов.

Из-за того, что область управления четвероногими роботами весьма обширна, анализ сконцентрирован в направлении оптимальных регуляторов. Выдающейся в этой области является команда MIT под руководством Сангбэ Кима. Анализ приведен с упором на работы этой команды.

Область в управлении четвероногими роботами активно развивается. Начиная с Райберта [1] появляется стремление подражать в передвижении четвероногим животным, а в частности обеспечивать динамическую ходьбу.

Райберт сделал возможным динамическую ходьбу благодаря особой конструкции прыгающего робота, движение которого было схоже с моделью пружинного перевернутого маятника (SLIP), и очень простого управления. Однако, робот имел малую маневренность, робастность, скорость передвижения и ограниченность воспроизведения типов походок.

Для решения этих проблем были разработаны более сложные по конструкции роботы с электрическими приводами: Cheetah 3, Mini Cheetah, ANYmal. Но вместе с возможностями увеличилась и сложность в управлении четвероногими роботами. Основная цель в управлении четвероногими роботами (рис. 1) состоит в расчете положений лап $p = (x, y, z)$ и сил реакции опоры f в момент контакта для достижения заданного характера передвижения (ходьба, рысь, галоп) с заданной скоростью.

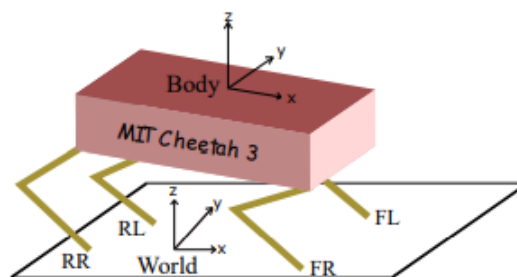


Рис. 1. Модель четвероногого робота [2]

Можно выделить два основных направления в управлении шагающими роботами:

1. Управление на основе оптимальных регуляторов.
2. Управление на основе обучения с подкреплением (Reinforcement Learning – RL).

Первое направление связано с особыми сложностями в разработке, реализации и настройке регуляторов. Так как требует значительный бэкграунд в теории управления, динамике, оптимального управления, оптимизации. Такие контроллеры трудоёмки в расчётах в реальном времени, но имеют прозрачную структуру и отрабатывают заложенный характер поведения робота.

Второе же направление в большей степени позволяет упростить работу с регуляторами и уменьшить сложность вычислений, однако, структура их не прозрачна.

Управление на основе оптимальных регуляторов в свою очередь имеет этапы развития:

1. Управление без прогнозирования. Достаточно для достижения статической ходьбы, в основе контроллера лежит решение задачи оптимизации на один шаг.

2. Управление с прогнозированием. Необходимо для достижения динамической ходьбы, в основе контроллера лежит решение задачи оптимизации на несколько шагов вперед – управление на основе прогнозирующих моделей (Model Predictive Control – MPC).

Для статической ходьбы (когда центр масс внутри области с вершинами контакта ног) достаточно решать оптимизационную задачу один раз, не прогнозируя поведение робота в следующий момент времени. В основу контроллера входит решение оптимизационной задачи (Quadratic Programming – QP). К таким контроллерам можно отнести WBC [3].

Для динамической же ходьбы становится необходимым прогнозировать поведение робота на несколько шагов вперед. В основу контроллера входит MPC. К таким контроллерам можно отнести cMPC [4], RPC [5], WHPC [6].

Необходимо решить какую модель использовать в MPC: простейшую модель пружинного перевернутого маятника (SLIP), упрощенную (линейную) модель тела робота (Rigid Body), модель тела робота (Rigid Body) или полную модель тела и ног робота (Whole-body) (рис. 2).

Имеются работы, сочетающие эти модели в разных вариациях оптимизационных задач:

1. Простая линейная модель тела робота в MPC – cMPC [4].
2. Простая линейная модель тела робота в MPC с регуляризацией – RPC [5].
3. Расчет Whole-body Control на один шаг после cMPC – cMPC+WBIC [7].
4. Для расчета половины цикла локомоции (7 шагов) используется полная модель тела и ног робота (Whole-body) в MPC, для другой половины используется упрощенная модель (SLIP или линейная Rigid Body) в MPC – WHPC [15].

Управление на основе обучения с подкреплением (Reinforcement learning) используют для решения следующих задач:

1. Замена MPC в алгоритмах на нейронную сеть, что позволит запускать алгоритмы на роботах с бортовым компьютером с меньшими вычислительными возможностями [9].
2. Нахождение положения лап и таймингов для преодоления препятствий [10].
3. Обучение робота ходить по простой модели. Обучение робота Cassie ходить по SLIP модели [11].
4. Имитация животных [12].

Из работы команды MIT можно выделить два направления развития контроллеров: cMPC [4] и RPC [5]. В результате появляются cMPC+WBIC [7], RPC+WBIC, DIC [10], Vision MPC [13].

Направление cMPC

Центральный алгоритм многих работ MIT [4], формулируется задача оптимизации с упрощенной линейной моделью робота, но без нахождения положения лап в MPC в отличие от RPC. Положения лап задаются заранее упрощенным алгоритмом используя эвристику Райберта.

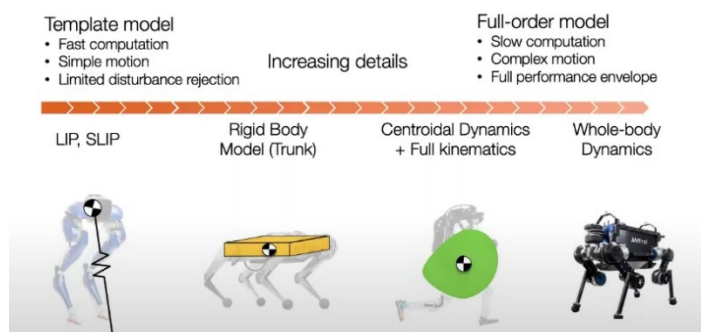


Рис. 2. Связь скорости вычислений и полноты математической модели [8]

К сMPC добавляется WBIC [7] что позволило увеличить скорость робота с 2,45 м/с до 3,7 м/с. Суть алгоритма состоит в релаксации найденных MPC сил реакции опоры.

Также добавляется компьютерное зрение [13] с применением сMPC и WBIC. И как развитие появляется DIC.

В работе [14] приводится алгоритм генерации типа походок адаптивно, а не жестко задавая тайминги как раньше.

Направление RPC

В работе [5] приводится алгоритм RPC опробованный в симуляторе. Суть алгоритма заключается в расчете не только сил ног, но и их положения в MPC а также добавления эвристик для уменьшения количества локальных минимумов.

После этот алгоритм был реализован на Cheetah 3 [15], а также разработан фреймворк для нахождения эвристик [16]. Более подробно можно посмотреть в Phd тезисе [17].

В работе [18] приводится алгоритм, основанный на RPC и WBIC с добавлением компьютерного зрения для детектирования и обхода препятствий.

Анализ работы команды MIT по четвероногим роботам показан в Mind Map по ссылке [19] (рис. 3).

Из дальнейшей работы упомянутой в статьях MIT формулирует следующие задачи дальнейшей работы:

1. Нахождение новой походки для достижения более высокой скорости и устойчивости [4].
2. Увеличение маневренности (пробовать более агрессивные характеры движения) [4].
3. Планирование траектории робота высокого уровня для достижения автономной работы (например, обойти препятствие) [20].
4. Оптимизация времени расчетов MPC [15].
5. Внедрение алгоритм [14] перехода между походками в RPC [14].
6. Замена RPC\MPC на нейронную сеть [18].
7. Отключение эвристик в некоторых ситуациях с помощью нейронной сети для RPC [16].
8. Увеличение скорости передвижения в среде с препятствиями для контроллера Vision MPC [13].
9. Определение твердости детектируемых объектов в Vision MPC [13].
10. Улучшение движения робота по лестницам [10].

Суммируя все вышеизложенное, можно составить таблицы сравнения контроллеров (табл. 1 и 2).

Таблица 1. Сравнение контроллеров управления движением четвероногими роботами

Контроллер	RL для нахождения положения лап p или сил f	Модель в MPC	Регуляризация	Внешний контроллер
сMPC [4]	–	Простая линейная	–	–
RPC [5]	–	Простая линейная	+	–
сMPC+WBIC [7]	–	Простая линейная	–	WBIC
RPC+WBIC	–	Простая линейная	+	WBIC
WHPC [6]	–	0.5 SLIP+0.5 полная	–	–
WBC [3]	–	–	–	–
DIC [10]	p	Простая линейная	–	WBIC
RL+SLIP [11]	f	–	–	–
MPC-Net [9]	f	–	–	–

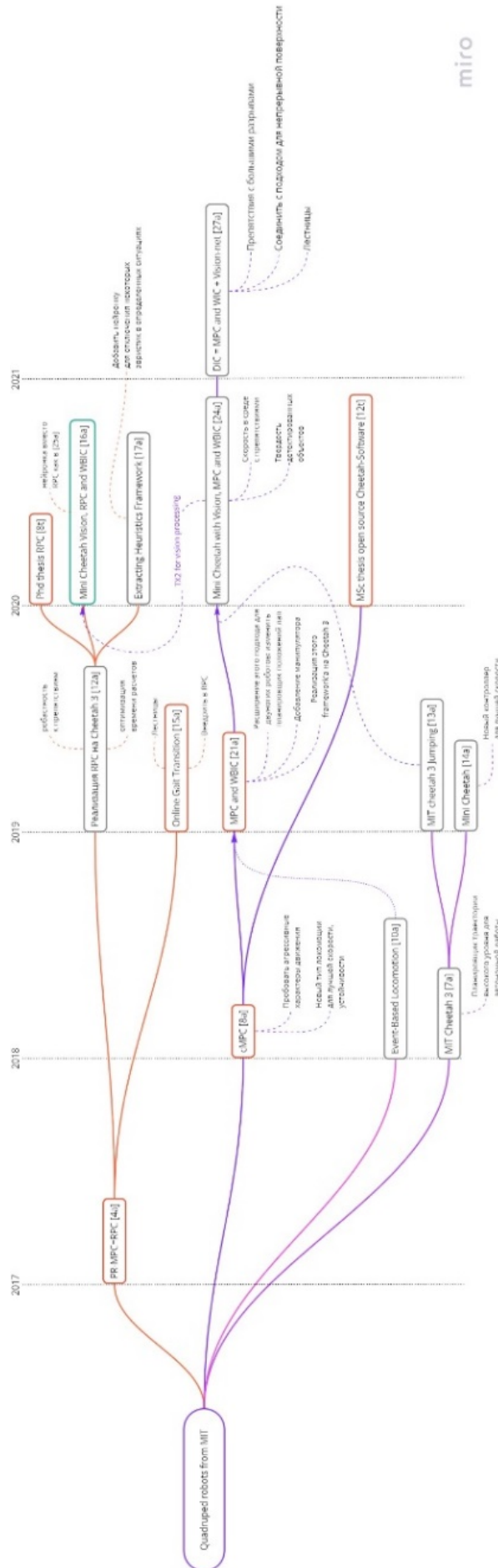


Рис. 3. Направления работы команды MIT по четвероногим роботам с 2018 года

Таблица 2. Преимущества и недостатки контроллеров управления движением четвероногими роботами

Контроллер	Преимущества	Недостатки
cMPC [4]	Относительно высокая скорость вычислений, маневренность	Средняя робастность, устойчивость, слабая естественность походки
RPC [5]	Высочайшая маневренность	Низкая скорость вычислений
cMPC+WBIC [7]	Относительно высокая скорость вычислений, маневренность, высокая робастность, устойчивость	Слабая естественность походки
RPC+WBIC	Высокая маневренность, робастность, устойчивость, естественность походки	Низкая скорость вычислений
WHPC [6]	Высокая робастность, устойчивость, маневренность	Средняя скорость вычислений
WBC [3]	Высокая скорость вычислений	Нет динамической ходьбы, низкая робастность, слабая естественность походки
DIC [10]	Относительно высокая скорость вычислений, маневренность, робастность, устойчивость	Слабая естественность походки
RL+SLIP [11]	Высочайшая скорость вычислений, высокая естественность походки	Низкая робастность
MPC-Net [9]	Высочайшая скорость вычислений	Слабая естественность походки

Алгоритмы управления движением четвероногими роботами условно можно разделить на два направления: оптимальные регуляторы и Reinforcement Learning.

К преимуществам оптимальных регуляторов можно отнести робастность, устойчивость, но они далеки от естественности походки и имеют низкую скорость вычислений.

К преимуществам регуляторов с Reinforcement Learning можно отнести относительную простоту настройки, высокую скорость вычислений, естественность походки, но они имеют слабую робастность.

Среди оптимальных регуляторов для управления движением четвероногих роботов: RPC [5] позволяет достичь большей маневренности чем cMPC [4], но более требовательный к ресурсам; WBIC [7] как дополнение к cMPC/RPC дает большую устойчивость и робастность.

Сочетание разной степени сложности моделей в MPC для оптимальных регуляторов имеет хорошую перспективу для дальнейшей работы.

Литература

1. Marc R. Legged Robots That Balance. The MIT Press. – 1986. – 205 p.
2. Di Carlo J. Software and control design for the MIT Cheetah quadruped robots. Massachusetts Institute of Technology, Department of Electrical Engineering and Computer Science. – 2020. – 101 p.
3. Kim D., Lee J., Ahn J., Campbell O., Hwang H., Sentis L. Computationally-Robust and Efficient Prioritized Whole-Body Controller with Contact Constraints // IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS. – 2018. – Pp. 5987–5994.
4. Di Carlo J., Wensing P.M., Katz B., Bledt G., Kim S. Dynamic locomotion in the MIT cheetah 3 through convex model-predictive control // IEEE International Conference on Intelligent Robots and Systems. – 2018. – Pp. 7440–7447.
5. Bledt G., Wensing P.M., Kim S. Policy-regularized model predictive control to stabilize diverse quadrupedal gaits for the MIT cheetah // IEEE International Conference on Intelligent Robots and Systems. – 2017. – Pp. 4102–4109.
6. He Li, Robert J. Frei, Patrick M. Wensing. Model Hierarchy Predictive Control of Robotic Systems // IEEE Robotics and Automation Letters. – 2021. – Pp. 3373–3380.
7. Kim D., Carlo J. Di, Katz B., Bledt G., Kim S. Highly dynamic quadruped locomotion via whole-body impulse control and model predictive control // arXiv. – 2019. DOI: 10.48550/arXiv.1909.06586.

8. Patrick Wensing (University of Notre Dame): Tailoring model complexity in MPC of legged locomotion. [Электронный ресурс]. – Режим доступа: <https://youtu.be/4EpH2fsXl88> (дата обращения: 11.03.2022).
9. Carius J., Farshidian F., Hutter M. MPC-Net: A First Principles Guided Policy Search // IEEE Robotics and Automation Letters. – 2020. – Pp. 2897–2904.
10. Gabriel B. Margolis, Tao Chen. Learning to Jump from Pixels. – 2021. DOI: 10.48550/arXiv.2110.15344.
11. Green K., Godse Y., Dao J., Hatton R.L., Fern A., Hurst J. Learning Spring Mass Locomotion: Guiding Policies with a Reduced-Order Model // IEEE Robotics and Automation Letters. – 2021. – Pp. 3926–3932.
12. Xue Bin Peng, Erwin Coumans, Tingnan Zhang, Tsang-Wei Edward Lee, Jie Tan, Sergey Levine. Learning Agile Robotic Locomotion Skills by Imitating Animals // arxiv. – 2020. DOI: 10.48550/arXiv.2004.00784.
13. Kim D., Carballo D., Di Carlo J., Katz B., Bledt G., Lim B., Kim S. Vision aided dynamic exploration of unstructured terrain with a small-scale quadruped robot // IEEE International Conference on Robotics and Automation. – 2020. – Pp. 2464–2470.
14. Boussema C., Powell M.J., Bledt G., Ijspeert A.J., Wensing P.M., Kim S. Online gait transitions and disturbance recovery for legged robots via the feasible impulse set // IEEE Robotics and Automation Letters. – 2019. – Pp. 1611–1618.
15. Bledt G., Kim, S. Implementing regularized predictive control for simultaneous real-time footstep and ground reaction force optimization // IEEE International Conference on Intelligent Robots and Systems. – 2019. – Pp. 6316–6323.
16. Bledt G., Kim, S. Extracting legged locomotion heuristics with regularized predictive control // IEEE International Conference on Robotics and Automation. – 2020. – Pp. 406–412.
17. Bledt G. Regularized predictive control framework for robust dynamic legged locomotion. Massachusetts Institute of Technology. – 2020. – 160 p.
18. Dudzik T., Chignoli M., Bledt G., Kim S. Robust autonomous navigation of a small-scale quadruped robot in real-world environments // IEEE International Conference on Intelligent Robots and Systems. – 2020. – Pp. 3664–3671.
19. Направления работы команды MIT по четвероногим роботам в виде Mind Map. [Электронный ресурс]. – Режим доступа: https://miro.com/app/board/uXjVOesowkY=?invite_link_id=728655214674 (дата обращения: 11.03.2022).
20. Bledt G., Powell M.J., Katz B., Di Carlo J., Wensing P.M., Kim S. MIT cheetah 3: Design and control of a robust, dynamic quadruped robot // IEEE International Conference on Intelligent Robots and Systems. – 2018. – Pp. 2245–2252.

УДК 004.05

ПОСТРОЕНИЕ СИСТЕМЫ ДЛЯ ОБНАРУЖЕНИЯ ПРЕПЯТСТВИЙ НА ЖЕЛЕЗНОДОРОЖНОМ ТРАНСПОРТЕ

Якимов Я.Д.¹ (студент)

Научный руководитель – канд. техн. наук Попов И.Ю.¹

¹*Университет ИТМО*

e-mail: yaroslav@itmo.ru

Статья посвящена построению систем, необходимых для развития беспилотных железнодорожных транспортных средств. Рассмотрены архитектура и компоненты систем для обнаружения объектов на железной дороге для установки на железнодорожный транспорт.

Ключевые слова: функциональная безопасность, железнодорожный транспорт, обнаружение препятствий, сенсоры, искусственный интеллект, снижение рисков.

Эффективная и безопасная работа железнодорожных систем в настоящее время зависит от автоматизированных систем управления (далее АСУ) поездами, основанных на взаимодействии бортового и наземного сигнального оборудования. При отказе АСУ машинист может двигаться только с крайне ограниченной скоростью, что приводит к нарушениям в работе транспорта в железнодорожной сети, как городских, так и в рамках междугороднего сообщения, что включает в себя как пассажирские, так и грузовые перевозки. В последние годы разрабатываются и устанавливаются в поездах интеллектуальные системы помощи при вождении на железной дороге (далее ИСП). Данные системы включают в себя множество различных систем, в том числе и системы камер, датчиков, лидаров и т. д. Информация обрабатывается с использованием алгоритмов искусственного интеллекта, для учета окружающей среды, инфраструктуры и для принятия решения, например, для торможения перед препятствием. Благодаря использованию ИСП машинист может управлять поездом на гораздо более высокой скорости. В то же время это также создает потенциальную проблему безопасности, когда водитель не успеет среагировать на те или иные возникающие проблемы такие как пожар, препятствие, застрявшие в дверях пассажиры. Основной целью системы является сокращение времени реакции для принятия мер по сохранению здоровья и жизни пассажиров в самом поезде, также снизить возможный риск причинения травмы субъекту, находящимся в опасной зоне движения электропоезда. Приведем наиболее известные международные решения в данной сфере.

Немецкий исследовательский проект КОМРА был направлен на разработку автоматизированной системы обнаружения препятствий. Система состояла из нескольких датчиков, включая видеокамеры, инфракрасные датчики и радарные системы [1]. Недостатком размещения датчиков на поезде является то, что они не охватывают необходимый диапазон на криволинейной дороге или холмистой местности.

Исследователи из Лаборатории прикладной физики Университета Джона Хопкинса [2] в Лореле разработали инструментальную систему контроля рельсов (IRiS). IRiS — это дистанционно управляемое железнодорожное транспортное средство, поэтому за ней должен следить оператор. IRiS является полуавтономным и может управляться из диспетчерской.

В [3] железнодорожные камеры были установлены в передней части поезда для определения препятствий на рельсах. Используя цифровые дополнительные препятствия, алгоритм обнаруживает большинство объектов и предупреждает, если объекты над рельсом могут создать опасность для безопасности поезда и его пассажиров.

С участием АО «НИИАС» были сформулированы требования к блоку обнаружения препятствий (далее БОП):

1. При нормальной видимости и уровне освещенности более 5 лк в соответствии с ГОСТ Р 54984-2012 БОП должен обнаруживать ростовую фигуру человека ростом более 120 см от уровня головки рельса на расстоянии не менее 600 метров при их нахождении на пути движения подвижного состава с вероятностью не ниже 95%.
2. При плохой видимости и очень плохой видимости БОП должен обнаруживать ростовую фигуру человека на расстоянии не менее, чем 30% от метеорологической дальности видимости.
3. БОП должен обнаруживать не менее четырех людей, представленных в виде ростовых фигур, в пределах дистанции обнаружения.
4. БОП должен определять дистанцию вдоль пути следования до обнаруженного человека с погрешностью не более 20%.
5. БОП должен определять зону местонахождения человека (красная, желтая) с вероятностью не ниже 95%.
6. БОП должен допускать не более 1 ложного обнаружения человека на 10 км пройденного пути.
7. При появлении человека в зоне видимости БОП должен обеспечивать формирование сообщения в систему управления за время не более 1 с.

Для детектирования объектов и соответственно препятствий, необходимо использовать набор различных аппаратных камер и датчиков, которые могли бы расширять информацию, которую получает программное обеспечение с использованием машинного обучения. После проведенных испытаний на полигоне и испытаний на участках Московского центрального кольца, были определены устройства, которые удовлетворяли бы требованиям ПО и вместе с ним соответствовали сформированным требованиям. Схема устройств представлена на рисунке 1.

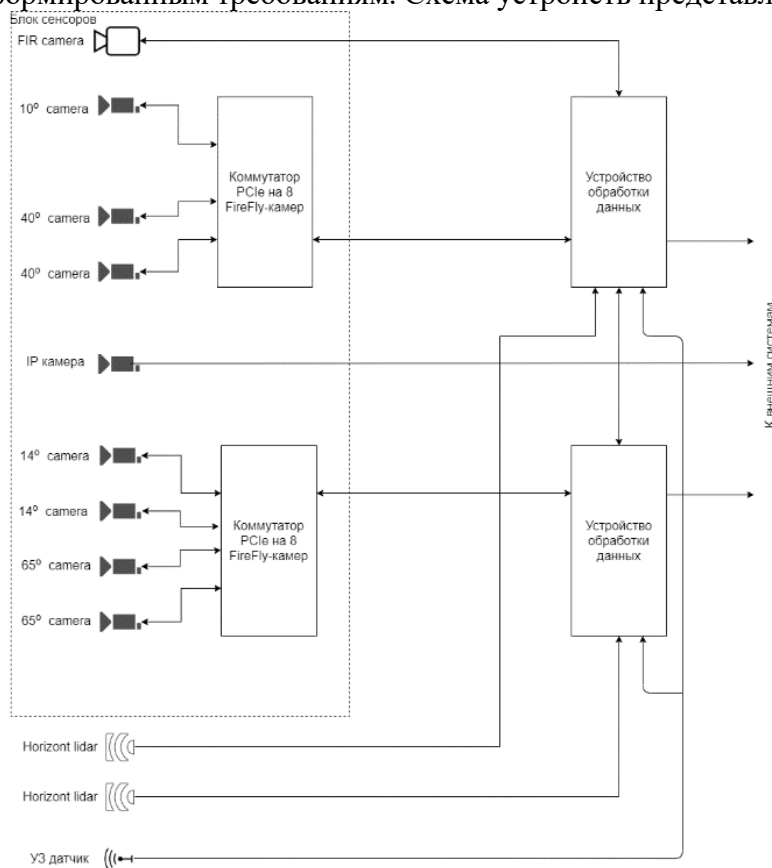


Рис. 1. Аппаратная составляющая БОП

Аппаратный набор включает в себя IP-камеру, 7 камер машинного зрения, парковочную систему с ультразвуковыми датчиками, тепловизор высокого разрешения и два лидара. Устройства позволяют построить пространственное изображение, в том числе с учетом

глубины и дальности до объекта. Тепловизоры необходимы для более точного определения живых объектов при плохой видимости. Так при использовании только камер во время дождя или снега снижается их дальность видимости и это напрямую влияет на точность работы алгоритмов машинного, а тепловизоры позволяют нивелировать данный недостаток.

Программное обеспечение призвано собирать данные с камер и датчиков, проводить их обработку на вычислительных комплексах, установленных внутри самого поезда также в 2-х экземплярах. Данные программные модули разрабатываются и модифицируются командами разработчиков АО НИИАС с учетом полевых тестов. На рисунке 2 представлен наглядный алгоритм работы установленного ПО. Важным критерием разработки и использования данного ПО является время обработки, и сообщение о наличии препятствия должно поступать в систему реагирования в течение 1 с.

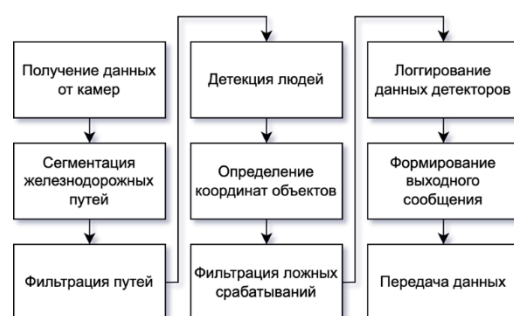


Рис. 2. Программная составляющая БОП

Программное обеспечение БОПа подразделяется на следующие основные классы:

- системное программное обеспечение, включающее драйверы сенсоров;
- прикладное программное обеспечение, решающее задачи обнаружения препятствий.

Системное ПО включает:

- драйвер захвата данных видеокамер;
- драйвер захвата данных от лидара.

Прикладное программное обеспечение включает:

- ПО обнаружения препятствий по данным камер;
- ПО обнаружения препятствий по данным лидара;
- ПО комплексирования данных о препятствиях от разных сенсоров.

На рисунке 3 представлены кадры из цифрового симулятора, который позволяет наглядно увидеть, какие данные поступают для обработки, и как беспилотный «видит» внешнюю среду.

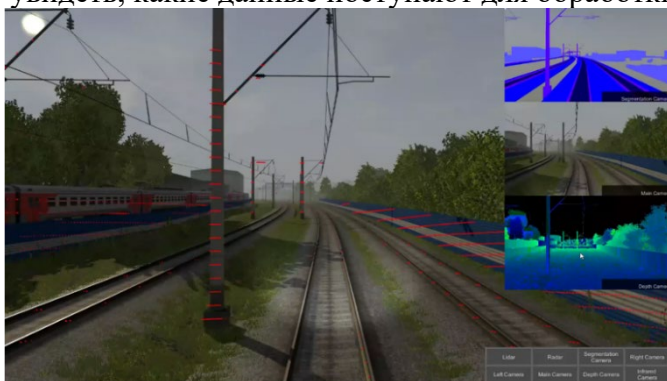


Рис. 3. Пример работы аппаратной и программной составляющих БОП

Проблема обнаружения препятствий остается всё еще достаточно сложной для однозначного решения. Так как существует множество факторов, влияющих на точность распознавания, а также не стоит забывать про стоимость установки необходимых компонентов и разработку ПО. На данный момент АО «НИИАС» является одним из лидеров в разработке систем для беспилотных железнодорожных транспортных средств и в планах есть запуск данной системы в серийное производство и внедрение его на текущую инфраструктуру АО «РЖД» для введения в эксплуатацию поездов GoA4.

Литература

1. Ruder M., Mohler N., Ahmed F. An obstacle detection system for automated trains // IEEE IV2003 Intelligent Vehicles Symposium. Proceedings (Cat. No. 03TH8683). – 2003. – С. 180–185.
2. Johns Hopkins University. Harsco Corp's Protran Technology, "Railway Gazette". [Электронный ресурс]. – Режим доступа: <https://www.railwaygazette.com/news/techn%20ology/single-view/view/johns-hopkins-develops-iris-remote-%20control-inspection-system.html> (дата обращения: 20.01.2022).
3. Rodriguez L A.F., Uribe J.A., Bonilla J.F.V. Obstacle detection over rails using hough transform // 2012 XVII Symposium of Image, Signal Processing, and Artificial Vision (STSIVA). – IEEE. – 2012. – С. 317–322.

ФОТОНИКА

УДК 666.221.6

РАЗРАБОТКА АКТИВНОГО ЭЛЕМЕНТА ОПТИЧЕСКОГО ДАТЧИКА ИСКРОВОГО ПРОБОЯ ДЛЯ РАСПРЕДЕЛИТЕЛЬНЫХ ЩИТОВ ЭЛЕКТРОСЕТЕЙ

Азина Л.В.¹ (аспирант), Панов Д.Ю.¹ (аспирант), Спиридонов В.А.¹ (аспирант), Волковицкий В.Е.¹ (студент)

Научный руководитель – д-р физ.–мат. наук, профессор Романов А.Е.¹

¹Университет ИТМО

e-mail: lvazina@itmo.ru, dmitriipnv@gmail.com

Работа выполнена в рамках темы НИР №421244 «Разработка оптического датчика искрового пробоя для распределительных щитов электросетей».

В работе рассмотрен этап изготовления образцов активного элемента оптического датчика искрового пробоя для комплектных распределительных устройств. Материал активного элемента представляет оптическое стекло на основе литий-бор-фосфатной матрицы с легированием Eu. Исследованы оптические свойства полученных стекол, также показана возможность применения таких материалов в качестве down-конвертеров (конвертеры оптического излучения, реализующие смещение спектра излучения в область больших длин волн), и, в частности, в разрабатываемом датчике.

Ключевые слова: датчик искры, дуговая защита, стеклокерамика, европий, стеклообразная матрица.

Представленные на сегодняшний день на рынке оптические датчики искрового пробоя для комплектных распределительных устройств основаны на регистрации вспышки оптического излучения открытой электрической дуги, которая возникает вследствие короткого замыкания. Как детектор и передатчик оптического излучения электрической дуги в таких устройствах используют оптические волокна – кварцевое оптическое волокно или пластиковое оптическое волокно. Но многократное использование таких датчиков невозможно, т.к. при каждом возникновении короткого замыкания выделяется большой поток энергии в виде УФ излучения и теплового излучения, что приводит к разрушению Polymer Optical Fiber – POF (пластикового волокна). Как приемник оптического сигнала применяют кремневые фотодиоды, которые имеют пик чувствительности в ИК области, а чувствительность в УФ диапазоне составляет порядка 5%, вследствие этого датчики регистрируют только часть спектра излучения дугового разряда (в видимой и ближней ИК области).

Проблему снижения порога чувствительности существующих оптических дуговых защит, возникающую вследствие регистрации излучения дугового разряда только в видимой области (т.е. части спектра излучения дугового разряда), можно решить регистрацией части спектра излучения открытой электрической дуги, лежащей в ближней УФ области спектра (~300 – 400 нм). Такое решение позволит повысить чувствительность устройства и избежать ложных срабатываний оптического датчика искрового пробоя. Для этого необходимо разработать активный элемент, способный конвертировать УФ излучение в видимое или ИК излучение. Стеклокерамика является одним из возможных материалов активного элемента.

Стеклокерамика и стеклокерамические материалы – это композитные материалы, образованные нанокристаллической фазой, распределенной в стеклообразной матрице. При правильном легировании стекольной матрицы редкоземельными материалами, получаемые стекла сочетают в себе оптическую прозрачность и свойства, необходимые для создания

приборов фотоники. Стеклообразные матрицы с новыми оптическими свойствами находят применение в различных устройствах фотоники, например, оптических усилителях, оптических волноводах, белых светодиодах, датчиках искры и т.д. [1, 2].

В последнее время внимание исследователей направлено на разработку материалов, в которых стеклообразная матрица влияет на свойства ионов из группы лантаноидов Ln^{3+} . Целью таких работ является уменьшение многофононной безизлучательной релаксации и, вследствие, повышение квантового выхода люминесценции иона Ln^{3+} . Оксидные стекла оказались наиболее подходящими в качестве матрицы для ионов Ln^{3+} благодаря их прозрачности, возможности варьирования состава и достаточной простоте технологии получения [3]. Среди оксидных стекол особое внимание привлекают фосфатные стекла, т.к. они обладают уникальными характеристиками, такими как: высокий коэффициент пропускания, низкая температура плавления (синтеза), малый показатель преломления и низкая дисперсность [4]. Присутствие Al_2O_3 в таких стеклах повышает не только механическую прочность, но и концентрацию ионов лантаноидов Ln^{3+} [5].

Среди ионов из группы лантаноидов Ln^{3+} ион европия (Eu^{3+}) является одним из лучших кандидатов для использования в качестве люминофора с красным излучением из-за узкого и монохроматического перехода $5\text{D}_0-7\text{F}_2$, соответствующего длине волны около 610 нм [6].

Цель данного исследования – разработать оптимальный состав матрицы $\text{Al}_2\text{O}_3\text{-K}_2\text{O-Li}_2\text{O-B}_2\text{O}_3\text{-P}_2\text{O}_5$, активированной ионами европия, исследовать оптические свойства получаемой стеклокерамики, и оценить возможность применения разработанного материала в оптическом датчике искрового пробоя.

Методология

Технология получения материала состоит из четырех этапов: синтез, отжиг, термообработка и итоговая подготовка образцов. Синтез получаемых стекол проводился методом высокотемпературной плавки шихты в тигле в шахтной электропечи «ТК.13.1500.Ш.1Ф». Для получения расплава использовались корундовые тигли, температура синтеза составляла $T = 1450^\circ\text{C}$. Пробподготовка образцов осуществлялась на пиле Buehler isomet 1000, которая позволяет получать плоскопараллельные образцы. Образцы подготавливались размером 15x15x10мм.

Спектры оптического пропускания были получены на спектрометрической установке Avantes. Спектры оптической люминесценции были исследованы на установке Horiba. Измерения проводились при комнатной температуре

Результаты и обсуждения

Были проведены исследования оптических свойств синтезированных стекол. Спектры оптического поглощения полученных образцов приведены на рисунке 1. Как видно из рисунка 1, синтезируемые литий-бор-фосфатные стекла, легированные Eu , прозрачны в видимом и ближнем ИК диапазоне, и имеют интенсивное поглощение в УФ диапазоне до ~300 нм. На графиках в области 393 нм и 461 нм наблюдается увеличение поглощения, что соответствует переходам $7\text{F}_0-7\text{F}_6$ и $7\text{F}_1-7\text{F}_6$ [7–9].

В литературе описывается наличие 11-ти переходов: $7\text{F}_1 \rightarrow 7\text{F}_6$, $7\text{F}_0 \rightarrow 7\text{F}_6$, $7\text{F}_1 \rightarrow 5\text{D}_1$, $7\text{F}_0 \rightarrow 5\text{D}_1$, $7\text{F}_0 \rightarrow 5\text{D}_2$, $7\text{F}_1 \rightarrow 5\text{D}_3$, $7\text{F}_1 \rightarrow 5\text{L}_6$, $7\text{F}_0 \rightarrow 5\text{L}_6$, $7\text{F}_0 \rightarrow 5\text{G}_2$, $7\text{F}_0 \rightarrow 5\text{G}_4$ и $7\text{F}_0 \rightarrow 5\text{D}_4$, которые соответствуют полосам поглощения на 2207, 2086, 533, 526, 465, 414, 400, 393, 382, 376 и 362 нм соответственно [7]. Как правило, полосы в спектрах поглощения ионов редкоземельных элементов возникают из-за внутриконтинуальных $f-f$ переходов. Переходы поглощения $7\text{F}_J \rightarrow 5\text{D}_J$ запрещены по спине, и, следовательно, они очень слабы, а переходы $7\text{F}_0 \rightarrow 5\text{L}_6$ и $7\text{F}_0,1 \rightarrow 7\text{F}_6$ оказываются более интенсивными, чем все остальные [7–9].

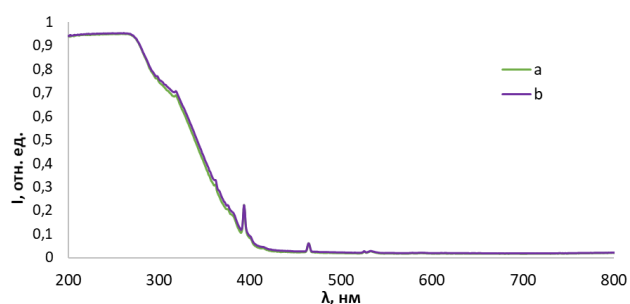


Рис. 1. Спектры поглощения бор-фосфатного стекла, легированного Eu , а – исходный состав, б – модифицированный состав с добавлением S

На рисунке 2 приведены спектры люминесценции литий-бор-фосфатного стекла, активированного Eu, при возбуждении длиной волны $\lambda_{\text{возб.}} = 400$ нм. На рисунке 2 наблюдаются полосы излучения с пиками 578, 589, 611, 655 и 702 нм. Как видно на графике, наиболее интенсивное излучение приходится на область 611 нм. При модификации состава с добавлением серы S спектры люминесценции схожи со спектром исходного состава, что говорит о том, что в ходе синтеза сера не встроилась в матрицу и, вероятно, во время синтеза улетучилась из состава. Полученные результаты соответствуют литературным данным: полосы излучения люминесценции в области длин волн 580, 590, 615, 655 и 700 нм соответствуют 4f-4f-переходам $5D_0 \rightarrow 7F_J$ ($J = 0, 1, 2, 3, 4$) в ионах Eu^{3+} , и при этом на длине волны ~ 615 нм ($5D_0 \rightarrow 7F_2$) наблюдается наиболее интенсивный пик в сравнении с другими пиками [7–11]. Таким образом, можно говорить о том, что в получаемом стекле образовались свободные ионы Eu^{3+} .

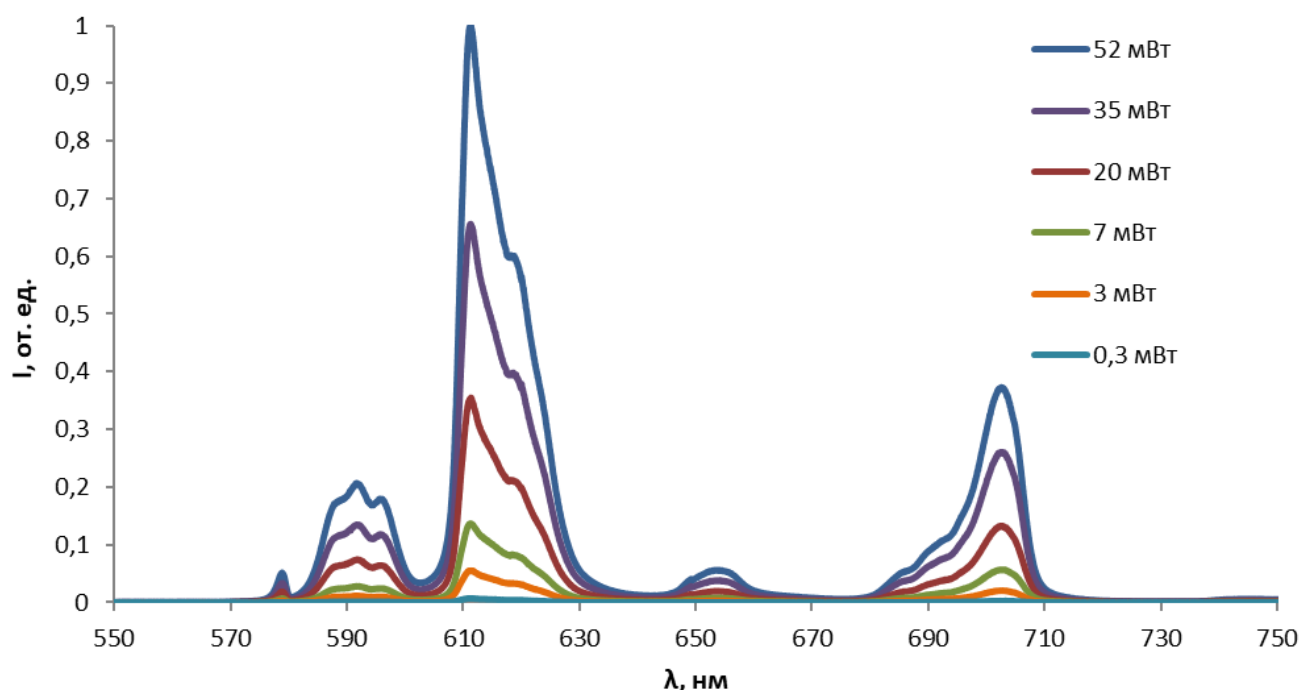


Рис. 2. Спектры фотолюминесценции литий-бор-фосфатного стекла, легированного Eu, при $\lambda_{\text{возб.}} = 400$ нм в зависимости от мощности возбуждающего излучения

Дополнительно была оценена зависимость интенсивности люминесценции от мощности возбуждающего излучения. Как видно из рисунка 2 интенсивность люминесценции прямо пропорционально зависит от мощности возбуждающего излучения. Максимальная мощность излучения лазера составляла 52 мВт и в последующем понижалась с помощью светофильтров.

Как видно из представленных данных, состав обладает конвертацией УФ излучения в диапазон ближнего ИК, имеет химическую стабильность и относительно невысокую температуру синтеза. Такие показатели делают данный состав возможным к использованию в качестве активного элемента оптического датчика искры и down-конвертеров.

В ходе работы были синтезированы оптические стекла литий-бор-фосфатной матрицы, активированные ионами европия. Исследованы оптические свойства полученных образцов, и оценены возможности дальнейшего практического применения стекла в качестве активного материала оптического датчика искрового пробоя.

Литература

1. Liu X., Zhou J., Zhou S., Yue Y., Qiu J. Transparent glass-ceramics functionalized by dispersed crystals // *Prog. Mater. Sci.* – 2018. – №97. – Pp. 38–96.
2. De Pablos-Martín A., Durán A., Pascual M.J. Nanocrystallisation in oxyfluoride systems: Mechanisms of crystallisation and photonic properties // *Int. Mater. Rev.* – 2018. – 57. – Pp. 165–186.

3. Wang J., Song H., Kong X., Peng H., Chen B., Zhang J., Xu W. Fluorescence properties of trivalent europium doped in various niobate codoped glasses // *J. Appl. Phys.* – 2003. – №93. – Pp. 1482–1486.
4. Suratwala T.I., Steele R.A., Wike G.D., Campbell J.H., Takenchi K. Effects of OH content, water vapor pressure, and temperature on sub-critical crack growth in phosphate glass // *J. Non Cryst. Solids.* – 2000. – №263–264. – Pp. 213–227.
5. Arai K., Namikawa H., Kumata K., Honda T., Ishii Y., Handa T. Aluminum or phosphorus codoping effects on the fluorescence and structural properties of neodymium doped silica glass // *J. Appl. Phys.* – 59. – 1986. – Pp. 3430–3436.
6. Day D.E. Mixed alkali glasses—their properties and uses // *Journal of Non-Crystalline Solids.* – 1976. – Т.21. – №3. – С. 343–372.
7. Nayab Rasool S., Rama Moorthy L., Kulala Jayasankar C. Optical and luminescence properties of Eu³⁺-doped phosphate based glasses // *Materials Express.* – 2013. – Т.3. – №3. – Pp. 231–24.
8. Marimuthu K. et al. Structural and optical studies of Eu³⁺ ions in alkali borate glasses // *Physica status solidi (a).* – 2009. – Т.206. – №1. – С. 131–139.
9. Swapna K. et al. Luminescence characterization of Eu³⁺ doped Zinc Alumino Bismuth Borate glasses for visible red emission applications // *Journal of luminescence.* – 2014. – Т.156. – С. 80–86.
10. Lian Z. et al. The reduction of Eu³⁺ to Eu²⁺ in air and luminescence properties of Eu²⁺ activated ZnO–B₂O₃–P₂O₅ glasses // *Journal of Alloys and Compounds.* – 2007. – Т.430. – №1-2. – С. 257–261.
11. Пустоваров В.А. Люминесценция твердых тел: учебное пособие / Пустоваров В.А. — Екатеринбург: Издательство Уральского университета. – 2017. – С. 128.

УДК 628.973

ВЛИЯНИЕ ИСКУССТВЕННОГО СВЕТА НА АТМОСФЕРУ ОБЩЕСТВЕННЫХ ПРОСТРАНСТВ

Астахов Ю.И.¹ (студент)

Научный руководитель – канд. архитектуры, доцент Соколова М.А.¹

¹*Университет ИТМО*

e-mail: iurii.ast@gmail.com

Атмосфера является ключевым параметром, влияющим на эмоциональную оценку пространства пользователем. В работе проведен анализ световых решений для определенных общественных пространств, рассмотрены основные характеристики, влияющие на атмосферу.

Ключевые слова: искусственный свет, общественные пространства, атмосфера, эмоциональный отклик.

Атмосфера является важным параметром в определении эмоционального отклика пользователя при его взаимодействии с пространством [1]. Плюсом её рассмотрения является то, что в основе восприятия атмосферы лежат эмоции, которые имеют большую скорость, чем мысли или осознание. Таким образом, существует возможность быстрой передачи идеи пространства пользователю.

Восприятие атмосферы зависит от предыдущего опыта пользователя и комплекса сенсорных стимулов: визуальных, звуковых, тактильных, обонятельных, что делает изучение влияния отдельных параметров проблематичным и требует комплексного рассмотрения каждой ситуации [2].

По мнению лауреата Притцкеровской премии Петера Цумтора показателем качества архитектурного пространства является именно атмосфера. Одним из важнейших параметров, формирующих атмосферу, архитектор называет свет [3]. Возможность формировать атмосферу при помощи искусственного света делает его эффективным инструментом, помогающим пользователю понять идею пространства и взаимодействовать с ним, и не вносящим изменения в интерьер или архитектуру пространства.

В настоящее время понятие атмосферы становится важным для музейных пространств [4]. Тем не менее, измерение этого параметра для других общественных пространств позволяет создать качественно новые дизайн решения. Понимание основных параметров, влияющих на атмосферу, позволяет проводить моделирование настроения пользователя при взаимодействии с пространством в процессе проектной работы.

В рамках исследования подходов к формированию атмосферы общественных интерьеров и открытых пространств средствами искусственного света проведен анализ актуальных световых решений, рассмотрены основные характеристики, влияющие на атмосферу. Рассмотренные световые решения воздействуют на восприятие не только характеристиками света – распределением света в пространстве, цветовой температурой источников, контрастом, но и их взаимодействием с материалами и формой. Совокупность используемых приемов позволяет создавать световые решения, вызывающие различный эмоциональный отклик пользователей – уютные и камерные световые среды, располагающие к длительному взаимодействию, либо напротив, побуждающие к активному времяпрепровождению. Искусственный свет, как эффективный инструмент формирования средовой атмосферы

требует от проектировщика точных и взвешенных решений, основанных на контекстуальном применении световых характеристик в соответствии с выбранным для определенного пространства сценарием и ожидаемым эмоциональным откликом пользователей.

Одними из первых исследований, посвященных атмосфере, являются исследования, связанные атмосферой магазинов. Окружающая среда магазинов должна повышать уровень продаж, привлекать покупателя. В маркетинговых исследованиях отмечается, что атмосфера влияет не только на клиента, но и на работников заведения, компании [5].

Свет, материалы, организация пространства работают как инструменты создания атмосферы. На примерах двух магазинов одежды видна возможная разница между дизайном пространств.

Камерное пространство магазина А.Р.С. создано с использованием материалов на основе дерева (рис. 1) Распределение света неравномерно, сильно заметны отбрасываемые тени от предметов интерьера. Пространство может быть в некотором смысле отвлечённым, в то же время побуждающим к исследованию. Распределение света ограничено с помощью подвешенных фанерных пластин под потолком. В противоположность неравномерному освещению магазина А.Р.С. можно представить освещение магазина Balenciaga (рис. 2) В магазине используются металлические материалы в качестве полок, бетонные стены без финиширования, зеркала. Освещение помещения равномерное и более близко к холодному, чем в магазинах А.Р.С. Источники света, направленные непосредственно на вещи, установлены с рассеивателями, которые повышают равномерность. Вдоль стен установлено контурное отраженное освещение. Взаимодействие материалов и света создаёт общую футуристическую стилистику пространства. Достаточно полярные примеры дизайна магазинов показывают возможности воплощения желаемой атмосферы с помощью не только пространственной организации и материалов, но и примененной световой концепции.



Рис. 1. Фотографии интерьера магазина А.Р.С. в Берлине [6]

Атмосфера является важным параметром не только для магазинов, но и для образовательных учреждений, коворкингов. Доказано, что для посетителей коворкингов атмосфера является одним из важнейших мотиваторов. Сохранение позитивного настроения учащегося и комфортная среда в свою очередь повышают качество изучения предметов.

Принципы освещения, которые могут быть применены как к световому дизайну в целом, так и к решению задач по формированию атмосферы были сформулированы Ричардом Келли. Американский светодизайнер предложил три элемента освещения: focal glow (световой акцент), ambient luminescence (обволакивающий свет), play of brilliants (игра света). По Ричарду Келли световой акцент повышает видимость и обращает внимание на объекты,

обволакивающий свет делает пространство безопасным и понятным, игра света создаёт особый дух пространства. Три перечисленных элемента являются основными составляющими, которые позволяют сформировать световой образ пространства. Важно отметить образный подход автора к описанию элементов. В статье приводятся аналогии не только с искусственным освещением, но и с природными световыми явлениями.



Рис. 2. Фотография интерьера магазина Balenciaga в Берлине

На данный момент проведено множество исследований, посвященных исследованию влияния световых характеристик на атмосферу пространств. Измеряются такие характеристики как цветовая температура, распределение света, яркость, структура света (рассеянный или прямой свет), скорость динамического освещения.

Таким образом, существуют данные, что прямой свет является более уютным и живым по сравнению с рассеянным. Холодные цветовые температуры высокой интенсивности повышают оживлённость пространства. В то же время при холодных цветовых температурах помещение воспринимается как более просторное. В динамическом освещении стоит отметить, что интенсивность, меняющаяся во времени по логарифмической зависимости, воспринимается как более живая, но менее комфортная.

При формировании атмосферы общественных пространств важно использовать знания из различных областей: светового дизайна, архитектуры, психологии окружающей среды. Современные исследования позволяют понять, как характеристики световой среды могут повлиять на атмосферу пространства. Для создания атмосферы требуется комплексный подход к дизайну. Должны учитываться не только фотометрические характеристики, но и структура пространства, характер взаимодействия света с материалами.

Литература

1. Vogels I. Atmosphere Metrics // Probing Experience: From Assessment of User Emotions and Behaviour to Development of Products. – 2008. – 250 с.
2. Myung E.C., Mi J.K. Measurement of User Emotion and Experience in Interaction with Space // Journal of Asian Architecture and Building Engineering. – 2017. – 8 с.
3. Zumtor P. Atmospheres // Bazel: Birkhäuser. – 2006. – 76 с.
4. Brigitte B., Dirk L. Aesthetics and Atmosphere in Museums: A Critical Marketing Perspective // Museum Media, Part 2. Mediation and Immersion. – 2015. – С. 251–257.
5. Bohl P. The effects of store atmosphere on shopping behaviour // A literature review. – 2012. – 18 с.
6. Фотографии интерьера магазина А.Р.С. в Берлине. [Электронный ресурс]. – Режим доступа: <https://de.foursquare.com/> (дата обращения: 22.02.2022).

УДК 681.7.013.3

ИССЛЕДОВАНИЕ ВОЗМОЖНОСТИ УВЕЛИЧЕНИЯ РАБОЧЕГО ПОЛЯ ПРИ ИСПОЛЬЗОВАНИИ НЕПЛОСКИХ ПРИЕМНЫХ ПЛОЩАДОК ЗЕРКАЛЬНЫХ КАМЕР

Безуглый А.М.^{1,2} (аспирант)

Научный руководитель – канд. техн. наук, доцент Бахолдин А.В.¹

¹Университет ИТМО

²АО «НПП «ЭЛАР»

e-mail: ambezuglyi@itmo.ru

В работе представлены основные результаты исследования возможности увеличения рабочего поля зеркальных камер с применением неплоских приемных площадок. Приводятся классификация причин ограничения поля изображения камер и технологически возможные решения исполнения неплоских площадок.

Ключевые слова: зеркальная камера, рабочее поле, неплоская приемная площадка, ограничение поля изображения, улучшение оптических систем.

В течение длительного времени в мире наблюдается устойчивое развитие сфер дистанционного зондирования Земли (ДЗЗ) и исследования космоса. Из-за увеличения количества решаемых задач космических аппаратов (КА), сложности проектирования и разработки, а также себестоимости будущих приборов для астрономических исследований и ДЗЗ требуются новые технологии, позволяющие создавать более компактные и простые приемные модели. Значительная часть проблем, возникающих при проектировании и создании КА может быть решена за счет улучшения оптической системы (ОС) приемного канала оптико-электронного устройства. Большое количество иностранной литературы утверждает, что применение неплоской структуры приемной площадки позволит решить проблемы, возникающие при проектировании и создании приемного тракта [1–3]. Установка криволинейных площадок с приемным устройством позволяет уменьшить количество оптических элементов в системе, при этом повышая ее оптические характеристики.

В каждой оптической системе имеются геометрические и оптические ограничения размера поля изображения. В качестве оптических ограничений выделяют полевые диафрагмы,

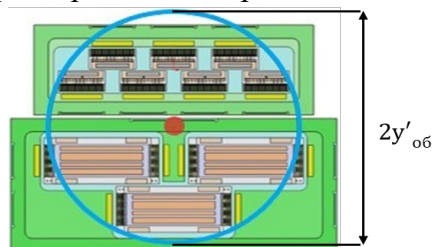


Рис. 1. Габаритные размеры приемной площадки

виньетирование и полевые аберрации. Такие аберрации проявляются вследствие использования не идеальной оптики и характеризуют собой нарушение гомоцентричности пучка света. Особый интерес представляют геометрические ограничения: габаритные размеры приемной площадки и направление угла падения излучения на поверхность фотозоны приемного устройства. На приемной площадке создается большое количество слепых зон, что негативно влияет на поле изображения и качество получаемых результатов работы. На рисунке 1 красным цветом выделена область наибольшей эффективности засветки, а в зоне, обозначенной синим цветом – наименьшие значения эффективности.

В качестве решения задачи изменения рабочего поля и коррекции aberrаций могут быть предложены различные структуры криволинейных приемных площадок. Важным условием является согласованность структуры криволинейной площадки и оптической системы.

Первая структура (рис. 2) представляет из себя единую криволинейную поверхность, для которой должно выполняться следующее условие: соотношение размера поля, создаваемого при помощи оптического тракта, меньше или равно размеру фотозоны приемного устройства. В процессе согласования приемной площадки и оптической системы, а именно при подборе эффективного и технологически возможного радиуса кривизны поверхности площадки можно скорректировать поле. Кроме того, полевые aberrации также подвергнутся корректировке. Такую структуру площадки можно осуществить при использовании КМОП фотоустройств, которые найдут свое применение в таких устройствах, как микроскопы, фототехника, датчики для астрономической навигации.

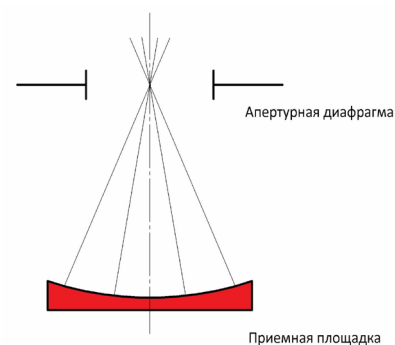


Рис. 2. Структура единой криволинейной площадки

Вторая структура (рис. 3) представляет из себя криволинейную поверхность с сегментированным расположением плоских ПЗС приемников, для которой выполняется следующее условие: соотношение размера поля, создаваемого при помощи оптического тракта, больше размера фотозоны приемного устройства. Свое применение такая конфигурация найдет в сфере дистанционного зондирования Земли. Все описанные выше преимущества первой структуры характерны и для второй.

В ходе работы были проанализированы геометрические и оптические ограничения размера поля изображения, различные конфигурации неплоских приемных площадок.

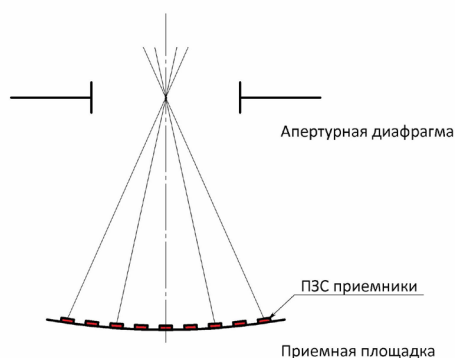


Рис. 3. Структура сегментированной криволинейной площадки

Для каждой оптической системы расчет эффективного радиуса кривизны приемной поверхности необходимо проводить индивидуально в соответствии с выполняемой задачей. В будущих работах планируется провести сравнительный анализ плоской и неплоской приемных площадок камеры с КМОП приемником, оценку эффективной площади засветки приемников излучением.

Литература

1. Chambion B., Gaschet C., Behaghel T., Vandeneynde A., Caplet S., Gétin S., Henry D., Hugot E., Jahn W., Lombardo S., Ferrari M. Curved sensors for compact high-resolution wide-field designs: prototype demonstration and optical characterization // Proc. SPIE 10539. Photonic Instrumentation Engineering. – V.5. – 2018. DOI: 10.1117/12.2272451.
2. Guenter B., Joshi N., Stoakley R., Keefe A., Geary K., Freeman R., Hundley J., Patterson P., Hammon D., Herrera G., Sherman E., Nowak A., Schubert R., Brewer P., Yang L., Mott R., McKnight G. Highly curved image sensors: a practical approach for improved optical performance // Optics Express. – 2017. – V.25. – №12. DOI: 10.1364/OE.25.013010.
3. Iwert O., Delabre B. The challenge of highly curved monolithic imaging detectors // Proc. SPIE 7742. High Energy, Optical, and Infrared Detectors for Astronomy. – V.6. – 2010. DOI: 10.1117/12.856719.

УДК 621.397.42

АНАЛИЗ ХАРАКТЕРИСТИК СИСТЕМ ВИДЕОНАБЛЮДЕНИЯ НА ОСНОВЕ АВТОНОМНЫХ И ОБЫЧНЫХ КАМЕР

Бердюгина Д.А.¹, Пастухов А.С.²

Научный руководитель – канд. техн. наук, ассистент Смирнов Ю.Ю.³

¹ООО “Софтланч”

²ООО “ТРАФО”

³Университет ИТМО

e-mail: aksamidnaya@gmail.com, artem.pastukhov1984@gmail.com, yuri.smirnov@itmo.ru

В работе рассматриваются варианты применения получающих в последние годы все большее внедрение автономных видеокамер в системах безопасности. Анализируется эффективность их применения в различных конфигурациях систем безопасности по сравнению с применением классических видеокамер.

Ключевые слова: системы безопасности, классические видеокамеры, автономные видеокамеры, система охраны периметра, глубина видеоархива.

В настоящее время растёт популярность внедрения автономных устройств в системах безопасности. Применение автономных устройств в системах безопасности дает системам безопасности целый ряд преимуществ, таких как невозможность отключить систему безопасности путем вывода из строя систем электроснабжения и передачи данных, масштабируемость и другие. Актуальной задачей является анализ эффективности применения подобных устройств для использования в системах безопасности различных конфигураций. Нашей задачей было исследовать системы видеонаблюдения, решающие различные задачи, построенные на основе классических и автономных видеокамер и оценить преимущества использования первых и вторых в данных системах видеонаблюдения. Для сравнения мы будем использовать объект типовой конфигурации, данный объект является базой отдыха элитного уровня. Общая площадь объекта 26 400 м², длина периметра 680 м. Исходя из параметров объекта нами были подобраны видеокамеры: классическая видеокамера Smart-камера Hikvision DS-2CD4585F-IZH [1] и автономная видеокамера Reolink Go [2]. Помимо видеокамеры Smart-камера Hikvision DS-2CD4585F-IZH, обладающей разрешением в 8 Мп, мы также подобрали для анализа 6 Мп видеокамеру Hikvision DS-2CD4565F-IZH [3]. В целом сравнивая классические и автономные видеокамеры, представленные на рынке, можно отметить, что классические камеры имеют более хорошие объективы, большую дальность ИК подсветки, дополнительные возможности по обнаружению лиц на уровне программного обеспечения. Рассмотрим системы видеонаблюдения объекта: территории объекта, систему охраны периметра и систему охраны прилегающей местности. На рисунке 1 представлена расстановка видеокамер системы видеонаблюдения на территории базы на основе классических видеокамер.

При замене классических видеокамер на автономные видеокамеры, чтобы обеспечить ту же эффективность работы системы, необходимо будет применить большее количество видеокамер, кроме того, передача данных по радиоканалу может прерваться при атмосферных осадках и других вызывающих помехи явлениях и событиях.

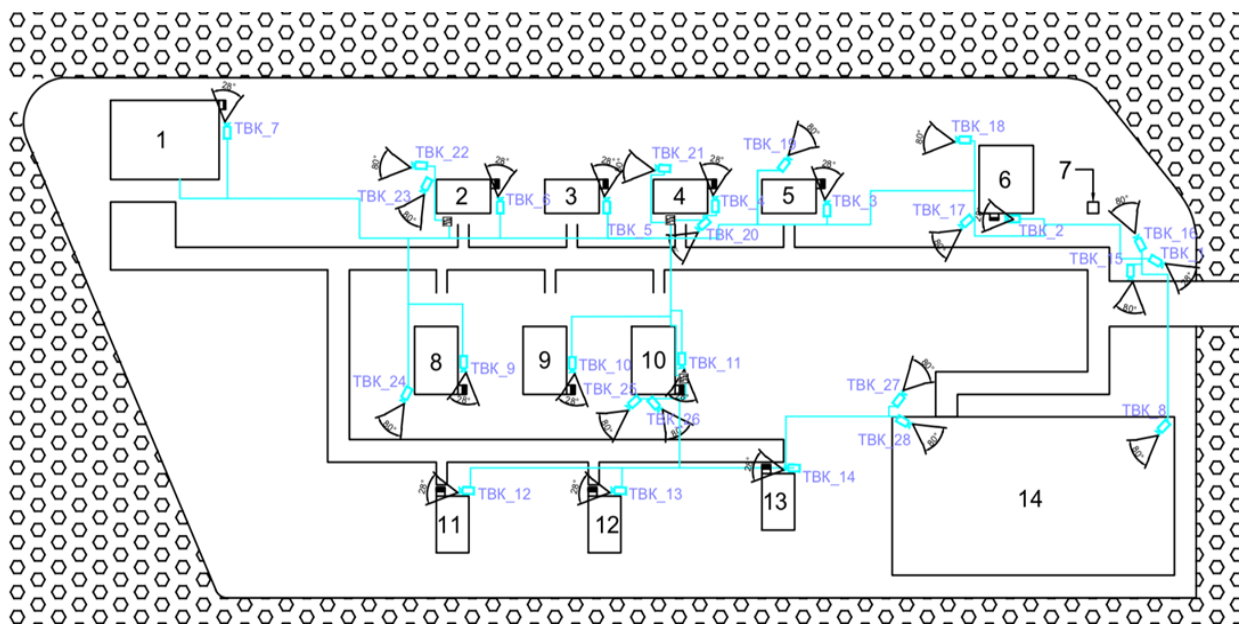


Рис. 1. Схема расстановки видеокамер системы видеонаблюдения на основе классических видеокамер

При этом стоит отметить, что система автономных видеокамер не требует наличие сетевой инфраструктуры, также менее уязвима при действиях квалифицированного нарушителя, чем систем на основе классических видеокамер, кроме того, к достоинствам применения автономных видеокамер стоит отнести возможность применения в местах, где классические видеокамеры не разместить и более низкую стоимость. Оценивая глубину видеоархива [4], которая необходима для хранения данных с видеокамер для системы видеонаблюдения территории в течении 7 дней, можем сказать, что для автономных видеокамер необходимая глубина видеоархива составит 14 ТБ, количество видеокамер 65 шт., а для классических видеокамер необходимая глубина видеоархива составит 49 ТБ, 6 Мп видеокамер 35 шт., 8 Мп видеокамер 13 шт.

Далее рассмотрим систему охраны периметра. Для анализа системы охраны периметра мы будем использовать автономную видеокамеру VolyGuard BG310-M 4G [5]. На рисунке 2 приведена расстановка автономных видеокамер по периметру объекта.

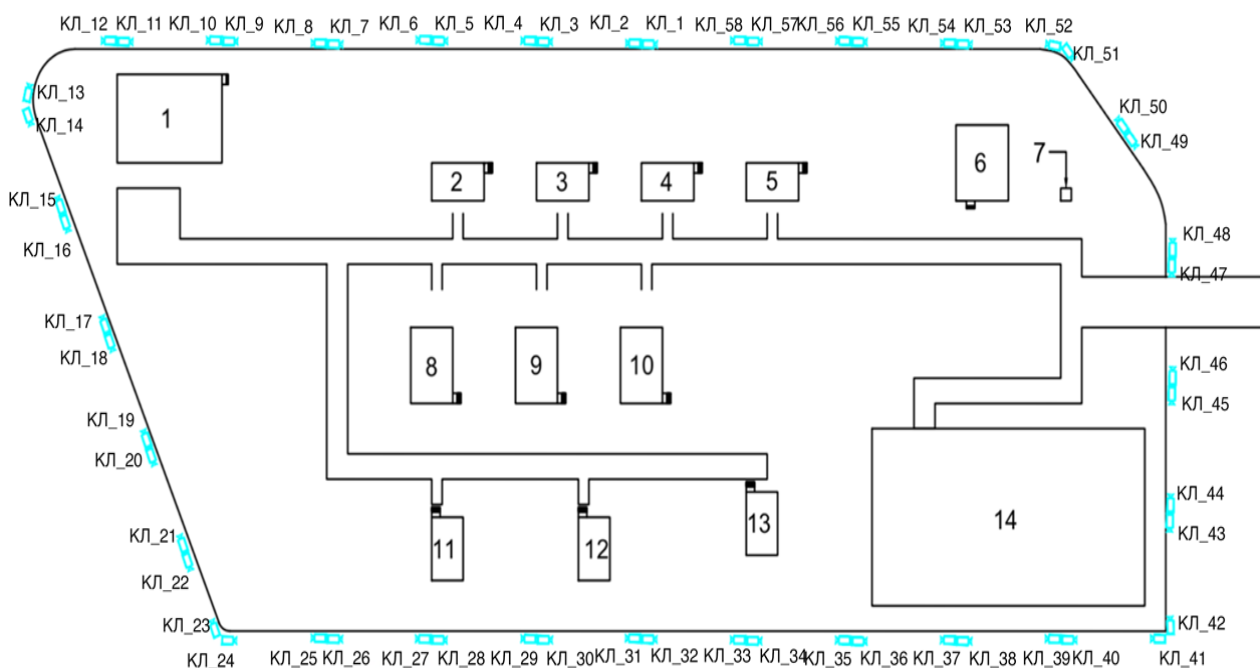


Рис. 2. Схема расстановки автономных видеокамер системы охраны периметра

Задачей системы в данном случае является регистрация нарушения периметра, поэтому камеры можно запустить в режиме регистрации событий по движению. В процессе эксплуатации базы периметр базы может быть расширен или изменен каким-либо образом. Оценивая необходимую глубину видеоархива, получаем при использовании автономных видеокамер 1,4 ТБ, количество видеокамер 58 шт., длина одного видеофрагмента 1 мин., количество видеофрагментов за день 300 шт., срок хранения данных 7 дней, при использовании классических видеокамер получаем 5,8 ТБ, количество видеокамер 46 шт., длина одного видеофрагмента 1 мин., количество видеофрагментов за день 300 шт., срок хранения данных 7 дней.

Далее рассмотрим систему видеонаблюдения прилегающей местности к объекту. Классические камеры при необходимости размещаются на подъездных путях в пределах возможности подключения к сетевой инфраструктуре. Автономные же камеры могут быть размещены на значительном удалении от сетевой инфраструктуры, а также от стационарных источников электропитания. Необходимая глубина видеоархива для системы видеонаблюдения прилегающей местности из 29 автономных видеокамер составит 88 ГБ, срок хранения данных 7 дней, длина одного видеофрагмента 1 мин., количество видеофрагментов за день 200 шт. Для расчёта количества автономных видеокамер для размещения на большой площади можно использовать простые модели в системе MatLab и других.

В данной работе были проанализированы системы видеонаблюдения на основе классических и автономных видеокамер, применения систем автономных камер для обеспечения безопасности объектов и территорий объектов менее эффективно, чем применения систем классических камер, в то же время, применения систем автономных камер для задач обеспечения охраны периметра, а также прилегающей местности может быть более эффективным решением, чем применения систем видеонаблюдения на основе классических видеокамер для этих задач.

Литература

1. Морозостойкая 4K Smart-камера Hikvision DS-2CD4585F-IZH с моторизированным объективом. [Электронный ресурс]. – Режим доступа: <https://www.dssl.ru/products/ds-2cd4585f-izh/> (дата обращения: 03.03.2022).
2. Беспроводная 4G-LTE камера безопасности. [Электронный ресурс]. – Режим доступа: <https://reolink.com/product/reolink-go/> (дата обращения: 03.03.2022).
3. DS-2CD4565F-IZH. [Электронный ресурс]. – Режим доступа: https://hikvision.ru/product/ds_2cd4565f_izh (дата обращения: 03.03.2022).
4. Как рассчитать объем дискового пространства для видеоархива? [Электронный ресурс]. – Режим доступа: <https://www.umx.ru/about-company/faq/videonablidenie/rasschitat-diskovogo-prostranstva-videoarkhiv/> (дата обращения: 03.03.2022).
5. BolyGuard BG310-M 4G. [Электронный ресурс]. – Режим доступа: https://bolyguard.ru/bolyguard_310m/ (дата обращения: 03.03.2022).

УДК 628.9

СВЕТОВЫЕ ХАРАКТЕРИСТИКИ КАК ИНСТРУМЕНТ СОЗДАНИЯ ВИЗУАЛЬНЫХ АКЦЕНТОВ В ПРОСТРАНСТВЕ ГОРОДА В ТЕМНОЕ ВРЕМЯ СУТОК

Захарова Ю.Е.¹ (студент)

Научный руководитель – канд. архитектуры Соколова М.А.¹

¹*Университет ИТМО*

e-mail: j.zakharova@inbox.ru

В работе рассмотрены светотехнические характеристики акцентного освещения вечернего города, выявлены критерии оценки влияния световых акцентов на повышение эффективности ориентирования человека в пространстве.

Ключевые слова: световой дизайн, визуальный акцент, световой акцент, доминанта, ориентир, контраст, городское пространство, светопространство, информационная эффективность.

Основное назначение световой архитектуры – организация светопространственной среды города. В темное время суток искусственный свет «создает видимый мир», а ночной город становится эмоциональной средой, которая отличается от образа дневного города и может восприниматься человеком иначе.

В современной практике светового дизайна отмечается тенденция к избыточному освещению главных улиц города, что приводит к нивелированию роли пространственных доминант. Также отмечается совместное использование цвета и света вместе с движущимися изображениями, растет интерес к динамическому освещению. Визуальный хаос приводит к серьезной проблеме – неспособности различить важное.

Многие исследователи отмечают, что световая среда должна обладать информационной эффективностью, что подразумевает под собой возможность создания системы акцентов и ориентиров средствами искусственного света. Проблемы освещения и создания комфортной световой среды города, создание световых акцентов, возможности использования их в качестве помощи человеку в ориентировании в пространстве представляют собой комплексный вопрос. Поэтому для раскрытия темы необходимо проанализировать подходы из разных профессиональных областей (архитектуры и урбанистики, светотехники и светового дизайна, социальных наук и видеоэкологии).

Целью работы стало создание типологии акцентов в освещении вечернего города в соотношении «фон – объект», выявление их количественных и качественных характеристик и критериев оценки влияния световых акцентов на повышение эффективности пространственного ориентирования человека.

Освещение городских пространств ориентировано на выполнение следующих функций: создание атмосферы города, коммуникация, социальная активность, архитектурное улучшение, распределение внимания и поиск пути. Визуальные акценты рассматриваются в качестве ориентиров в пространстве, и их ключевым свойством становится возможность сохраняться в памяти человека и способствовать комфортному передвижению человека в пространстве.

Выделяют три типа световых пространств [1]:

1. Светопространство движения – основная роль в восприятии передвижения по городу в пространстве и времени (городские коммуникации, пешеходные дороги).
 2. Светопространство общения – внутренняя динамика и значимость, имеет относительную дискретность и развитые связи с окружением (общественные центры, визуальная информация).
-

3. Светопространство отдыха – социальное и функциональное назначение освещения (локальные участки рекреационных зон, участки жилых зон).

Основные элементы среды, которые могут быть использованы в роли визуального акцента:

- Доминанты – господствующие компоненты, контрастно отличающиеся от своего окружения рядом параметров: размерами, формой, цветом.
- Акценты – элемент композиции, выделенный из общего ряда деталей объекта за счет особого решения их отдельных изобразительных характеристик.
- Ориентиры – запоминающиеся объекты городской среды; привлекающие внимание элементы городского пространства; объекты, композиционно формирующие среду; объекты, видимые с больших расстояний, видимые длительное время в процессе передвижения по среде [2].

Анализ качества световой среды можно осуществить по таким критериям, как комфорт, безопасность, эстетические свойства. Средняя яркость и освещенность, равномерность, контраст, спектральный состав отвечают за комфортность среды. Освещенность должна соответствовать конкретным пространственным ситуациям, но отсутствие или снижение данного показателя не всегда снижает качество жизни. Темнота может нести положительные эффекты, такие как уменьшение негативного воздействия света, снижение затрат на электроэнергию.

Безопасность среды может быть описана показателями дискомфорта, ослепленности, блескости. Данные показатели актуальны при восприятии световых пространств завышенной яркости или при попадании бликов неправильно размещенных осветительных устройств в поле зрения человека. Показатель дискомфорта служит критерием оценки дискомфортной блескости (UGR), а показатель ослепленности – критерием оценки слепящего действия. С точки зрения эстетических свойств рассматривается световая структура, исследуется организованность световой структуры фрагментов городской среды, построение иерархии видов освещения в соответствии с объемно-пространственной композицией; архитектурно-планировочным решением и масштабом восприятия [3]. Все показатели представлены в таблице 1.

Таблица 1. Светотехнические показатели световой среды

Комфорт	Средняя освещенность	$E_{ср}$ (Лк)
	Средняя яркость	$L_{ср}$ (кд/м ²)
	Контрастность	$L_{мин}/L_{макс}$
	Равномерность	$L_{мин}/L_{ср}$
	Полуцилиндрическая освещенность	$E_{пл}$ (Лк)
	Спектральный состав	$T_{цв}, Ra$
Безопасность	Ослепленность	
	Дискомфорт	UGR
	Блескость	

Визуальная заметность (привлекательность) — это перцептивное качество, благодаря которому некоторые объекты в сцене выделяются и привлекают внимание. Традиционно в городском освещении увеличение яркости было одним из основных инструментов, используемых для улучшения заметности объекта.

Согласно нормам наружного архитектурного освещения городских объектов сооружений зданий, монументов центральных площадей, зон общегородских доминант, улиц, парков, пешеходных улиц общегородского значения заливающее освещение и локальное освещение составляет $L_{ф} - 5 - 10$ кд/м². При этом средняя яркость акцентированного светом элемента $L_{э} - 15 - 30$ кд/м².

Памятники, монументы, сооружения, элементы ландшафта на территориях окружного и районного значения имеют показатели ниже: заливающее и локальное освещение $L_{ф} 3 - 7$ кд/м², $L_{э} 10 - 20$ кд/м².

Памятники, монументы, сооружения, элементы ландшафта на территориях местного значения имеют следующие показатели: заливающее и локальное освещение $L_{ф} - 3 - 5$ кд/м², $L_{э} - 8 - 10$ кд/м².

При равномерном заливающим освещении фасада отношение максимальной освещенности к минимальной должно быть не более 3:1, а на рельефных и многоцветных фасадах – до 5:1, а максимальная освещенность должна создаваться на основных композиционно-пластических элементах.

При неравномерном заливающим освещении фасада соотношение максимальной и минимальной освещенности в пределах освещаемой зоны принимается не менее 10:1 и не более 30:1, при этом максимальная освещенность создается на акцентуемом светом элементе [4].

Однако, при неточном планировании и недостаточном контроле, тенденции завышения яркости объектов способствуют формированию визуального «хаоса» и светового загрязнения. Существующие нормативные документы определяют соотношение яркости между объектом и его фоном в зависимости от требуемой степени заметности (привлекательности) при этом более высокий контраст яркости рекомендуется для более высокой степени заметности. Эти рекомендации обычно интерпретируются как основанные на заметности объекта на однородном фоне. В реальных городских условиях объекты, как правило, имеют сложный фон. Заметность объекта зависит от свойств контекста, в котором находится объект в дополнение к свойствам объекта, таким как локальный контраст признаков (например, цветовой контраст и контраст яркости) [5]. В архитектурном освещении активно используются спектральные характеристики. Элементы «цветного» освещения создают визуальное лидерство. Их относят к категории более заметных и интересных. Светоцветовые акценты можно описать с помощью характеристик цвета (табл.2).

Таблица 2. Параметры цвета (Щепетков Н.И.) [3]

Параметры цвета	Объективные	Субъективные
Качественные	Спектральный состав длина волны излучения λ (нм)	Цветовой тон (красный, синий и т.д.)
	Чистота P%	Насыщенность H
Количественные	Яркость L (кд/м ²)	Светлота V

Стоит обратить внимание, что каждый объективный параметр имеет «аналог» субъективный, что говорит о важности учета восприятия цвета человеком. Любой акцент формируется на основании контраста. Иоганнес Иттен предложил следующую типологию цветовых контрастов:

1. Контраст по цвету.
2. Контраст светлого и темного.
3. Контраст холодного и теплого.
4. Контраст дополнительных цветов.
5. Симультанный контраст, основанный на физиологической особенности зрения придавать рядом находящемуся цвету оттенок дополнительного.
6. Контраст по насыщенности, заключающийся в противопоставлении насыщенных и блеклых цветов.
7. Контраст по площади цветового пятна, возникающий при сопоставлении цветовых поверхностей разной величины.

Выявление влияния факторов дизайна фонового освещения на внешний вид городских объектов говорит о том, что значимые факторы освещения не ограничены характеристиками, представленными в литературе по светотехнике, такими как яркость и цветовой контраст, равномерность или неравномерность фонового освещения.

Для технического описания акцента и влияния его параметров на восприятие среды человеком необходимо комплексное рассмотрение показателей: среды (фона) и светового эффекта (акцента, объекта), источника света или системы источников для создания каждого светового решения. Описание системы освещения (источников света) необходимо для оценки энергоэффективности светового решения.

Различимость любого объекта зависит от яркостных характеристик, углового размера, контраста между объектом и фоном, спектра освещения, прозрачности воздуха и продолжительности наблюдения.

В вечернее время внимание полностью сосредоточено на световых эффектах, а освещенные точки выступают в роли ориентиров. Светлые и темные поверхности создают впечатления «районов», роль главных путей выполняют освещенные коридоры. Распознавание узлов и границ зависит от яркости освещения. Наличие ночного освещения повышает видимость дорожной ситуации, повышает личную безопасность, снижает уровень вандализма и делает возможным ориентацию в пространстве.

Существующие исследования подтверждают функциональную роль наличия визуальных акцентов (ориентиров) в запоминании маршрута и поиске пути. Запоминающейся средой делают структурные особенности и опыт каждого человека, полученный в процессе пространственного ориентирования. Таким образом, наиболее эффективный способ помочь людям, передвигающимся в незнакомой среде, заключается в создании модели, которая содержит визуальные ориентиры. Люди успешно определяют ориентиры, содержащие примечательные особенности. Заметность, отличительность и привлекательность (с точки зрения контраста между объектом и фоном) являются наиболее часто упоминаемыми характеристиками объектов или зданий, которые должны выполнять функцию ориентиров. Эти характеристики имеют решающее значение для навигации.

Богатство среды ориентирами влияет на ее мысленное представление, на возможность построения общей когнитивной карты навигационной среды с помощью преобразования отношений между маршрутами и окружающими объектами. Таким образом, визуальные акценты вносят значительный вклад в построение конфигурационных знаний о среде.

В данной статье рассмотрено использование акцентного освещения в качестве инструмента способствующего ориентации в пространстве города в темное время суток. Главной задачей создания визуального акцента является привлечение внимания к объекту, его «оживление» с помощью освещения, посредством дифференциации яркостей, спектральных характеристик, особенностей светового «рисунка». По аналогии с театральным светом, данный инструмент призван завершить образ путем добавления тонких нюансов освещения, выделения областей, на которых должно быть сосредоточено особое внимание зрителя. С помощью света есть возможность воздействовать на эмоции человека, придавать объектам дополнительные смыслы, таким образом, свет становится средством коммуникации человека с городской средой, позволяет сделать акцент на архитектуре и влиять на человека, неся в себе социокультурное и философское содержание.

Объемно-пространственная система, структурные связи ландшафтных, архитектурных и цветоцветовых доминант создают уникальную вертикальную и пространственную композицию фрагментов города. Использование акцентного освещения может стать тем самым балансом между функциональностью и визуальным комфортом, поможет создать спланированное освещение среды и создания понятной структуры. В данном случае немаловажным преимуществом использования данного принципа является уменьшение установленной мощности световой системы, сокращение стоимости системы освещения и затрат на ее обслуживание, уменьшение негативного влияния света на природную среду.

Литература

1. Щепетков Н.И. Световой дизайн города // Архитектура-С. – 2006. – 319 с.
2. Робежник Л.В. Ориентиры в современной городской среде: Архитектурно-дизайнерская трактовка и философское осмысление // Праксема. Проблемы визуальной семиотики. – 2018. – №3. – С. 114–127.
3. Быстрянцева Н.В. Комплексный подход в создании световой среды вечернего города // М.: МАРХИ. – 2015. – 21 с.
4. СП 52.13330.2016 Естественное и искусственное освещение. Дата введения 2017-05-08. – 135 с.
5. Davoudian N. Visual saliency of urban objects at night: Impact of the density of background light patterns // Leukos. – 2011. – Т.8. – №2. – С. 137–152.

УДК 621.315.592, 544.032.65

ВЛИЯНИЕ ЛЕГИРОВАНИЯ ПЕРЕХОДНЫМИ МЕТАЛЛАМИ НА ФОТОПОТЕМНЕНИЕ В ХАЛЬКОГЕНИДНЫХ ПЛЁНКАХ

Капустина Е.В.¹ (студент), Гресько В.Р.¹ (аспирант)

Научный руководитель – канд. техн. наук Сергей М.М.¹

¹*Университет ИТМО*

e-mail: kapustina01@inbox.ru, gresko.97@mail.ru, maxim.m.sergeev@gmail.com

В настоящей статье показана зависимость светочувствительности пленки халькогенидного стекла в зависимости от концентрации легированного в пленку металла. Были исследованы пленки семейства AsS, которые востребованы в качестве фоточувствительных элементов в оптоэлектронике. Показано, что пятиминутное лазерное воздействие плотностью мощности 20 Вт/см² на халькогенидную пленку без металла приводит к значительному (в среднем до 14%) фотопотемнению пленки. С повышением концентрации легированного металла процесс фотопотемнения происходит медленнее. Результаты исследования могут быть полезны для выявления устойчивых связей между режимами лазерного воздействия и модификации оптических свойств материала, в том числе, показателя преломления и экстинкции.

Ключевые слова: Халькогенидные плёнки, фотопотемнение, непрерывное лазерное излучение, оптические свойства.

Полупроводниковые халькогенидные плёнки, основанные на соединениях S, Se, Te, за счет своей высокой фоточувствительности широко используются в качестве сверхбыстрых оптических переключателей, преобразователей частоты, оптических усилителей, оптических записывающих устройств и т. д. [1, 2].

Особенность халькогенидных стекол заключается в проявлении фотоиндуцированных явлений, например, обратимых фотоструктурных изменений, фотоиндуцированной анизотропии и фотоиндуцированного объемного расширения [3, 4]. Обратимое фотоструктурное изменение проявляется как фотопотемнение [5], когда оптическая ширина запрещенной зоны уменьшается при потемнении. Исходные параметры могут быть восстановлены последующим отжигом при температурах, близких к температуре стеклования.

Строение халькогенидов уникально тем, что только два из их четырех р-электронов используются для связи, оставляя пару несвязывающихся р-электронов (одиноких электронных пар). Наиболее заполненная электронная зона, валентная зона халькогенидных материалов, образуется в результате перекрытия этих орбиталей одиноких пар. Так, взаимодействия одиноких пар играют важную роль в структурных и оптических свойствах этих материалов [6].

Для модификации оптических свойств бинарные халькогениды также могут быть легированы металлами, например молибден (Mo) или серебро (Ag). Поскольку применение халькогенидов во многом основано на их взаимодействии с электромагнитным излучением, важной задачей оказывается исследование влияния легирующих элементов на фоточувствительность материала.

В частности, в данной работе исследуется кинетика фотопотемнения As₃₀S₇₀ плёнок, легированных различным содержанием Mo при воздействии непрерывного лазерного излучения. На данный момент проведено множество исследований воздействия лазерного излучения на халькогенидные пленки, например фемтосекундная лазерная запись на стёклах Ge-As-Se [7], а также изучены свойства халькогенидных пленок, легированных металлом, в том числе молекулярные модели ковалентного присоединения кластеров халькогенидов,

легированных переходным металлом [8]. Однако, влияние лазерного воздействия на халькогенидные пленки, легированные металлом – малоизученный вопрос.

Целью настоящего исследования является определение основных закономерностей кинетики фотопотемнения халькогенидных плёнок, легированных молибденом, в зависимости от параметров воздействия и концентрации легирующего компонента. Образцами исследования были пять халькогенидных плёнок толщиной 500 нм сульфида мышьяка $As_{30}S_{70}$ со следующими концентрациями молибдена: 1/5, 1/2, 2/1, 5/1 и плёнка без примесей.

Пленки, легированные молибденом, были получены следующим образом: осколки объемного стекла $As_{40}S_{60}$, синтезированного методом закалки в расплаве, и порошка $(NH_4)_2MoS_4$ использовались в качестве прекурсоров для образования исходных растворов. Растворы тщательно перемешивали в соотношении 5:1, 2:1, 1:1, 1:2 и 1:5 и оставляли для гомогенизации и реакции друг с другом в течение 24 часов. Тонкие слои были нанесены методом центрифугирования. Подготовленные тонкие пленки помещали в вакуумную печь, предварительно нагретую до $200^\circ C$, продували чистым газообразным аргоном и быстро вакуумировали до остаточного давления 1 Па. Слои подвергали отжигу в течение 2 ч, а затем охлаждали до комнатной температуры.

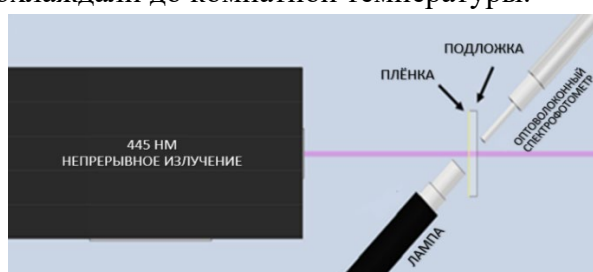


Рис. 1. Схема экспериментальной установки

Для выявления зависимости светочувствительности пленки от концентрации металла была использована экспериментальная схема, представленная на рисунке 1. В качестве лазерного источника был выбран непрерывный лазерный источник на длине волны 445 нм.

На одной линии с источником света за пленкой был расположен спектрофотометр AvaSpec ULS4096CL-EVO, так, чтобы снимались показания модифицированной области. Область модификации составила $4,5 \text{ мм} \times 0,6 \text{ мм}$.

В ходе эксперимента плотность излучения лазера варьировалась от 8 до 25 Вт/см^2 , спектрофотометр детектировал область взаимодействия с лазерным веществом каждую минуту от 0 до 5 минут.

На рисунке 2 показано, что в результате лазерного воздействия длительностью 5 минут происходит смещение спектральных линий в сторону снижения значений пропускания плёнок (фотопотемнение). При этом зарегистрированное изменение пропускания в среднем по видимой части спектра у пленок чистого сульфида мышьяка без примесей составляет 14%, против 4% у плёнки, содержащей примесь в соотношении 2/1. При этом образец с самым высоким содержанием молибдена (5/1) практически не чувствителен к воздействию (среднее смещение составило 0,02%).

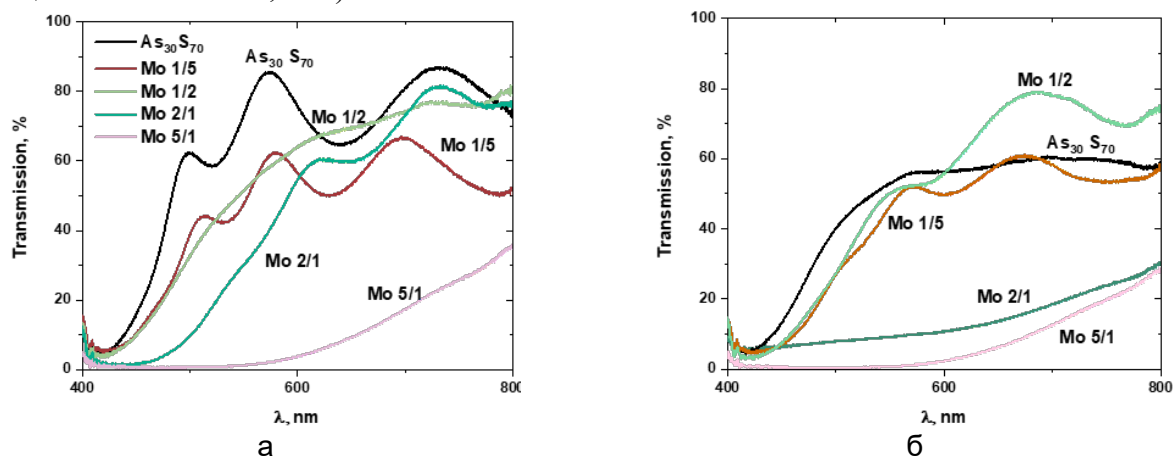


Рис. 2. Спектры: а) исходные всех образцов, б) после 5-мин обработки при плотности излучения 20 Вт/см^2

На рисунке 3 показано воздействие на пленки с концентрацией Mo 1/2 и Mo 2/1 тремя различными режимами плотности излучения: 8, 14 и 20 Вт/см^2 . Для исходного образца без

примесей запрещенная зона находится примерно на 400 нм, для образца с наибольшим содержанием молибдена (5/1) – примерно на 600 нм.

Увеличение концентрации легированного металла не приводит к монотонному уменьшению фоточувствительности пленки. На рисунке также заметно, что пленка с концентрацией 2/1 более чувствительна к лазерному воздействию, чем пленка с концентрацией 1/2.

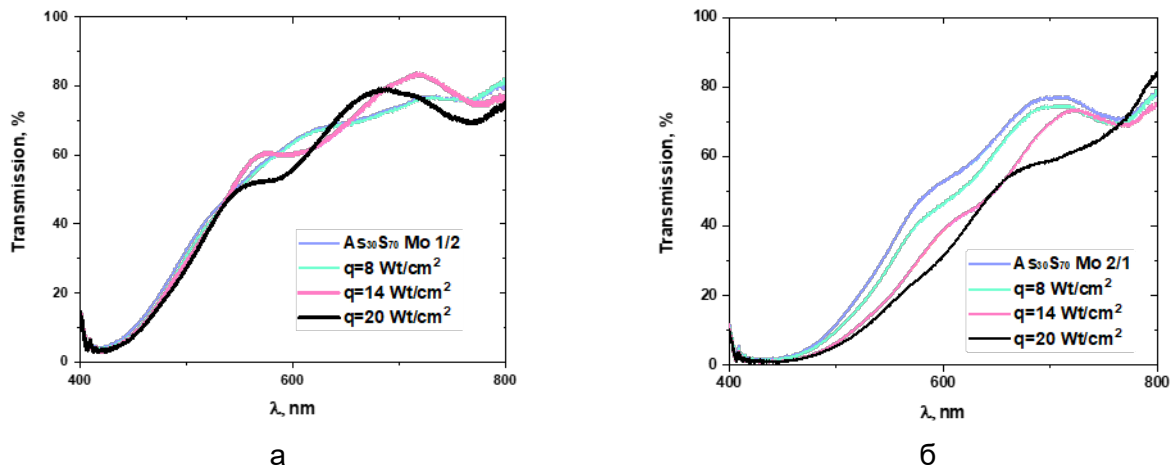


Рис. 3. Спектры пропускания при воздействии на пленки в течение 5 мин с разной плотностью излучения а) $As_{30}S_{70}$: Mo $\frac{1}{2}$ плёнка, б) $As_{30}S_{70}$: Mo $\frac{2}{1}$ плёнка

Для количественного анализа кинетики фотопотемнения были построены зависимости величины изменения пропускания $dT = T_0 - T_i$ от времени экспонирования, где T_0 это значение исходного пропускания пленки на длине волны 500 нм, T_i – значение пропускания на длине волны 500 нм в момент времени t_i . Данная длина волны близка к значению толщины образцов, в связи с чем влияние интерференционных эффектов в тонких пленках было минимизировано.

На рисунке 4 представлена зависимость dT для плёнок $As_{30}S_{70}$ и $As_{30}S_{70}$:Mo1/2 при различных q . Подтверждено, что повышение интенсивности модифицирующего излучения приводит к кратному росту значения dT . При этом наибольшее изменение спектра пропускания происходило в течение первой минуты.

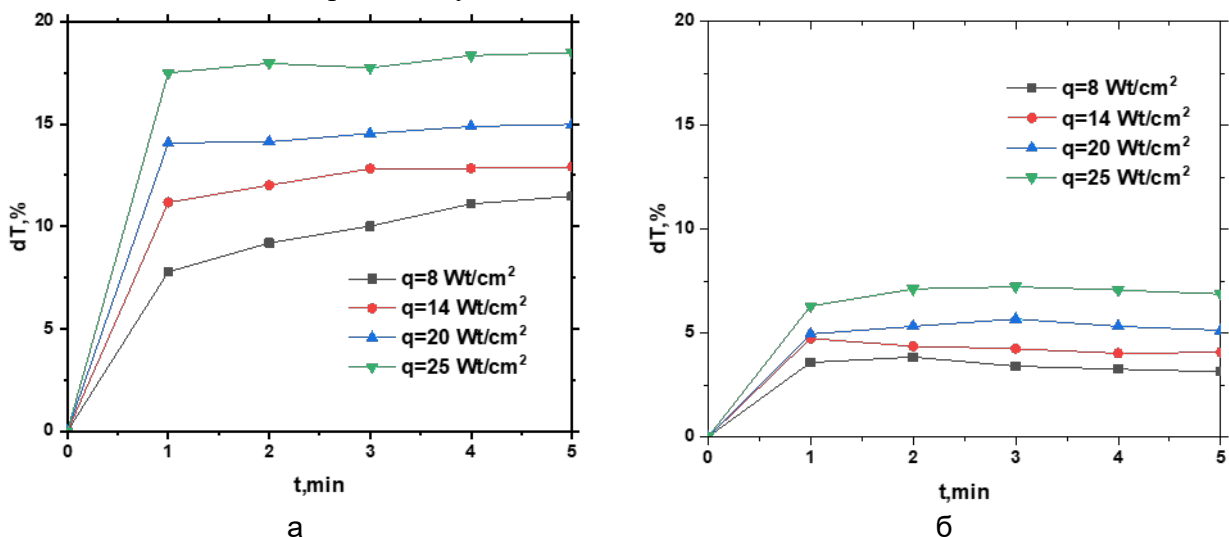


Рис. 4. Кинетика изменения dT при лазерном воздействии. а) $As_{30}S_{70}$ плёнка, б) $As_{30}S_{70}$: Mo1/2 плёнка

На рисунке 5 представлены зависимости dT при плотности мощности 8 Вт/см² и 14 Вт/см² для плёнок с различной концентрацией Mo. Показанные кривые дополнительно демонстрируют, что при наибольшей концентрации металла dT была близка к нулю.

Как было замечено ранее, для халькогенидов характерно наличие одиноких пар электронов. При легировании металлом одинокие пары расходуются для образования связей с переходным металлом, тем самым фотоиндуцированные свойства уменьшаются.

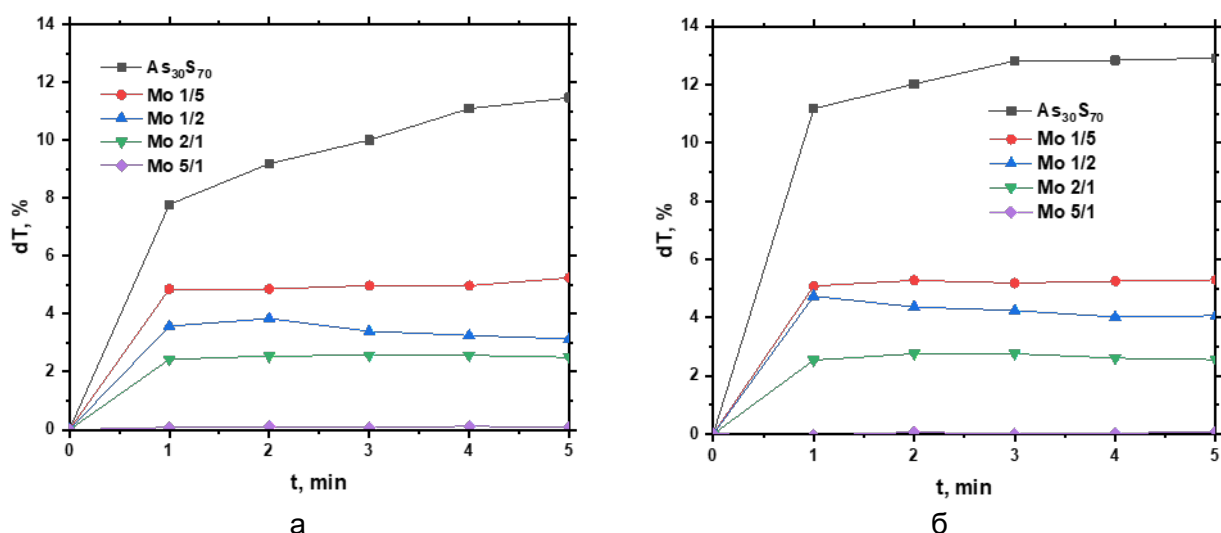


Рис. 5. Кинетика изменения dT при лазерном воздействии а) 8 Вт/см^2 , б) 14 Вт/см^2

Одним из наиболее интересных применений лазерного воздействия на такие плёнки является запись оптической информации, где модифицированная область несет в себе информацию по интенсивности прошедшего поляризованного излучения. Кодирование информации по амплитуде осуществляется за счет режима фотопотемнения, фотопросветления плёнки, а также за счет анизотропных свойств модифицированной области.

В статье исследована кинетика фотопотемнения плёнок в течение 5 мин под действием непрерывного излучения с длинами волн 445 нм. Показано, что в течение 1 мин способность к пропусканию образцов значительно увеличивалась, после чего её изменение замедлялось.

Подтверждено что изменение содержания молибдена в составе ведет к изменению чувствительности пленки к лазерному воздействию. Меньшая ширина запрещенной зоны для такой структуры объясняет эффект фотопотемнения. Легирование халькогенидного стекла переходным металлом приводит к уменьшению концентрации одиноких парных электронов и, следовательно, к исчезновению фотоиндуцированных эффектов.

Эксперимент доказал что процесса фотопотемнения проще добиться лазерной обработкой пленки без переходного металла или с его низким содержанием. Это также связано с уменьшением количества одиноких электронных пар, которые задействуются для образования металлических связей и ухудшают фоточувствительные свойства типичного халькогенидного соединения.

Лазерное воздействие производилось при различных значениях плотностей мощности: 8, 14, 20 и 25 Вт/см^2 . Увеличение плотности мощности приводит к ускорению процесса фотопотемнения.

Дальнейшие исследования направлены на выявление устойчивых связей между режимами лазерного воздействия и модификации оптических свойств материала, в том числе, показателя преломления и экстинкции.

Благодарности

Авторы благодарят канд. техн. наук Колобова А.В., ИПХФ РАН за предоставление образцов плёнок и обсуждение результатов.

Литература

1. Нежданов А.В. Структурные и оптические свойства пленок халькогенидных стеклообразных полупроводников на основе S, Se и Te: диссерт. Канд. Техн. Наук. – Нижний Новгород. – 2019. – 162 с.
2. Непомнящих В.А., Подгорный К.А. Порождение правил поискового поведения динамической системой // IV-я Всероссийская научно-техническая конференция «Нейроинформатика-2002». Сб. научных трудов. – 2012. – Ч. 1. – С. 110–116.

3. Shimakawa K., Kolobov A., Elliott S. R. Photoinduced effects and metastability in amorphous semiconductors and insulators // *Advances in Physics*. – 1995. – Т.44. – №6. – С. 475–588.
4. Tanaka K. Photo-induced metastability in amorphous semiconductors // *Journal of non-crystalline solids*. – 1991. – Т.137. – С. 1–10.
5. Berkes J.S. SW I ng, Jr., WJ H i 11 ega s // *Appl. Phys.* – 1971. – Т.42. – С. 4908.
6. Fritzsche H. The origin of reversible and irreversible photostructural changes in chalcogenide glasses // *Philosophical Magazine B*. – 1993. – Т.68. – №4. – С. 561–572.
7. Zhu L. et al. Optical and thermal stability of Ge-As-Se chalcogenide glasses for femtosecond laser writing // *Optical Materials*. – 2018. – Т.85. – С. 220–225.
8. Dornsiepen E. et al. [$\{(PhSn)_3SnS_6\} \{(MCp)_3S_4\}$] (M= W, Mo): Minimal Molecular Models of the Covalent Attachment of Metal Chalcogenide Clusters on Doped Transition Metal Dichalcogenide Layers // *Journal of the American Chemical Society*. – 2019. – Т.141. – №41. – С. 16494–16500.

УДК 004.932, 523.9

ИССЛЕДОВАНИЕ КОЛЕБАНИЙ В СОЛНЕЧНОЙ ПЛАЗМЕ ПУТЁМ ПОСТРОЕНИЯ ПРОСТРАНСТВЕННО- ВРЕМЕННЫХ ДИАГРАММ

Каракотов Р.Р.¹ (студент), Анфиногентов С.А.²

Научный руководитель – д-р физ.–мат. наук Кузнецов А.А.²

¹*Университет ИТМО*

²*Институт солнечно-земной физики СО РАН*

e-mail: karakrus97@mail.ru, a_kuzn@mail.iszf.irk.ru, anfinogentov@iszf.irk.ru

Данная работа посвящена аспектам обработки спутниковых данных вспышечных событий в атмосфере Солнца. В процессе работы был написан набор программ для построения и их последующей обработки пространственно-временных диаграмм по изображениям, полученным прибором Atmospheric Imaging Assembly (AIA) космического аппарата Solar Dynamics Observatory (SDO). Были проанализированы захваченные колебания в корональной петле, вызванные вспышкой класса M1.0 22.10.2013 г. и дана возможная интерпретация наблюдаемой эволюции температуры плазмы.

Ключевые слова: Солнце, солнечная корона, пространственно-временная диаграмма, колебания, выравнивание гистограммы, выравнивание гистограммы с адаптивным порогом контраста, алгоритм Кэнни.

Исследование солнечной короны дает возможность получить ответы на вопросы, связанные с фундаментальной физикой плазмы, в том числе таких, как магнитное пересоединение, взаимодействие волна-частица, плазменные неустойчивости, ускорение заряженных частиц. Солнечная активность (вспышки, корональные выбросы масс, потоки высокоэнергичных частиц) оказывает значительное влияние на космическую погоду и, как следствие, на космические аппараты и станции, – снижая их срок эксплуатации и нанося вред электронике. Изучение и прогноз солнечной активности невозможны без наблюдения солнечной атмосферы, для чего используются как наземные телескопы, так и космические обсерватории, ведущие наблюдения в диапазонах длин волн, поглощаемых земной атмосферой.

Данные, используемые в работе, получены прибором Atmospheric Imaging Assembly, установленным на космической обсерватории SDO [1]. Он позволяет проводить одновременные наблюдения всего солнечного диска и области $0,5R_{sun}$ над уровнем лимба в нескольких линиях ультрафиолетового (УФ) и крайнего ультрафиолетового (КУФ) диапазонов с высокими пространственным ($1,5''$) и временным (до 12 сек) разрешением для одной длины волны наблюдений. Наблюдения при помощи AIA позволяют регистрировать излучение плазмы в диапазоне температур от 0,06 до 20 МК. Это обеспечивается четырьмя двухсотмиллиметровыми двухканальными телескопами с полем зрения $41'$, десятью УФ и КУФ фильтрами, приёмными ПЗС-матрицами 4096×4096 с разрешением $0,6''$ на пиксель.

Наличие верхнего предела динамического диапазона матрицы приводит к тому, что матрица уходит в насыщение во время вспышек M и X классов. Для частичного устранения этого эффекта используется переменное время экспозиции, которое уменьшается с ростом потока, что усложняет исследование плазменных процессов.

Колебания и волны в корональной петле можно определить при помощи построения пространственно-временной диаграммы. Пространственно-временная диаграмма –

распределение яркости по времени и пространственной координате, построенная вдоль виртуальной щели. Для исследования изгибных (кинк) колебаний виртуальная щель располагается поперек петли, а для изучения продольных волновых движений – вдоль нее.

Исследование солнечной короны проводилось на длине волны 131\AA , на которой хорошо визуализируются горячие корональные структуры (петли) с температурой порядка 10^7 K [1]. Поиск активной области проводился путём визуального анализа изображений солнечной короны в программе JHelioviewer [2]. Более подробная информация о наблюдении содержится в таблице 1.

Таблица 1. Информация о наблюдениях солнечной короны при помощи SDO/AIA

Начало наблюдения	22-10-2013 00:00:00
Конец наблюдения	22-10-2013 02:00:00
Длительность, с	7200
Количество кадров	600
Шаг между кадрами, с	12

Построение пространственно-временной диаграммы было реализовано в виде программы на языке программирования Python. Пошаговое выполнение поставленной задачи сделано следующим образом:

1. Загрузка данных.
2. Выбор единичного кадра.
3. Ввод координат точек на визуально различимой петле и определение параметров эллипса, на дуге которого они лежат.
4. Вычисление координат точек, равномерно распределенных на дуге полученного ранее эллипса (количество точек определяется поставленной задачей).
5. Вычисление координат точек, лежащих на нормалях к точкам дуге эллипса.
6. Масштабирование изображения с последующей интерполяцией.
7. Построение двумерной матрицы интенсивностей по полученным ранее координатам точек на единичном кадре.
8. Построение трехмерной массива значений интенсивности по всем кадрам.
9. Суммирование по одной из пространственных координат в зависимости от поставленной задачи и получение итоговой пространственно-временной диаграммы.
10. Визуализация и при необходимости последующая обработка.

Из 600 кадров был выбран один, на котором петля была отчётливо видна. Результатом работы программы стала пространственно-временная диаграмма, изображенная на рисунке 1. Для улучшения качества картинки было применено линейное контрастирование уровней интенсивности $l_1 = 1500$ и $l_2 = 6500$ (рис. 2).

В результате динамический диапазон сузился, что позволило выделить характерные для захваченных продольных волн структуры на диаграмме (V- и W-подобные структуры). Интервалы времени, в которых наблюдались захваченные колебания, приведены в таблице 2.

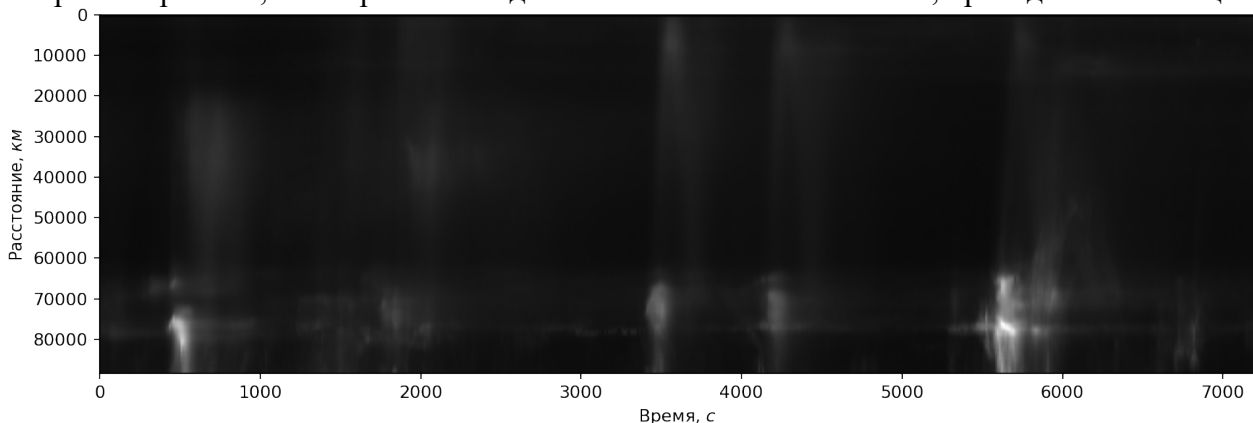


Рис. 1. Пространственно-временная диаграмма до обработки. Визуальное выделение колебаний возможно, но затруднено

Таблица 2. Временные интервалы, соответствующие захваченным колебаниям на диаграмме

Номер интервала	t_{starts} с	t_{ends} с
1	540	792
2	3444	3756
3	4140	4500
4	5604	5976

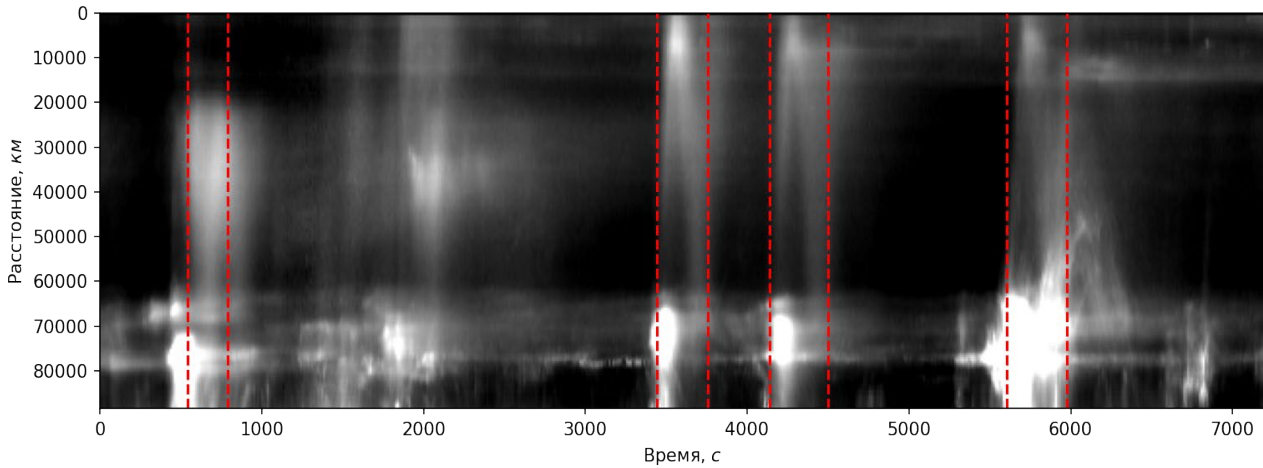


Рис. 2. Пространственная диаграмма после линейного контрастирования. Красным выделены визуально определенные границы наблюдаемых захваченных колебаний

Рассмотрим первый временной интервал. На диаграмме видно, что в данный временной промежуток в петле длиной 75,5 Мм наблюдалось колебание волны с периодом 252 с. Волна движется с трубочной скоростью, которая в условиях солнечной короны становится скоростью звука c_s :

$$c_s = \sqrt{\frac{\gamma k_B T}{\widehat{\mu}_p m_p}},$$

где $\gamma = \frac{5}{3}$, k – постоянная Больцмана,

T – температура плазмы в K ,

$\widehat{\mu}_p = 0.6$ – средняя масса частицы в корональной плазме, измеренная в массах протона,

m_p – масса протона.

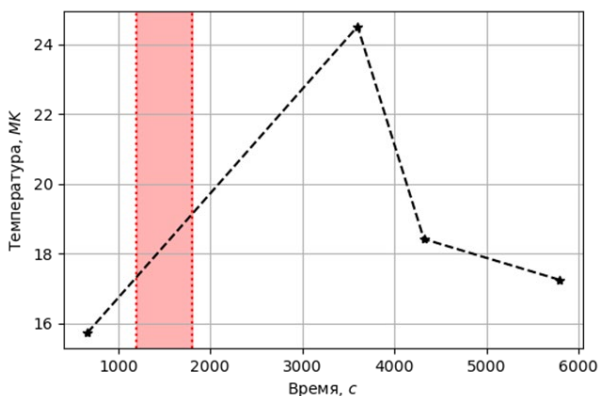


Рис. 3. Эволюция во времени температуры плазмы в корональной петле. Красным прямоугольником обозначен временной интервал развития энерговыделения вспышки класса M1.0

Таким образом, было получено значение температуры плазмы $T = 15,7$ МК. Полученная температура близка к температуре возбуждения линии 131 Å (~16 МК).

Аналогичную процедуру провели для последующих участков. Результаты приведены в таблице 3. Был построен график (рис. 3) изменения температуры плазмы в исследуемой петле. Мы можем видеть, что изначально наблюдаемая плазма имела температуру, близкую к температуре возбуждения спектральной линии. Эту температуру можно считать фоновой. В определенный момент времени произошла вспышка (показана красным прямоугольником на рисунке 3), нагревшая плазму в петле, которая в дальнейшем остыла за счёт излучения.

Таблица 3. Параметры захваченных колебаний

Номер интервала	Период, с	Скорость, км/с	Температура плазмы, МК
1	252	600	15,724
2	312	750	24,521
3	360	650	18,417
4	372	629	17,248

Для повышения качества визуализации пространственно-временной диаграммы было проведено выравнивание гистограммы (рис. 4) [3]. Видно, что выравнивание гистограммы даёт более засвеченную картину, чем выравнивание гистограммы с адаптивным порогом контраста (CLANE) [4]. При этом структуры, характерные для захваченных колебаний, становятся хорошо различимы, что облегчает определение пространственных и временных границ отдельных событий. Поскольку существенной разницы в изображениях, полученных в ходе выравнивания интенсивности с использованием методов гистограммы и CLANE, обнаружено не было, мы решили для автоматизации процесса обработки изображений использовать метод выравнивания яркости гистограммы.

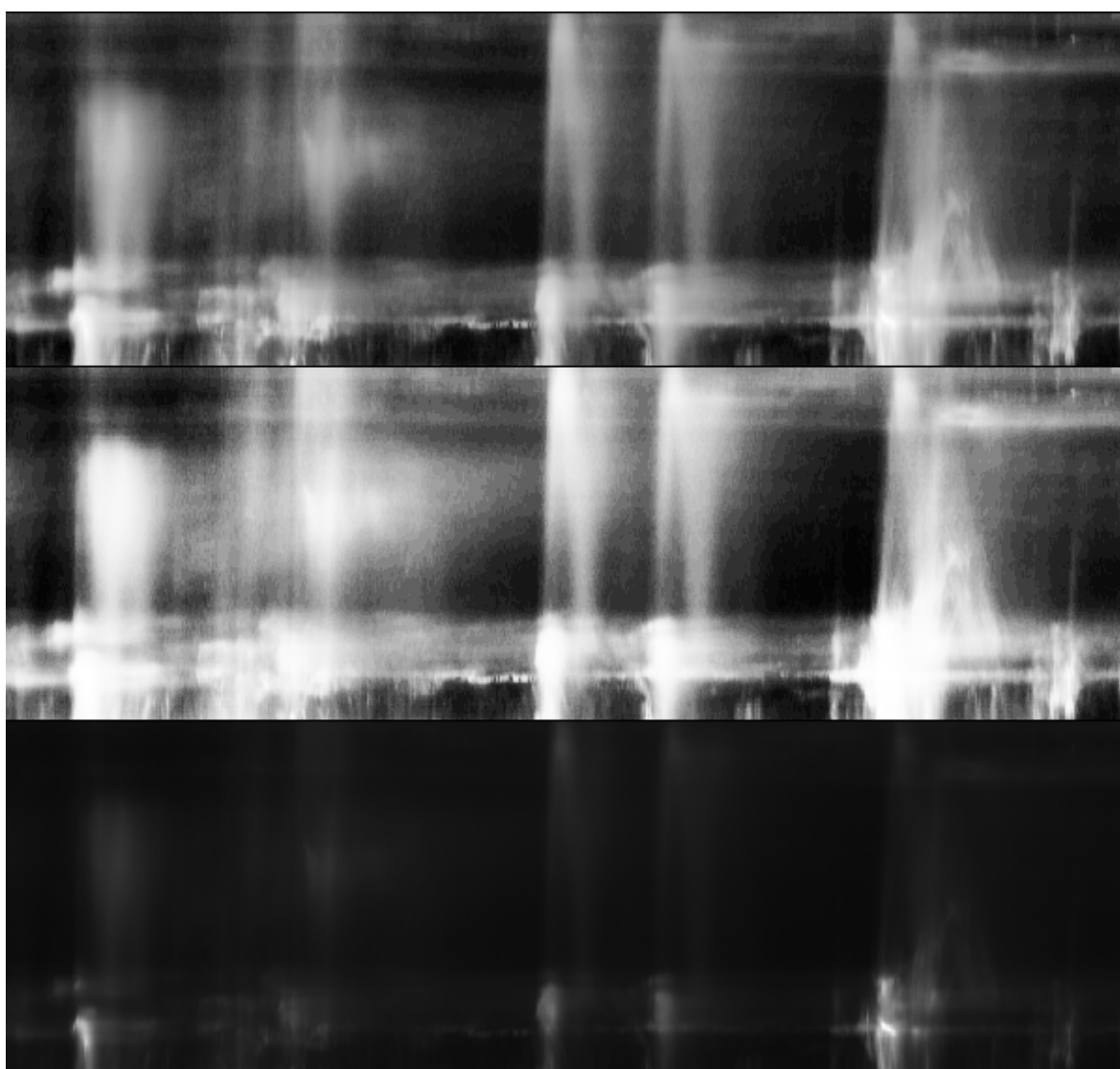


Рис. 4. Пространственно-временная диаграмма до обработки (снизу), после выравнивания яркости гистограммы (посередине) и после применения CLANE (сверху)

Для выделения петли были применены следующие преобразования:

- выравнивание гистограммы;
- CLANE;
- применение фильтра Кэнни [3] для выделения границ.

Для сравнения на рисунке 5 приведен результат работы преобразований и исходное изображение. Выделить петлю на исходном изображении сложнее, чем на изображении, полученном методом CLAHE. Видно, что использовать выравнивание яркости гистограммы нельзя, вследствие сильного уровня яркости фона по отношению к видимым петельным структурам. Поэтому при визуальном выделении корональных петель имеет смысл использовать метод CLAHE для выделения петель в короне.

В дальнейшем планируется снизить влияние исследователя путём автоматизации поиска сигнатур захваченных колебаний на пространственно-временной диаграмме при помощи техник машинного обучения.

Результатом проведенного исследования является разработка набора программного обеспечения, которое в будущем будет использовано для исследования солнечной короны с анализом короны активной области во время солнечной вспышки класса M1.0 и вычислением параметров захваченных колебаний.

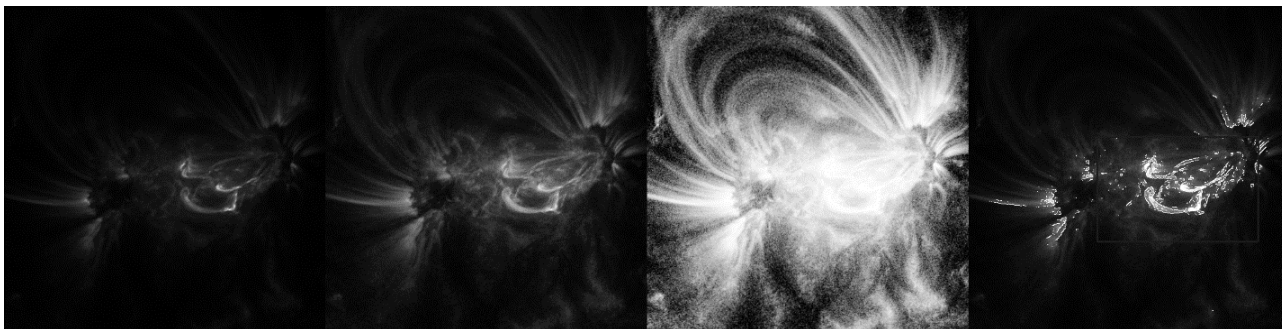


Рис. 5. Слева-направо: исходное изображение, изображение после CLAHE, изображение после эквализации гистограммы, изображение с контурами, выделенными фильтром Кэнни

Литература

1. Lemen J.R. et al. The Atmospheric Imaging Assembly (AIA) on the Solar Dynamics Observatory (SDO) // *Solar Physics*. – 2012. – Vol.275. – Pp. 17–40.
2. JHelioviewer. [Электронный ресурс]. – Режим доступа: <https://www.jhelioviewer.org/> (дата обращения: 10.10.2021).
3. Ежова К.В. Моделирование и обработка изображений. Учебное пособие. – СПб: НИУ ИТМО. – 2011. – 93с.
4. Эквализация гистограмм для повышения качества изображений. [Электронный ресурс]. – Режим доступа: <https://habr.com/ru/post/244507> (дата обращения: 03.12.2021).

УДК 528.8.044.6

ОЦЕНКА ЭНЕРГЕТИЧЕСКИХ ХАРАКТЕРИСТИК ЛАЗЕРНОГО СКАНИРУЮЩЕГО ДАЛЬНОМЕРА ДЛЯ БЕСПИЛОТНОГО ЛЕТАТЕЛЬНОГО АППАРАТА

Карзов М.М.¹ (студент), Пименов А.Ю.², Никандров Г.В.²

Научный руководитель – канд. техн. наук, доцент Маргарянц Н.Б.¹

¹Университет ИТМО

²АО «КТ-Беспилотные Системы»

*e-mail: karzovm1999@gmail.com, pimenovay@yandex.ru,
nikandrov.spb@mail.ru, nbmargaryants@itmo.ru*

В данной работе выполнен расчет минимальной мощности зондирования лазерного сканирующего дальномера, необходимой для проведения сканирования зоны посадки беспилотного вертолета. Рассмотрены существующие технические решения для задач лазерного зондирования и произведен подбор компонентной базы для передающего канала. На основе полученных данных сформированы основные требования к оптической системе устройства.

Ключевые слова: LiDAR, лидар, лазерное сканирование, мощность зондирования, посадка БЛА, подбор компонентной базы.

На сегодняшний день лазерные сканирующие системы нашли свое применение во многих гражданских и военных отраслях для решения задач локации, для сбора информации о подстилающей поверхности, построения рельефной карты местности. В частности, такие системы применяются для проведения посадки беспилотного летательного аппарата (БЛА) на неподготовленную площадку, когда заранее неизвестна форма её рельефа.

Принцип действия лазерного сканирующего дальномера (лидара) основан на измерении времени пролёта импульса лазерного излучения от устройства до отражающего объекта, а также получения набора углов отклонения лазерного луча для построения набора точек, формирующих трехмерное изображение пространства.

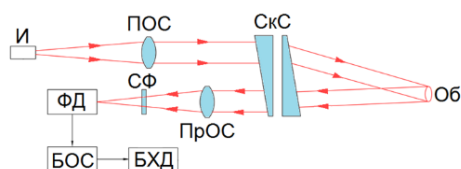


Рис. 1. Блок-схема лидара

Благодаря наличию систем навигации на борту БЛА, возможно построение трехмерной рельефной карты местности, что позволяет БЛА в автоматизированном режиме определять подходящие зоны для посадки [1, 2]. На рисунке 1 приведена блок-схема лидара, где И – источник излучения, ПОС – передающая оптическая система, СкС – сканирующая система, Об – отражающий объект, ПрОС – приемная оптическая система, СФ – светофильтр, ФД – фотодетектор, БОС – блок обработки сигнала, БХД – блок хранения данных.

Излучение от источника излучения (обычно лазер ближнего ИК диапазона) коллимируется с помощью передающей оптической системы. Сканирующая система отклоняет пучок излучения на заданный угол и направляет его на отражающий объект. Приемная оптическая система фокусирует отраженное излучение на фотодетекторе. Блок обработки сигнала преобразует полученный сигнал в массив данных с информацией о дистанции до отражающего объекта и его координатах, сохраняемый в блоке хранения данных.

Целью данной работы является анализ энергетических характеристик и определение основных требования к оптико-электронной системе лидара, предназначенного для построения

трехмерной карты местности предполагаемой зоны посадки беспилотного вертолета. С учетом особенностей данной задачи такая система должна обеспечивать угловое поле сканирования $90^\circ \times 90^\circ$. Диапазон дальности от 5 до 200 м, при этом точность определения расстояния — единиц сантиметров, а угла должна быть не хуже единиц угловых минут. Габариты устройства должны быть минимальными, а масса изделия — не более 1 кг.

На данный момент на рынке лидаров имеются готовые решения как для наблюдения при малых угловых полях, так и с возможностью сканирования на 360° в горизонтальной плоскости. В таблице 1 приведены характеристики имеющихся на сегодняшний день импульсных лидаров.

Таблица 1. Характеристики лидаров

Производитель	Модель	Диапазон сканирования, м	Габариты, мм
Phoenix LiDAR	SCOUT-16	1 – 100	160 × 116 × 116
Quanergy	M8-Ultra	до 200	103 × 89 × 83
Teledyne Tech.	CL-90	1,5 – 1000	300 × 213 × 209
Ibeo Automotive	ibeo LUX 4L	до 50	164 × 93,2 × 88
Velodyne	Alpha Prime	до 300	165 × 165 × 141
	Длина волны, нм	Угловое поле (Г × В), град	Угловое разрешение, град
Phoenix LiDAR	905	360 × 30	0,085
Quanergy	905	360 × 20	0,033
Teledyne Tech.	1550	90 × –	0,017
Ibeo Automotive	905	110 × 3,2	0,25
Velodyne	905	360 × 40	0,1

Обзор существующих решений показал, что имеющиеся решения не обеспечивают требуемого углового поля сканирования. Также при подвешивании системы с круговым сканированием на 360° в нижней полусфере БЛА будут присутствовать "мертвые зоны" при формировании набора координатных точек.

Расчет минимальной мощности зондирования $P_{пер}$ проводился по методике [3] с учетом специфики поставленной в данной работе задачи:

- поверхность цели является диффузно-отражающей по закону Ламберта;
- размеры отражающей поверхности много больше диаметра пятна излучения;
- потери в атмосфере распределены равномерно по всей траектории движения светового пучка, так как дистанция сканирования относительно мала (до 200 м).

При расчете учитывалось два типа шумов: шум фоновой засветки, создаваемый солнечным излучением, а также темновой шум, вызванный спонтанным возникновением носителей тока при отсутствии засветки фотоприемника. На рисунке 2 наглядно продемонстрировано влияние шума фоновой засветки.



Рис. 2. Влияние шума фоновой засветки: 1 — передающая оптическая система; 2 — приемная оптическая система

Минимальную регистрируемую фотодетектором энергия излучения определяется формулой:

$$E_{пр.м.} = \frac{c}{\omega} \cdot P_{итог} \cdot t_{имп}, \quad (1)$$

где $\frac{c}{\omega}$ – соотношение сигнал-шум (для достоверной регистрации сигнала его величина должна быть равна 6...10),

$P_{итог} = \sqrt{P_{фон}^2 + P_{темн}^2}$ – результирующая эквивалентная мощность шума фоновой засветки $P_{фон}$ (зависящая от фокусного расстояния приемной оптической системы f' и размера фотодетектора h) и темнового шума $P_{темн}$,
 $t_{имп}$ – длительность импульса.

Минимальная энергия и мощность источника излучения лидара составят:

$$E_{\text{пер.}} = \frac{E_{\text{пр.м.}} \cdot 4 \cdot L_{\text{макс.}}^2 \cdot e^{2 \cdot \alpha \cdot L_{\text{макс.}}}}{\zeta \cdot D_{\text{пр.}}^2 \cdot \tau_1 \cdot \tau_2 \cdot \rho}, \quad (2)$$

$$P_{\text{пер.}} = \frac{E_{\text{пер.}}}{t_{\text{имп.}}}, \quad (3)$$

где $L_{\text{макс}}$ – максимальная дальность до отражающего объекта,
 α – коэффициент потерь, определяемый метеорологической дальностью видимости (МДВ),
 $D_{\text{пр}}$ – диаметр входного зрачка приемной оптической системы,
 τ_1, τ_2 – коэффициент пропускания передающей и приемной оптической системы,
 ρ – коэффициент отражения объекта.

Определим величину минимальной мощности зондирования в зависимости от фокусного расстояния для ряда относительных отверстий приемного канала ($D/f' = 1:1; 1:2; 1:5$).

Расчет проводился для следующих параметров:

- МДВ = 10 км;
- длина волны источника излучения $\lambda = 0,905$ мкм;
- максимальная дальность сканирования $L_{\text{макс.}} = 200$ м;
- коэффициент отражения объекта (подстилающей поверхности) $\rho = 0,5$;
- коэффициенты пропускания передающей и приемной оптических систем $\tau_1 = \tau_2 = 0,9$;
- длительность импульса $t_{\text{имп.}} = 100$ нс;
- пороговое отношение сигнал-шум $\frac{С}{Ш} = 10$;
- диаметр фотодетектора 0,5 мм.

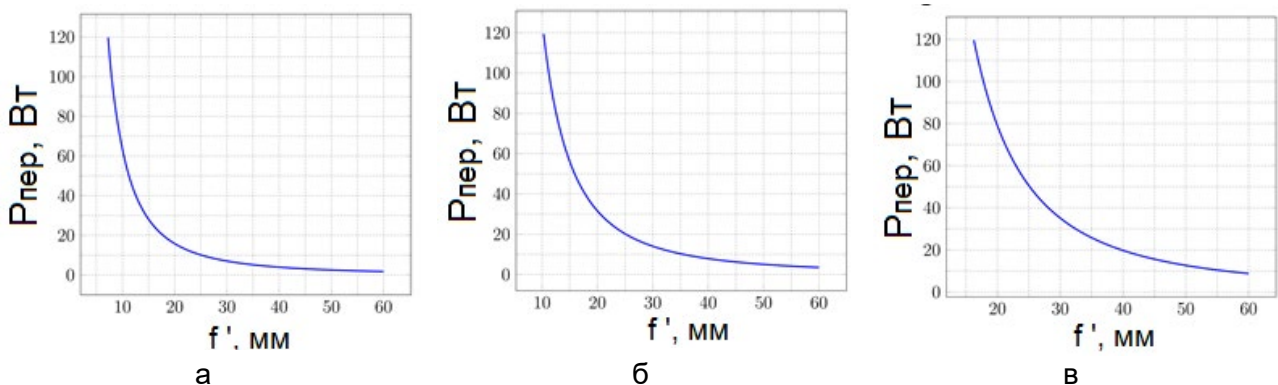


Рис. 3. Зависимость минимальной мощности зондирования от фокусного расстояния при а) $D/f' = 1:1$; б) $D/f' = 1:2$; в) $D/f' = 1:5$

Из рисунка 3 видно, что необходимая минимальная мощность сканирования уменьшается с увеличением выбранного фокусного расстояния. Мощность импульса излучения современных лазерных диодов находится в диапазоне от 20 до 120 Вт, следовательно, имеет смысл рассматривать оптические системы с фокусным расстоянием в диапазоне $f' = 20 - 60$ мм, которые позволят разработать систему с габаритами, не превышающими габариты аналогов. Для больших значений f' габариты системы будут выходить за рамки допустимых.

При выборе учитывались следующие параметры лазеров: размер излучаемой области, мощность импульса и расходимость излучения. В таблице приведены сравнительные характеристики лазерных источников для ближней ИК области спектра (около 900 нм). Выбор спектрального диапазона обусловлен большим количеством представленных моделей, а также упрощением процесса юстировки разрабатываемой системы за счет применения камер на основе кремниевого фотоприемника, при наблюдении излучения лазера.

Таблица 2. Характеристики лазерных диодов

Производитель, Модель	Мощность импульса, Вт	Длина волны, нм	Размер излучающей площадки, мкм × мкм	Расходимость вдоль OX × OY, °
Hamamatsu, L11348-307-05	21	870	70 × 10	24 × 11
Hamamatsu, L11649-120-04	22	870	230 × 10	30 × 10
Hamamatsu, L11854-323-51	75	905	230 × 10	24 × 10
Laser Components, 905D1S3J06X	52	905	160 × 10	20 × 12
Excelitas, TPGAD1S03H	20	905	76 × 10	25 × 10

Из таблицы видно, что наибольшая расходимость лазерного излучения диодов составляет около $\alpha_X = 24^\circ$, следовательно, относительное отверстие передающего канала должно быть не менее:

$$\frac{D}{f'} = \frac{1}{2 \operatorname{tg}(\alpha_X/2)} \approx 1:2,4. \quad (4)$$

Величина излучающей площадки лазера определяют расходимость лазерного излучения после передающего канала, и, следовательно, размер формируемого пятна d на объекте (фактически определяющему разрешение системы) при $f' = 40$ мм и $x = 100$ мкм:

$$d = \frac{L_{\max} \cdot x}{f'} = 0,5 \text{ м}, \quad (5)$$

где x – размер излучающей площадки.

С учетом вышеобозначенных критериев оптимальными являются лазеры Excelitas TPGAD1S03H или Hamamatsu L11348-307-05, удовлетворяющих условию необходимой мощности излучения.

На основе анализа энергетических характеристик системы были сформированы основные требования к оптической системе и компонентной базе лазерного сканирующего дальномера:

- фокусное расстояние оптической системы в диапазоне от 20 до 60 мм, относительное отверстие не менее 1/2,4;
- минимальная оптическая мощность импульса зондирования - не менее 5 – 10 Вт.

Таким образом, система может быть построена с использованием компактных полупроводниковых лазерных диодов.

Литература

1. Медведев Е.М., Данилин И.М., Мельников С.Р. Лазерная локация земли и леса: учеб. пособие. – М.: Геолидар: Геокосмос. – 2007. – С. 16–22.
2. Dong Pinliang, Chen Qi. LiDAR remote sensing and applications // CRC Press. – 2017. – 200 p.
3. Назаров В.Н., Балашов И.Ф. Энергетическая оценка импульсных лазерных дальномеров // СПб.: СПбГУ ИТМО. – 2002. – Т.38. – 32 с.

УДК 628.974

ПРИНЦИПЫ ФОРМИРОВАНИЯ ГОРОДСКИХ СОБЫТИЙНЫХ ПРОСТРАНСТВ СРЕДСТВАМИ ИСКУССТВЕННОГО ОСВЕЩЕНИЯ

Кутько Г.А.¹ (студент)

Научный руководитель – канд. архитектуры Соколова М.А.¹

¹Университет ИТМО

e-mail: kutkog@mail.ru

В работе рассмотрены проблемы организации событийных пространств средствами искусственного освещения, его влияния на восприятие пользователей и развитие городских территорий. Был проведен сбор данных и сравнительный анализ событийных пространств, организованных с помощью средств искусственного освещения в Санкт-Петербурге.

Ключевые слова: событийное пространство, искусственное освещение, визуальное восприятие, удержание внимания, социальное освещение, привлекательность территории.

В связи с глобальной урбанизацией, города ищут новые инструменты маркетинга для формирования событийных пространств, которые должны существенно отличаться от своих конкурентов и продвигать себя как уникальные продукты на мировом рынке. Сегодня необходимо создавать места с неповторимым характером с точки зрения экономики и культуры, делая их более привлекательными для горожан и туристов. Событийные пространства вынуждены позиционировать себя как узнаваемые городские бренды, привлекать молодые таланты, бизнес и инвесторов.

Одним из инструментов формирования досуговых вечерних мероприятий является творческое применение искусственного света. Все большее число городов начинает признавать важность искусственного освещения после наступления темноты не только с точки зрения безопасности, но и как средства создания нового имиджа и привлекательности территорий, развитию уникальных событийных пространств, которыми могут восхищаться не только коренные жители, но и туристы [1]. Это, в свою очередь, может помочь формированию у людей чувства принадлежности месту, конкретному городскому фрагменту. Использование искусственного освещения на мероприятиях привлекает внимание и значительно увеличивает количество посетителей, делая популярными места их проведения. Также это шанс привлечь людей в те районы города, которые не имеют какого-либо особого исторического наследия. Уникальное световое решение может также обеспечить дополнительную экономическую ценность в ночное время, такие проекты рассматриваются как значимый источник дохода и развития туризма на городском, региональном, национальном и международном уровнях.

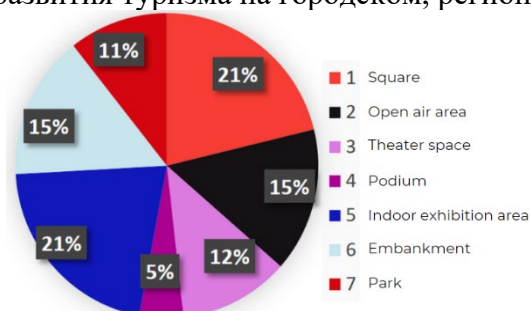


Рис. 1. Тип пространства

Это положительно сказалось на их посещаемости из-за возможности быть замеченными при обычной прогулке по городу.

За период 2017 – 2021 гг. автором статьи было рассмотрено 93 крупных события в Санкт-Петербурге, относящихся к информационным мероприятиям и ориентированным на организацию досуга. Рассматривались крытые и открытые общественные пространства. В результате проведенного исследования было выявлено, что 49 мероприятий проводились на открытом воздухе, на обширной территории (рис. 1).

Оформление всех событийных пространств достигло своих целей по привлечению посетителей, хотя мероприятия отличались по следующим параметрам:

1. Продолжительностью от 1 дня («Алые Паруса» - сценическое освещение, подсветка парусов на корабле, салют) до 5 лет (ожившие полотна в «Люмьер-Холле» - технология видеомэппинг).
2. Сезонностью – ежегодное (ярмарка на Манежной площади – наружные фигуры и инсталляции, декоративная подсветка), ежедневное (ожившие полотна в «Люмьер-Холле») или единичное мероприятие («Моя остановочка - диодная инсталляция»).
3. Масштабностью, где количество посетителей, включая туристов могло достигать 10 000 000 (звуковое шоу «Поющие мосты» - сочетание архитектурной подсветки моста и музыки).
4. Бюджетностью – от очень большого (Ярмарка на Манежной площади) до почти полного отсутствия финансирования (акция «Свеча памяти»).

Автором установлено, что время года не является важным аспектом, определяющим успех событийного пространства. Это можно проследить по фестивалю Wonder of Light, который был проведен в 2019 году, два раза, в весенний и осенний периоды, и в обоих случаях фестиваль посетило около 200 000 человек. Кроме этого замечено, что большинство досуговых мероприятий под открытым небом, где используются инструменты искусственного освещения, проходят осенью и зимой, это связано с длительностью светового дня в Санкт-Петербурге, а также попыткой компенсировать недостаток осенне-зимнего туризма.

Правильно организованные событийные пространства, могут сделать город привлекательным на более длительный период времени, в случае расширения временных рамок проведения мероприятий, независимо от того, в какое время года они будут организованы.

Из минусов было выявлено, что практически все площадки находились в центре Санкт-Петербурга, а шестьдесят одно мероприятие имело продолжительность до четырнадцати дней, что приводило к перегруженности локаций, отрицательно влияя на восприятие среды посетителем (рис. 2). Выход площадок за пределы центра города, может создать благоприятную среду для жителей и туристов, которые не будут испытывать дискомфорт от перегруженности локации, при этом принесет дополнительный доход для этого места, за счет увеличения проходимости. Необходимо обеспечить легкодоступное расположение площадки (вблизи метро, автобусных остановок). Однако, не стоит организовывать проекты на постоянной основе вблизи жилой зоны, так как через некоторое время их цвет и динамика, станут утомительными для жителей района.

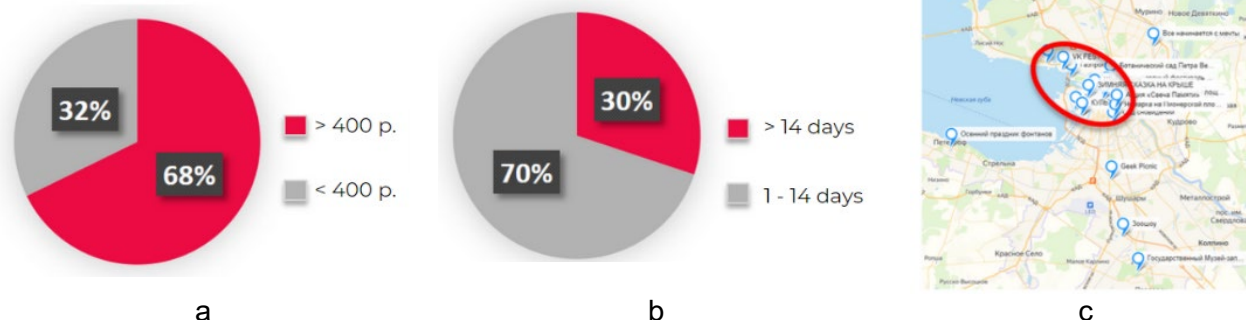


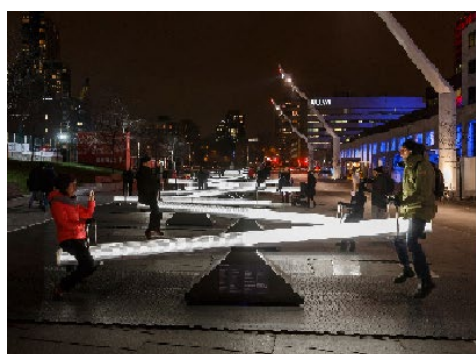
Рис. 2. Статистические данные по мероприятиям в Санкт-Петербурге: а) посещаемость; б) длительность; с) расположение

Важно подчеркнуть, что оригинальные разовые художественные концепции освещения привлекают к себе большее внимание, чем стандартно повторяющиеся. Например, в случае с фестивалем Wonder of Light, в разработке концепции и оформлении мероприятия участвовали профессиональные дизайнеры по свету, но это единичный случай. На большинстве недавних проектах было выявлено, что чаще всего использовалось светотехническое оборудование для сценического освещения, технология “3d mapping” и пиротехническое оборудование (рис. 3). И в связи с тем, что большинство программ было ограничено 1–3 применяемыми группами

средств искусственного освещения, это отрицательно сказалось на способности к длительному удержанию внимания пользователя, и привело к нежеланию повторного посещения мероприятий.

Одним из основных недостатков является тот факт, что все световые решения были направлены только на зрительное восприятие, и практически не задействовались интерактивные сценарии, которые как известно, существенно влияют на успех мероприятия в части привлечения и удержания внимания пользователя. Следовательно, возникает необходимость вывести работу с искусственным освещением на новый уровень, где оно сможет взаимодействовать с дополнительными элементами, благодаря которым можно будет влиять на формирование нового опыта коммуникации с окружающей средой.

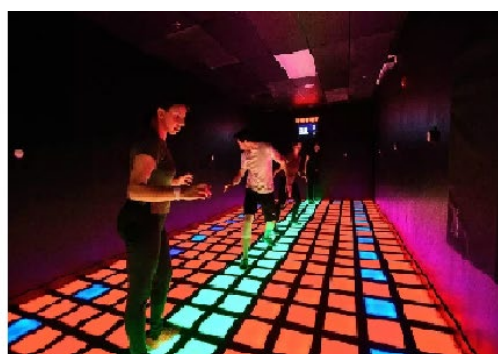
Для решения этой проблемы необходимо внедрить в искусственное освещение процесс взаимного интерактива с пользователем, то есть геймификацию. Геймификация – это использование игровых механик, поведенческих элементов и методов дизайна в неигровом контексте [2]. Некоторые игровые концепции смогут пригласить участников к общению («Импульс» (рис. 4.а), другие - подтолкнуть игроков к исследованиям (Музей Хет Норд Брабант - Редизайн цифровых медиа — Студия Louter (рис. 4.б), третьи - больше внимания уделяют достижениям (интерактивная комната в Берлингтоне (рис. 4.с).



а



б



с

Рис. 4. Примеры световых сценариев, основанные на методе геймификации [3-5]

Свет должен призывать к действию, общению, способствовать интерактивности, веселью в ночной обстановке, которые обычно затихают после наступления темноты. Событийные пространства с применением светового оборудования, безусловно, могут поощрять совместное «игровое» участие пользователей. Но при этом следует учесть, что лишь небольшие мероприятия способствуют социальной интеграции, созданию сетей и укреплению связей [6]. В больших масштабах чрезмерная активность может привести к хаосу.

Общественные пространства должны быть пригодны для всех возрастных групп. Как отмечает Д. Джейкобс, городская жизнь создается городами, которые имеют как можно большее разнообразие. Слабая функциональность пространств в городе создает монотонность и снижает их привлекательность. Если план пространственного разнообразия составлен правильно, он не вызовет хаоса, а будет побуждать людей проводить больше времени в общественных местах. Их дизайн является частью «эволюционного процесса проектирования», и эти места должны быть адаптированными, гибкими и долговечными, чтобы хорошо служить обществу [7].

На основе полученных данных актуальной представляется необходимость разработки общей методики: с рекомендациями по используемому осветительному оборудованию; его

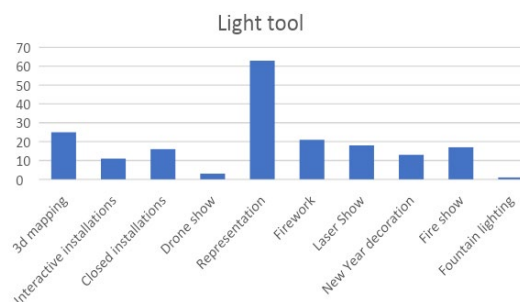


Рис. 3. Светотехническое оснащение событийных пространств города Санкт-Петербург, за период с 2017 — 2021 гг.

художественно-выразительному применению; способным обеспечить повышение пользовательского интереса при формировании светового сценария для вечерних событийных пространств открытого типа. Такой подход должен быть направлен на поддержание долгого «социального» контакта пользователя со средой, а также перераспределять потоки и формировать новые «точки роста» в нецентральных районах Санкт-Петербурга.

Подход, ориентированный не только на эстетическую сторону событийного освещения и его оригинальность, но и на визуальное восприятие, интерес и эмоции человека, сможет сформировать новые точки притяжения, актуализировать городские пространства за пределами исторического центра города, сможет повысить качество жизни горожан, а также расширит область понимания воздействия света на поведенческие особенности людей и их психологию.

Литература

1. Zielinska-Dabkowska K. The Role of Cultural Institutions and Events in the Marketing of Cities and Regions. [Электронный ресурс]. – Режим доступа: https://www.researchgate.net/publication/305315339_Night_in_a_big_city_Light_festivals_as_a_creative_medium_used_at_night_and_their_impact_on_the_authority_significance_and_prestige_of_a_city (дата обращения: 25.01.2022).
2. Мастюгин А.Ю., Суранов А.Ю. Привлечение, удержание и возвращение потребителя посредством геймификации // Практический маркетинг. – №8. – 2020. – С. 19–24.
3. [Электронный ресурс]. – Режим доступа: https://static.wixstatic.com/media/ffcf7_0490ac4c3faf4612b24a80c61c909f15~mv2.png/v1/fit/w_920%2Ch_720%2Cal_c/file.png (дата обращения: 25.01.2022).
4. [Электронный ресурс]. – Режим доступа: <https://images.squarespace-cdn.com/content/v1/5f6b34f1d2d9182bf9cdc0d0/1611655791805N51KLVK5TTNNQB2H96KG/studio-louter-het-noordbrabants-museum-digital-media-redesign-behind-the-scenes-siege-of-den-bosch-film?format=750w> (дата обращения: 25.01.2022).
5. [Электронный ресурс]. – Режим доступа: <https://a.mktgcdn.com/p/afWTVePfgdi81V8IFUyYCIWoVDWmbfchGM4sbS1k5Y/8256x5504.jpg> (дата обращения: 25.01.2022).
6. Endensor T. The virtues of contemporary light festivals. [Электронный ресурс]. – Режим доступа: <https://www.luciassociation.org/magazine/Cities-Lighting-008/> (дата обращения: 25.01.2022).
7. Емлина Д.И. Видеоэппинг в пластическом театре // Современный танец: дискурс и практики : сб. ст. / под общ. ред. канд. культурологии Н.В. Курюмовой. – Екатеринбург: Гуманитарный ун-т. – 2017. – 157 с.

УДК 628.974.8

ВЛИЯНИЕ ПРИЕМОВ АРХИТЕКТУРНОГО ОСВЕЩЕНИЯ НА ВЫЯВЛЕНИЕ ДОМИНАНТНОСТИ ОБЪЕКТОВ ГОРОДСКОЙ СРЕДЫ

Николаева Н.В.¹ (студент)

Научный руководитель – канд. архитектуры Соколова М.А.¹

¹*Университет ИТМО*

e-mail: nikolaeva.natvit@mail.ru

Рассматриваются проблемы сумеречного и ночного освещения архитектурных доминант города и визуального восприятия пользователей на основе критериев оценки параметров освещенности и привлекательности.

Ключевые слова: архитектурные доминанты, световые доминанты, планировочная структура, визуальное восприятие, потоки пешеходов.

Освещение играет ключевую роль в формировании архитектурного пространства. Долгое время его относили к инженерно-технологической отрасли, не придавая значения психологической и эстетической стороне вопроса. Однако, разработка световых сценариев улиц позволяет направлять пешеходные потоки, привлекая туристов, снизить утомляемость и число стрессов у населения, повысить общий престиж города и его властей. Учет соотношения фона и доминанты позволяет создавать городскую среду, которая наиболее читаема и пригодна для жизнедеятельности.

Современные исследования архитектурного освещения находятся на стыке технологий и эстетики. Они базируются на выявлении параметров, обеспечивающих адекватное восприятие пользователем среды, отношений элементов городского пространства и иерархии объектов. Сегодня также существует необходимость выработки энергоэффективных решений для снижения светового загрязнения.

Освещение всегда определяло наше отношение к окружающей среде и друг к другу. Оно было важным условием создания гостеприимных, безопасных и привлекательных городов. Например, в средневековой Европе люди по закону должны были носить фонари, чтобы сообщать о своем присутствии согражданам. Сегодня, как и в прошлом, освещение создает чувство сопричастности пространству, как у жителей, так и у гостей.

Световая доминанта, как элемент светопланировочной структуры, наряду со световым ансамблем и световым акцентом выделены в работах Н.И. Щепеткова. Им разработана теоретическая модель цветоцветовой структуры среды – «система параметров структурно-иерархической и масштабной дифференциации городского светопространства» [1]. Отсюда определяются световые параметры яркости и освещенности, которые дают необходимый световой эффект, а также имеют оптические показатели, они и определяют разнообразие света, образов и эстетики вечернего города.

Для достижения баланса между желаемым эффектом и его влиянием на траекторию движения пользователей необходимо определить параметры гармоничного восприятия среды и сохранения характеристик здания, таких как его цвет и материал, объем, идентичность и роль в городском пространстве [2].

В результате исследования исторических источников было выявлено, что при строительстве Санкт-Петербурга для каждой доминанты выделялось открытое ландшафтное

пространство; таким образом, формировался «принцип городского ансамбля». Открытые пространства и доминанты сформировали историческую застройку центральной части города. Кроме пространственно-силуэтной гармонии городские акценты формируют еще и чувство безопасности граждан – ориентируясь на какой-либо доминантный объект среди площадей и проспектов, всегда знаешь, где находишься. Статус охранной зоны имеет не только система открытых архитектурных пространств, но и так называемые «открытые» зоны: пути обзора и точки восприятия объектов.

Разработка светового мастер-плана города тесно связана со структурой генплана. Главной целью при формировании системы доминант становится сохранение исторического силуэта Санкт-Петербурга, его целостной панорамы, выразительного, запоминающегося внутреннего и внешнего силуэтов города и создание системы ориентирования с помощью доминант в периферийных районах. В 2008 году при разработке Правил землепользования и застройки Петербурга был создан первый полноценный регламент высотности построек в городе, который коснулся и архитектурных доминант.

Существует несколько типов освещения: утилитарное, архитектурное, ландшафтное, информационное. Яркостные характеристики световых доминант и расчеты их освещенности относятся к архитектурному типу и приведены в нормативных требованиях; данное освещение описывается как акцентирующее, заливающее, локальное, комбинированное. Световые доминанты – это не только преобладающие по высоте элементы сооружений, но и малые архитектурные формы, служащие пространственными ориентирами в темное время суток.

В зависимости от размеров общегородского центра и его планировочной структуры следует организовывать системы взаимосвязанных общественных сред (главных улиц, площадей, пешеходных зон), составляя ядро общегородского центра. Особое внимание сейчас уделяется благоустройству пешеходных улиц (к ним относятся аллеи, бульвары, набережные). По этой причине стоит учитывать разность восприятия световых доминант водителями и пешеходами. Скорость движения влияет на цели, которым может соответствовать освещение. Для водителей приоритетными характеристиками восприятия являются: концентрация внимания, поддержание работоспособности и комфорта зрения. Для пешеходов: возможность среднего ориентирования, безопасность, художественная привлекательность [1].

Высотные доминанты и ансамбли определяют вертикальную композицию города и влияют на его панораму и силуэт [3]. Их структурообразующая роль может иметь компактную, вытянутую или вертикально развитую форму. В то же время высота доминант формирует высоту градостроительной композиции. Доминанты могут быть объединены в одну или две вертикали, ряд, кольцо, группу вертикалей. Композиционные центры и оси определяют градостроительную структуру и эстетику города, влияют на движение потока пешеходов [4].

Световые и архитектурные доминанты можно отнести как к ландшафтному масштабу, если это касается крупных градостроительных образований, так и к ансамблевому при восприятии силуэта города (набережные, улицы). Восприятие пешеходом общественного пространства, фасадов зданий, а также взаимодействие с малыми архитектурными формами – это черты камерного масштаба. В нем человек перемещается в соизмеримом с ним пространстве двора, улицы, площади. Важным и конкретным определением композиции вечернего света является распределение световых пятен по яркости.

Цветовая температура света влияет на психику человека, создаёт атмосферу, активизирует его или расслабляет. Теплый свет создает атмосферу защищенности и уюта, а холодные оттенки помогают взгляду концентрироваться. Современная урбанистика отдает предпочтение теплоте свету в диапазоне 2700 – 3000 К.

Немаловажную роль играет и цвет фасада. Важно отметить, что теплые источники света плохо сказываются на «холодных» фасадах – в таком случае к ним рекомендуют подбирать более холодный свет, например, 4000 – 5000 К, нейтральный белый свет считается универсальным, а для натуральных деревянных, природных оттенков подходит теплый свет, для металлических, соответственно, — холодный.

При формировании системы доминантной застройки предлагается использовать традиционные методы и принципы формирования застройки Санкт-Петербурга.

1. Застройку необходимо разделять на фоновую и доминантную с соблюдением главных для Санкт-Петербурга геометрических и качественных особенностей силуэта, обеспечивающих контраст восприятия. Минимальный разрыв между фоновой и доминантной застройкой – 1/3 высоты фоновой застройки. Разрыв может быть и больше, если выше статус размещаемой доминанты. Доминантная застройка должна превышать фоновый объект более, чем в 2 раза.
 - для территорий зон компактного размещения доминантной застройки – более чем на 75%;
 - для главных городских доминант – от 40 до 75%;
 - для основных доминант районного уровня – от 30 до 50%.
2. Доминантная застройка и ее фон должны создавать сложную линию горизонта с включением акцентных силуэтов.
3. Необходимо учитывать восприятие доминанты из открытых общественных территорий с ближних и дальних точек обзора при выборе места размещения.
4. Расположение доминанты должно быть обосновано планировочной композицией транспортных маршрутов.
5. Доминантная застройка также может быть расположена в наиболее значимых точках города (входные зоны, пересечения основных магистралей), способствуя пространственному ориентированию, проявляя основные направления и пути.
6. Доминанты предлагается разделять на главные, основные, прочие и акценты.
7. Необходимо соблюдать традиционную плотность размещения доминант.
8. Соответствовать сложившейся в Санкт-Петербурге с XX века (70-х годов) методике регулирования силуэта.

На данный момент применяются различные способы измерения качества световой среды и ее восприятия с помощью технических средств. Это означает, что оценка архитектурного освещения не должна производиться на основе субъективного анализа ее визуальных качеств. Существует также возможность введения новых параметров оценки и показателей, направленных на обеспечение энергоэффективности и снижение светового загрязнения.

Современные исследования показывают, что с помощью применения технологии айтрекинга, яркомера и спектрометра можно доказать, что оптимизация яркости доминантной единицы не влияет на визуальную привлекательность с точки зрения конечного пользователя, а объект продолжает оставаться архитектурным и световым ориентиром [5], в том числе сохраняя свое структурно-иерархическое значение в градостроительной композиции.

Исследования в области светового дизайна города позволили выделить несколько важных факторов:

- на восприятие городских доминант в темное время суток наиболее влияет яркость и цветовая температура;
- разные световые характеристики зданий создают неравномерность восприятия улицы, что позволяет подчеркнуть уникальность доминантных объектов в сумерках и ночью;
- дизайн с соблюдением масштаба человека и световые акценты способствуют безопасности пешеходов и созданию комфортных пространств.

Таким образом, разработка комплексного подхода к формированию световой среды вечернего города включает выявление критериев и определение факторов, влияющих на построение иерархии взаимодействия всех видов освещения с задачей художественной гармонизации городского пространства.

Литература

1. Skarżyński K. The balance between visual effect and engineering correctness in architectural lighting / K. Skarżyński // Proceedings of the 29th Quadrennial Session of the CIE – 2019. [Электронный ресурс]. – Режим доступа: https://www.researchgate.net/publication/333982969_the_balance_between_visual_effect_and_engineering_correctness_in_architectural_lighting (дата обращения: 03.09.2021).

2. Щепетков Н.И. Световой дизайн города. М.: Архитектура-С. – 2006. – 317 с.
3. Бархин М.Г. Город. Структура и композиция. Москва. – 1986. – 264 с.
4. Косицкий Я.В. Композиционные основы планировочной структуры города. Москва. – 1984. – 124 с.
5. Noguera M., Pernão J. Colour and light in exterior architectural illumination: From efficiency to aesthetics/ M.Noguera, J. Pernão. // AIC Midterm Meeting Color and Landscape – 2019. [Электронный ресурс]. – Режим доступа: https://www.researchgate.net/publication/336848178_Colour_and_light_in_exterior_architectural_illumination_From_efficiency_to_aesthetics (дата обращения: 18.10.2021).

УДК 004.932; 550.388.8

ПОДГОТОВКА ИЗОБРАЖЕНИЙ, ПОЛУЧЕННЫХ КАМЕРОЙ ВСЕГО НЕБА, ДЛЯ КОМПЛЕКСНОГО АНАЛИЗА ВОЛНОВЫХ СОБЫТИЙ В МАГНИТОСФЕРЕ

Смотрова Е.Е.¹ (студент), Михайлова О.С.², Сыренова Т.Е.²

Научный руководитель – канд. техн. наук Ежова В.В.¹

¹*Университет ИТМО*

²*Институт солнечно-земной физики СО РАН*

e-mail: katerina.smotrova@mail.ru, o_mikhailova@iszf.irk.ru, angata@mail.iszf.irk.ru, evv_foist@mail.ru

Данная работа посвящена предварительной обработке изображений, получаемых камерой всего неба станции Исток (ИСЗФ СО РАН), установленной в рамках проекта PWING. Были написаны программы, направленные на восстановление интенсивности и улучшение визуального качества изображений. Написана программа для перевода географических координат изображения в геомагнитные с использованием моделей магнитного поля вблизи Земли и на дальних расстояниях.

Ключевые слова: камера всего неба, полярные сияния, стереоизображения, удаление звезд, атмосферные эффекты, гамма-коррекция.

Изучение явлений, происходящих во внутренней магнитосфере, представляет большой интерес для научного сообщества. Одним из важных является исследование взаимодействия энергичных частиц (электронов и ионов) с волнами, в том числе ультранизкочастотными (УНЧ) волнами с диапазоном частот порядка нескольких мГц. Считается, что процессы взаимодействия волна-частица могут ускорять электроны до энергий порядка нескольких МэВ и способствовать высыпанию частиц в атмосферу.

Для изучения подобных процессов существует множество инструментов как наземных, так и спутниковых. Чтобы восстановить полную цепочку процессов от магнитосферы до атмосферы, необходим совместный анализ спутниковых и наземных наблюдений. Однако, подобных работ на сегодняшний день мало, так как эта задача является трудоемкой [1].

Среди наземных инструментов особое место занимают камеры всего неба. С их помощью получают изображения полярных сияний, представляющие собой свечения нейтральных частиц атмосферы из-за воздействия ускоренных и промодулированных УНЧ-волнами потоков заряженных электронов. Для возможности проведения совместного анализа процессов взаимодействия волн и частиц с использованием наземных наблюдений полярных сияний и спутниковых данных необходимо проводить исследования событий в привязке не к географическим координатам, а к геомагнитным. Так, привязка изображений полярных сияний к геомагнитным координатам позволит локализовать область магнитосферы, где произошло взаимодействие волны с энергичными частицами, вызвавшие наблюдаемое полярное сияние. В свою очередь знание о возможной области локализации события даст возможность для более быстрого отбора тех спутниковых данных, которые проходили в данное время через выделенную область, для дальнейшего исследования и установления взаимосвязи наблюдаемых событий в магнитосфере и атмосфере.

При этом для качественного анализа необходима подготовка изображений и создание наиболее эффективного представления временной зависимости изменения полярных сияний. На данный момент исследования полярных сияний проводятся лишь в географической системе

координат, где временную зависимость изменения полярных сияний представляют с помощью кеограмм (срез пикселей изображений по линии север-юг) по широте или эвограмм (срез пикселей изображений по линии восток-запад) по долготе [1]. В данной статье мы представим первые результаты предварительной обработки изображений полярных сияний для проведения последующего анализа волновых событий.

Для исследования полярных сияний нами была выбрана камера всего неба, размещенная на высокоширотной станции Исток (70.03N, 88.01E) Института солнечно-земной физики СО РАН. Данная камера вместе с другими инструментами была установлена в рамках международного проекта PWING [2]. Камера оборудована пятью интерференционными фильтрами, три из которых имеют временное разрешение 1,5 мин, что является достаточным для исследования волновых событий в миллигерцовом диапазоне.

Изображения, полученные рассматриваемой камерой всего неба, представлены в градациях серого с глубиной цвета в 16 бит и размером 511×511 пикселей. Для оптимизации времени обработки изображения была создана круговая маска. Через данную маску проходят все пиксели, входящие в окружность с радиусом в 255 пикселей, остальным присваивается значение *NaN*. Последние в дальнейшем при обработке изображения не рассматриваются.

Для того, чтобы использовать изображения камеры для дальнейшей обработки, необходимо производить калибровку камеры всего неба. Этот шаг нужен для привязки пикселей изображения камеры к географическим координатам, где необходимо учитывать возможные нарушения юстировки камеры и аберраций при использовании широкоугольных систем. В силу того, что на данном этапе основной интерес представляет обработка самого изображения, то для отработки алгоритмов была создана произвольная координатная сетка. В дальнейшем планируется использование программ калибровки камеры всего неба по звездам, разработанных в лаборатории физики нижней и средней атмосферы ИСЗФ СО РАН [3].

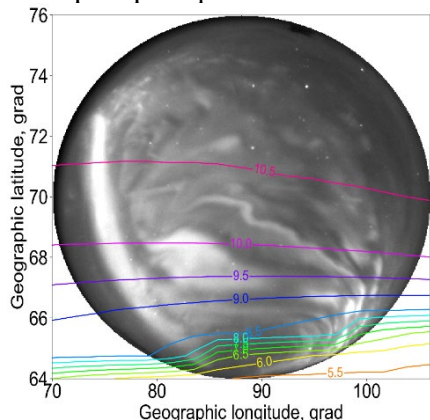


Рис. 1. Изображение камеры всего неба с географическими координатами. Снимок зарегистрирован 02.12.2018 в 19:59:57 UT широкополосным фильтром излучения полосы гидроксила ОН (>720 нм)

В рамках данной работы написана программа для преобразования географических координат в геомагнитные. В качестве геомагнитных координат были выбраны геомагнитная широта и долгота, а также параметр МакИлвейна, который характеризует максимальное удаление силовой линии от центра Земли. Для расчета первых двух координат использовалась модель магнитного поля вблизи поверхности Земли IGRF-13, в то время как для расчета параметра МакИлвейна – модели Цыганенко, описывающие магнитное поле вдали от поверхности Земли. Для наших целей применялась модель T96, имеющая явно заданную реалистичную магнитопаузу, крупномасштабные токовые системы Биркеланда и учитывает проникновение межпланетного магнитного поля через границу [4]. На рисунке 1 приведен пример изображения полярного сияния без обработки с рассчитанным параметром МакИлвейна.

Стоит отметить, что на рисунке 1 представлено изображение до проведения обработки.

Обработку изображений необходимо провести с целью автоматического выделения дуг полярных сияний на фоне неба. На данный момент в программу обработки изображений полярных сияний включены следующие функции: удаление звезд, восстановление интенсивности изображения и гамма-коррекция. Перейдем к более подробному описанию данных этапов.

Процедура удаления звезд производится для того, чтобы исключить появления резкого локального изменения интенсивности, определяемой функцией рассеяния точек системы, или широкополосных помех при непосредственном анализе динамики полярных сияний, например, при получении двумерного Фурье-спектра. Путем анализа возможных способов нами был выбран метод статистической вариации интенсивности пикселей [5]. В данном методе

используется построчный расчет среднего значения и стандартного отклонения пикселей, находящихся справа и слева от рассматриваемого пикселя. В случае, если значение отклоняется от предела в два стандартных отклонения, то его значение вычисляется с помощью интерполяции. Главное преимущество данного метода в отличие, например, от методов с использованием фильтра Винера или медианного фильтра, заключается в том, что при его использовании не происходит ухудшение качества самого изображения. На рисунке 2 приведен результат реализации данного алгоритма. Стоит отметить, что при обработке данной процедуры может некорректно работать для звезд, наблюдавшихся на фоне полярных сияний. Данную проблему можно будет решить в дальнейшем с использованием вышеупомянутой программы калибровки камеры по звездам.

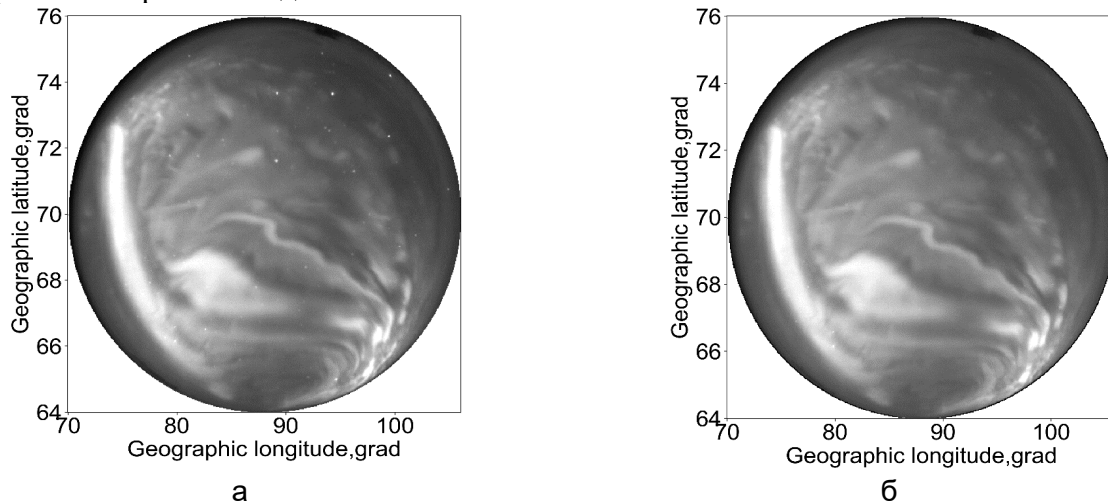


Рис. 2. Тестовое изображение, полученное камерой всего неба а) до удаления звезд; б) после проведения удаления звезд

Следующий шаг предполагает восстановление интенсивности изображения путем учета шумов и искажений, которые вносятся физическими характеристиками широкоугольного объектива. Данным искажениям подвержены объекты, расположенные на значительном расстоянии от оптической оси объектива, поэтому в основном они локализируются по краям изображения. Восстановление интенсивности проводится путем учета эффекта Ван Рина и атмосферного поглощения [5].

Первый эффект возникает в случае, когда объект, находящийся вдали от оптической оси, концентрируется на нескольких пикселях по краям изображения. В результате интенсивность этих пикселей становится ярче пикселей, расположенных ближе к оптической оси. Возникающая неравномерность баланса яркости задается уравнением:

$$I(\theta) = \left[1 - \left(\frac{R_E}{R_E + h_{ag}} \right)^2 \sin^2 \theta \right]^{-\frac{1}{2}} \cdot I(0), \quad (1)$$

где $I(\theta)$ – интенсивность под углом наклона θ ,
 R_E и h_{ag} – радиус Земли и высота наблюдаемого свечения соответственно [5].

Второй эффект ослабляет наблюдаемую интенсивность на краях изображения. Причиной служит тот факт, что свет, исходящий из объекта, расположенного вдали от оптической оси, ослабляется атмосферой. Таким образом, из двух одинаковых по интенсивности астрономических объектов сильнее будет ослабляться интенсивность того, который находится дальше от оптической оси, так как он будет иметь большую оптическую длину пути. Данный процесс сам по себе зависит от множества параметров атмосферы, поэтому на практике пользуются эмпирической формулой:

$$\begin{cases} I(\theta) = I_{true}(\theta) \cdot 10^{-0.4a \cdot F(\theta)} \\ F(\theta) = (\cos \theta + 0.15 \cdot (93.885 - \theta))^{-1} \end{cases} \quad (2)$$

где a – коэффициент атмосферного поглощения,
 $I_{true}(\theta)$ – интенсивность от астрономического объекта,
 $I(\theta)$ – интенсивность под углом наклона θ .

Стоит отметить, что коэффициент атмосферного поглощения a зависит от длины волны и должен рассчитываться отдельно [5].

Данная коррекция была проведена на тестовом изображении, при этом угол наклона был задан так, что он уменьшается обратно пропорционально от центра изображения (рис. 3). Результат коррекции приведен на рисунке 4, где можно отметить не только падение интенсивности изображения по краям, но также увеличение интенсивности некоторых дуг полярных сияний к центру изображения.

При расчетах был принят коэффициент атмосферного поглощения $a = 0,3$.

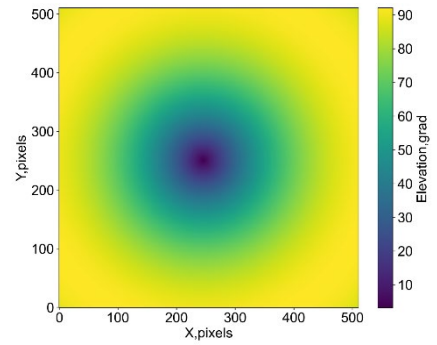


Рис. 3. Распределение значений угла наклона для восстановления интенсивности изображения путем учета атмосферных эффектов

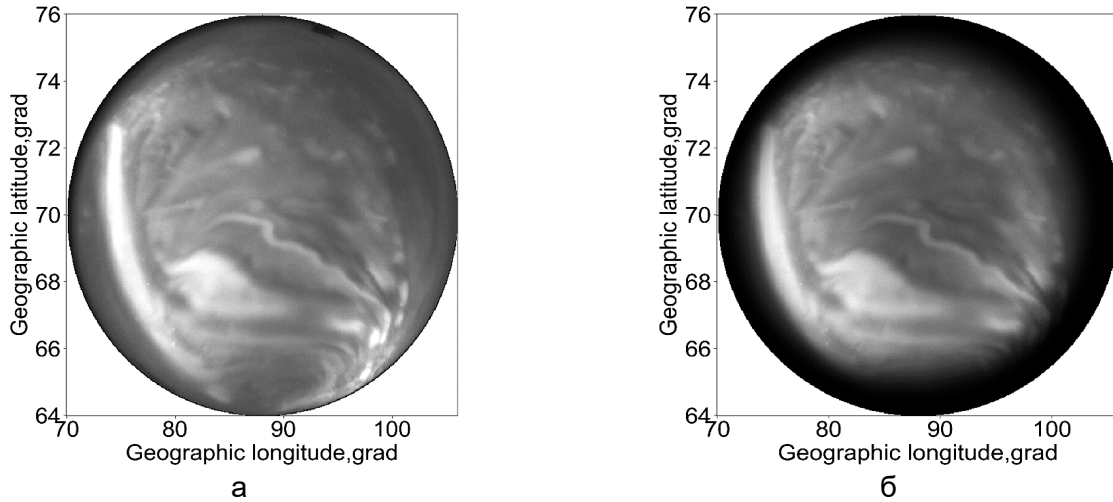


Рис. 4. Тестовое изображение, полученное камерой всего неба а) до обработки; б) после восстановления интенсивности изображения путем учета атмосферных эффектов

Последним пунктом обработки изображения стало проведение гамма-коррекции, представляющая из себя степенное преобразование интенсивности пикселей изображения и служащая для точного воспроизведения изображения на экране монитора. Эта процедура позволит более четко в визуальном плане выделить волновые авроральные структуры за счет улучшения динамического диапазона в темной области. Гамма-коррекция проводилась с использованием следующей формулы:

$$I_y = I_{\max} \cdot \left(\frac{I}{I_{\max}} \right)^{\frac{1}{\gamma}} \tag{3}$$

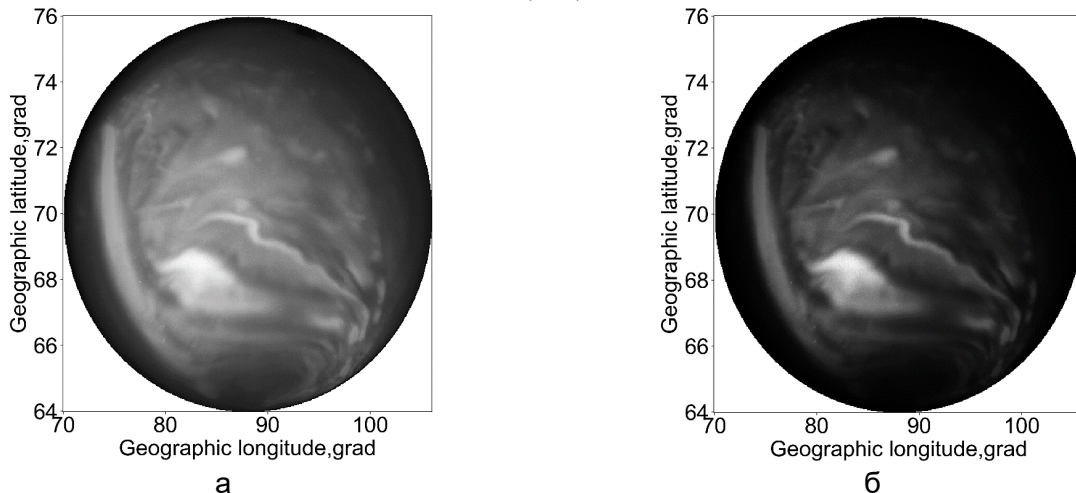


Рис. 5. Тестовое изображение, полученное камерой всего неба а) до коррекции; б) после гамма-коррекции изображения

Пример проведенной процедуры гамма-коррекции представлен на рисунке 5. При расчетах было принято $\gamma = 0,45$, в силу того, что данный параметр дает наилучшее визуальное качество для наблюдения волновых авроральных структур.

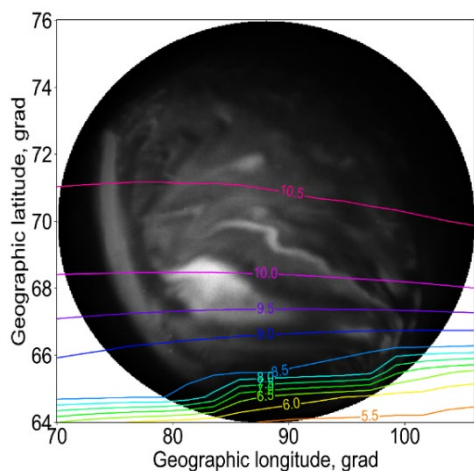


Рис. 6. Обработанное изображение с наложенными геомагнитными оболочками, рассчитанными по модели Цыганенко Т96

По результатам проделанной работы нами была разработана программа для обработки изображений, полученных камерой всего неба, для исследования полярных сияний в привязке к геомагнитным координатам. Пример конечного варианта изображения полярных сияний с наложенными значениями параметров МакИлвейна представлен на рисунке 6.

В дальнейшем планируется объединение написанных программ в единый модуль с некоторыми доработками. Так, планируется добавить в обработку изображения коррекцию плоского поля, целью которого является исправление систематических пространственных неоднородностей, связанных с работой ПЗС-матрицы. Для каждого интерференционного фильтра в будущем будет проведена оценка коэффициента атмосферного поглощения a с помощью спектрофотометрического каталога звезд.

Результаты программы преобразования географических координат в геомагнитные планируется использовать для представления временной зависимости изменения полярных сияний. Для этого будет написан ряд программ, в которых будет производиться срез интенсивности полярных сияний в зависимости от заданных параметров. Например, можно будет производить срез интенсивности полярных сияний в фиксированной области магнитосферы, где мог находиться космический аппарат на момент начала волнового события в магнитосфере.

Литература

1. Motoba T., Takahashi K., Ukhorskiy A., Gkioulidou M., Mitchell D.G., Lanzerotti L.J., Korotova G.I., Donovan E.F., Wygant J.R., Kletzing C.A., Kurth W.S., Blake J.B. Link between premidnight second harmonic poloidal waves and auroral undulations: conjugate observations with a Van Allen Probe spacecraft and a THEMIS all-sky imager // *J. Geophys. Res. Space Phys.* – 2015. – №120. – Pp. 1814-1831. DOI: 10.1002/2014JA020863.
2. Shiokawa K., Katoh Y., Hamaguchi Y. et al. Ground-based instruments of the PWING project to investigate dynamics of the inner magnetosphere at subauroral latitudes as a part of the ERG-ground coordinated observation network // *Earth Planets Space.* – 2017. – №69. DOI: 10.1186/s40623-017-0745-9.
3. Сыренова Т.Е., Белецкий А.Б., Васильев Р.В. Географическая привязка кадров широкоугольных систем // *Журнал технической физики.* – 2021. – Т.91. – Вып.12. – С. 1990–1996. DOI: 10.21883/JTF.2021.12.51765.172-21.
4. Tsyganenko N.A. Effects of the solar wind conditions on the global magnetospheric configuration as deduced from data-based field models // *Proceedings of the Third International Conference on Substorms (ICS-3).* Eur. Space Agency Spec. Publ., ESA-SP. – 1996. – №389. – Pp. 181–185.
5. Kengne F.F., Mundra R., Rangel M.A. Mid-latitude All sky imager Network for Geophysical Observation // *WPI.* – 2013. DOI: 10.13140/2.1.1789.7762.

УДК 535.361

ПОРТАТИВНАЯ ОПТИКО-ЭЛЕКТРОННАЯ СИСТЕМА ДЛЯ МОНИТОРИНГА ОПТИЧЕСКИ НЕОДНОРОДНЫХ СРЕД

Хлынов Р.Д.¹ (студент), Чан Ван Тхань¹ (студент)

Научный руководитель – канд. техн. наук, доцент Рыжова В.А.¹

¹Университет ИТМО

e-mail: super.ryslan3@yandex.ru, victoria_ryz@itmo.ru, tranthanh.ht.1203@gmail.com

Работа финансируется Министерством науки и высшего образования РФ (грант 08-08).

В статье рассмотрены теоретические аспекты методик калибровки поляриметров. Определены перспективы и возможности использования представленных методик для портативных оптико-электронных систем, направленных на реализацию поляризационного метода в биологических и медицинских сферах. Анализ точности вычислительных процедур установил, что использование произвольного набора линейно-независимых калибровочных векторов для оптимизации результатов измерения параметров вектора Стокса и элементов матрицы Мюллера позволяет существенно снизить погрешность определения приборной матрицы системы.

Ключевые слова: смартфон, калибровка поляриметров, приборная матрица, токовые отклики, вектор Стокса.

Устройства формирования изображений на базе смартфонов – это платформы, которые используют возможности смартфона для преобразования оптических изображений с последующей их обработкой программно-аппаратными средствами и визуализацией на экране. Видеоинформационные системы смартфонов широко используются для различных целей в качестве сенсорных устройств ввода и отображения информации в составе микроскопических устройств, аналитических блоков систем обнаружения, а также устройств для мониторинга атмосферных загрязнений и определения типов загрязняющих веществ [1].

Портативность оптико-электронной системы специального назначения на базе смартфона достигается с помощью либо дополнительной оптической насадки для камеры, либо мобильных приложений для обработки изображений от встроенной камеры. Современный смартфон с улучшенной интеграцией аппаратного и программного обеспечения может использоваться как автономная платформа со встроенными источником, детектором и вычислительной платформой для разработки автономных мобильных сенсорных устройств, которые находят применение в различных аналитических приложениях [1, 2].

Быстрое развитие смартфонов определило их основной выбор в области доступных биомедицинских услуг. Визуализация на базе смартфонов была предложена для многочисленных биомедицинских приложений, многие из которых используют оптическую насадку для расширения возможностей оптико-электронных устройств. В последнее время наиболее распространенным применением диагностических визуализаций на базе смартфонов является неинвазивный мониторинг биологических тканей и сред, доступных снаружи, особенно в дерматологии. Сообщается также о применении портативных оптико-электронных систем на базе смартфонов для минимизации инвазивных процедур в спектроскопии, эндоскопии, микроскопии и хирургии [3].

Главными преимуществами таких систем являются подтверждение предварительного диагноза пациента и быстрый обмен информацией и данными для получения оперативного анализа, что делает всю диагностическую систему доступной для большей части населения.

Интерес научных сотрудников к применению смартфонов в пунктах оказания медицинской помощи возникает из-за многих факторов, таких как мониторинг на месте в режиме реального времени, быстрое тестирование и низкая стоимость лабораторных приборов [4, 5]. Возможность быстрой количественной и качественной оценки биологических образцов на основе анализа их изображений делает программную среду устройств смартфонов очень перспективной, в особенности для мониторинга и диагностики различных заболеваний.

Проведение биоаналитических исследований на базе смартфонов основано на реализации усовершенствованных оптических схем и технологий обработки данных. При этом смартфон можно рассматривать как микроскопическое устройство лабораторного класса. Многими авторами было продемонстрировано, что использование смартфона в схемах поляриметров в качестве недорогого микроскопического устройства позволяет проводить различные биоаналитические исследования, а в качестве экономически эффективного инструмента визуализации позволяет проводить быстрый анализ крови [6].

Так как поляриметры применяются для измерения параметров оптически неоднородных структур биологических объектов, возрастают требования к метрологическому обеспечению измерительных средств на базе смартфонов. Одной из важных процедур, направленных на снижение погрешности измерения параметров оптически анизотропных сред с использованием поляриметров, является калибровка, которая позволяет учесть в качестве поправок влияние «неидеальности» элементов оптического тракта приборов на результат измерения.

По результатам анализа информационных источников зарубежной литературы было установлено, что поляриметрию диффузно рассеивающих сред, в том числе биологических тканей, можно разделить на две категории: поляриметрия Стокса и поляриметрия Мюллера, а реализуемые при этом оптические схемы построения на Стокс- и Мюллер-поляриметры.

Стокс-поляриметры используются для анализа прохождения частично поляризованного излучения через исследуемые деполаризующие оптические системы и позволяют определять их параметры на основе выявления их зависимости от измеренных параметров поляризации выходного излучения. Мюллер-поляриметры используются для анализа распространения излучения через различные оптические системы и позволяют определять шестнадцатиеlementную матрицу Мюллера объекта, которая содержит всю информацию об оптических параметрах его поверхностной и внутренней структуры [7]. Стокс-поляриметры представляют наибольший интерес в связи с тем, что измерительная процедура легко автоматизируется и обработка измерений параметров занимает намного меньше времени, чем при использовании Мюллер-поляриметра, что очень важно для портативных оптико-электронных систем, используемых для мониторинга оптически неоднородных сред. Поэтому целью данного исследования является критический обзор современной методики калибровки Стокс-поляриметров.

Одна из методик калибровки визуального Стокс-поляриметра заключается в осуществлении запуска серии известных состояний поляризации излучения во входной зрачок поляриметра, оптическая установка которой приведена на рисунке.

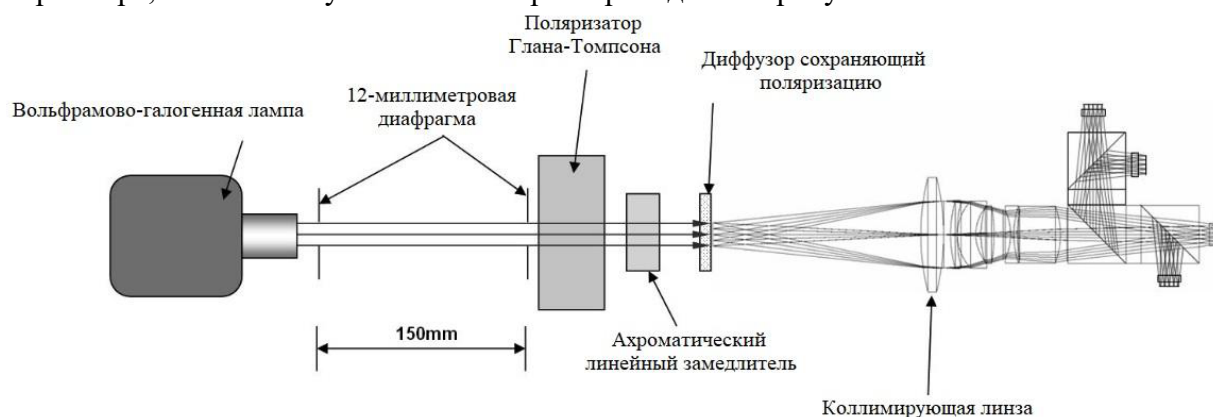


Рисунок. Оптическая схема для калибровки визуального Стокс-поляриметра [8]

Предложенная оптическая схема является универсальной для калибровки четырехканальных систем на основе пространственного разделения интенсивности входного излучения по амплитуде без вращающихся элементов. В состав схемы входит вольфрамово-галогенная лампа, две двенадцатимиллиметровые диафрагмы, поляризатор, ахроматический линейный замедлитель, диффузор, коллимирующая линза, светоделительная поляризационная система и четыре фотоприемника. Ориентация ахроматически-линейного замедлителя изменяется для последовательного введения набора известных состояний поляризации, которые описываются вектором Стокса Sq , во входное отверстие поляриметра. Вектор Стокса входного излучения определяется с помощью уравнения [8]:

$$Sq = [1(\theta) + (\theta) \cdot \cos \cos \delta \sin \sin(\theta) \cdot \cos \cos(\theta) \cdot (1 - \cos \cos \delta) \sin \sin(\theta) \cdot \sin \sin \delta] \cdot i_{0,x,y}. \quad (1)$$

При разных ориентациях ахроматического замедлителя на выходе каждого приемного канала получают набор из N -изображений. Матричное уравнение для каждого пикселя изображения с координатами (x,y) , связывающее параметры известных состояний поляризации входного излучения с измеренной облученностью области пикселя (x,y) на выходе матричного фотоприемника выглядит так [8]:

$$\begin{bmatrix} c_{00} & c_{01} & c_{10} & c_{11} & c_{20} & c_{21} & c_{30} & c_{31} & \dots & c_{0N} & \dots & c_{1N} & \dots & \dots & c_{2N} & c_{3N} \end{bmatrix}_{x,y} = \\ \begin{bmatrix} a_{00} & a_{01} & a_{10} & a_{11} & a_{20} & a_{21} & a_{30} & a_{31} & a_{02} & a_{03} & a_{12} & a_{13} & a_{22} & a_{23} & a_{32} & a_{33} \end{bmatrix}_{x,y} \cdot \\ [1 \ 1 \ s_{1,0} \ s_{1,1} \ s_{2,0} \ s_{3,0} \ s_{2,1} \ s_{3,1} \ 1 \ \dots \ s_{1,2} \ \dots \ s_{2,2} \ s_{3,2} \ \dots \ \dots \ 1 \ s_{1,N} \ s_{2,N} \ s_{3,N}] \cdot i_{0,x,y}. \quad (2)$$

Уравнение (2) может быть представлено в упрощенном виде [8]:

$$C_{x,y} = A_{x,y} \cdot S \cdot i_{0,x,y}, \quad (3)$$

где $A_{x,y}$ – приборная матрица системы, в параксиальном приближении одинаковая для всех пикселей изображений;

S – полный вектор Стокса входного излучения;

$i_{0,x,y}$ – интенсивность излучения, падающего на заданную область пикселя с координатами (x,y) .

Приборная матрица системы затем рассчитывается через уравнение (4) или с помощью определения матриц Мюллера для каждого канала четырехканальной системы и подстановки первых строк матриц Мюллера в приборную матрицу системы [9]:

$$A_{x,y} = (S^T \cdot S)^{-1} \cdot S^T \cdot C_{x,y} \cdot i_{0,x,y}. \quad (4)$$

Таким образом, представленная методика калибровки поляриметра позволяет получать итоговую реальную приборную матрицу системы для всех пикселей изображения, в которой будут учтены погрешности измерения параметров вектора Стокса и степени поляризации. Расчеты, выполненные по данной методике, показали, что указанная погрешность составляет несколько процентов [8]. Для достижения такой точности необходимо проводить измерения с использованием известных состояний поляризации, а калибровочные векторы Стокса должны быть выбраны так, чтобы приборная матрица поляриметра была не сингулярной.

Вторая методика калибровки Стокс-поляриметров заключается в использовании произвольного набора линейно независимых калибровочных векторов. Для этого вводится вектор для соответствующих токовых откликов фотодетекторов $I_k = (I_{0k}, I_{1k}, I_{2k}, I_{3k})^T$. Тогда классическое матричное уравнение, описывающее работу Стокс-поляриметра ($I = A \cdot S$), переписывается в следующем виде:

$$I_k = A \cdot S_k, \quad (k = 0, \dots, 3), \quad (5)$$

где $A = [a_{ij}]$ – приборная матрица системы.

Задача калибровки состоит в выражении приборной матрицы A и ее относительной погрешности $\delta A = \frac{\|\Delta A\|}{\|A\|}$ в евклидовой норме через векторы S , I , и их погрешности δS , δI . Существует два варианта решения этой задачи.

Первый вариант заключается в изначальном формировании блочно-диагональной матрицы шестнадцатого порядка для вектора Стокса s и преобразовании приборной (a) и токовой (i) матриц в блочно-диагональные матрицы шестнадцатого порядка, которые содержат

столбцы, составленные из строк, соответственно, для матриц A и I . Затем классическое матричное уравнение Стокс-поляриметра представляется как линейная комбинация уравнений и определяется условие для расчета числа обусловленности:

$$\delta a \leq \frac{\nu_s}{1-\nu_s \cdot \delta s} (\delta s + \delta i), \quad (6)$$

где ν_s – число обусловленности матрицы s .

Тогда, определяя относительную погрешность приборной матрицы $\delta a = \delta A$, относительную погрешность вектора Стокса $\delta s = \delta S$, относительную погрешность характеристической матрицы поляриметра $\delta i = \delta I$, число обусловленности можно будет рассчитать следующим образом:

$$\nu_s = \|s\| \cdot \|s^{-1}\| = 2\|S\| \cdot 2\|S^{-1}\| = 4\nu_s. \quad (7)$$

Подстановка полученных значений в формулу (6) позволяет рассчитать искомую относительную погрешность для приборной матрицы:

$$\delta A \leq \frac{4 \cdot \nu_s}{1-\nu_s \cdot \delta S} \cdot (\delta I + \delta S), \quad (8)$$

Второй вариант заключается в записи классического уравнения для Стокс-поляриметра в возмущенном виде с учетом абсолютных погрешностей для токовой матрицы I , приборной матрицы поляриметра A и вектора Стокса S . Предполагая, что итоговая матрица, которая рассчитывается как результат суммы вектора Стокса и его абсолютной погрешности, является невырожденной, можно использовать свойства нормы, описанные в работе [10], для расчета относительной погрешности приборной матрицы поляриметра:

$$\delta A \leq \frac{\nu_s}{1-\nu_s \cdot \delta S} \cdot (\delta I + \delta S). \quad (9)$$

Таким образом, оценка неравенства для относительной погрешности приборной матрицы Стокс-поляриметра (9), даже если считать $\delta I = \delta S = 0$, в четыре раза ниже оценки неравенства (6). Причина этого заключается в том, что оценка неравенства (6) получается при условии независимости возмущений всех элементов матрицы S , что не соответствует реальным условиям работы системы.

Третья методика калибровки Стокс-поляриметров, которая заключается в использовании определенного набора линейно независимых калибровочных векторов для определения числа обусловленности матрицы S . Методика основывается на рассмотрении нормированного набора калибровочных векторов Стокса:

$$S_k = (1, S_{1k}S_{2k}, S_{3k}), S_{1k} + S_{2k} + S_{3k} = 1. \quad (10)$$

Однако, такой ортогональной системы не существует, в связи с тем, что определитель, составленный из векторов S_k , по модулю равен шестикратному объему трехмерной пирамиды, которая вписана в трехмерную сферу единичного радиуса с центром в начале координат [11]. В связи с этим проводится корректировка координат вершин тетраэдра таким образом, чтобы вершина тетраэдра находилась на оси OZ , в точке с координатами $(0,0,1)$. Тогда матрица S принимает следующий вид [11]:

$$S(\varphi) = \begin{bmatrix} 1 & 1 & 0 & \frac{2\sqrt{2}}{3} \cos \cos \varphi & 0 & 1 & \frac{2\sqrt{2}}{3} \sin \sin \varphi & -\frac{1}{3} & 1 & 1 & \frac{2\sqrt{2}}{3} \cos \cos \left(\varphi + \frac{2\pi}{3}\right) & \frac{2\sqrt{2}}{3} \cos \cos \left(\varphi + \frac{4\pi}{3}\right) \\ \frac{2\sqrt{2}}{3} \sin \sin \left(\varphi + \frac{2\pi}{3}\right) & -\frac{1}{3} & \frac{2\sqrt{2}}{3} \sin \sin \left(\varphi + \frac{4\pi}{3}\right) & -\frac{1}{3} \end{bmatrix}. \quad (11)$$

где φ – угол между положительным направлением оси OX и радиус-вектором.

Независимо от угла φ для евклидовой нормы матрицы $S(\varphi)$ и ее определителя, получаем следующее уравнение:

$$\|S(\varphi)\| = 2\sqrt{2}, |S(\varphi)| = \frac{16\sqrt{3}}{9}. \quad (12)$$

Обратная матрица $S(\varphi)^{-1}$ может быть найдена аналитически. Тогда число обусловленности ν_s матрицы $S(\varphi)$ не зависит от угла между положительным направлением оси OX и радиус-вектором и равно: $\nu_s = \|S\| \cdot \|S^{-1}\| = 2\sqrt{5}$.

Проведенный сравнительный анализ методик калибровки Стокс-поляриметров позволил установить, что первая методика калибровки реализована на использовании известного набора калибровочных векторов и основана на определении реальной приборной матрицы системы, в которой будут учтены все относительные погрешности для всех пикселей изображения. Погрешность методики составляет несколько процентов. Вторая и третья методики калибровки с произвольным и с определенным набором линейно независимых калибровочных векторов позволили установить значительно более точную оценку относительных погрешностей при нахождении элементов приборной матрицы Стокс-поляриметров путем учета погрешностей элементов калибровочных векторов и получении на их основе токовых откликов. При этом вторая методика калибровки Стокс-поляриметров позволяет добиться в четыре раза большей точности определения погрешностей, обусловленных не идеальностью оптических компонентов схемы, по сравнению с третьей методикой.

На основании обзора рассмотренных методик калибровки Стокс-поляриметров установлено, что для портативных поляризационных систем, направленных на исследование структуры диффузно рассеивающих оптически неоднородных сред в медицинской сфере, наиболее целесообразно использовать вторую методику калибровки, которая основана на определении реальной приборной матрицы системы и использовании произвольного набора линейно независимых калибровочных векторов. Данная методика позволяет добиться в четыре раза большей точности определения погрешностей, которые существенно влияют на процесс измерения, и которые необходимо учитывать при работе поляриметра в биологических и медицинских сферах для получения достоверных результатов.

Литература

1. Banik S., Melanthota S.K., Arbaaz, Vaz J.M., Kadambalithaya V.M., Hussain I., Dutta S., Mazumder N. Recent trends in smartphone-based detection for biomedical applications: a review // *Analytical and Bioanalytical Chemistry*, 413. – 2021. – Pp. 2389–2406.
2. Yang K., Peretz-Soroka H., Liu Y., Lin F. Novel developments in mobile sensing in resource-limited settings: a brief review // *Analytical Methods*. – 2016. – №8(36). – Pp. 6591–6601.
3. Hunt B., Ruiz A.J., Pogue B.W. Smartphone-based imaging systems for medical applications: a critical review // *J. of Biomedical Optics*. – 2021. – №26(4). 040902. DOI: 10.1117/1.JBO.26.4.040902.
4. Alawsi T., Al-Bawi Z. A review of smartphone point-of-care adapter design // *Engineering Reports*. 1(e12039). – 2019. – Pp. 1–30.
5. Quesada-Gonzalez D., Merkoci A. Mobile phone-based biosensing: an emerging «diagnostic and communication» technology // *Biosensors and Bioelectronics*. – 2017. – №92. – Pp. 549–562.
6. Zhu H., Sencan I., Wong J., Dimitrov S., Tseng D., Nagashima K., Ozcan A. Cost-effective and rapid blood analysis on a cell-probe // *Lab on a Chip*. – 2013. – №13(7). – Pp. 1282–1288.
7. Kolokolova L., Hough J., Levasseur-Regourd A. *Polarimetry of stars and planetary systems* // Cambridge University Press. – 2015. – №503. DOI: 10.1017/CBO9781107358249.
8. Pezzaniti J.L., Chenault D., Roche M., Reinhardt J., Schultz H. Wave slope measurement using imaging polarimetry // *Proc. of SPIE Ocean Sensing and Monitoring*. – 7317(73170B). – 2009. DOI: 10.1117/12.819031.
9. Azzam R.M.A. Division-of-amplitude photopolarimeter (DOAP) for the simultaneous measurement of all four stokes parameters of light // *Optical Acta: International Journal of Optics*. – 1982. – №29. – Pp. 685–689.
10. Воеводин В.В., Кузнецов Ю.А. Матрицы и вычисления М.: Наука. – 1984. – 320 с.
11. Vidmachenko A., Ivanov Y.S., Siniavskiy I.I., Sergeev A.V. Development and manufacturing of panoramic Stokes polarimeter using the polarization films // 15 Ukrainian Conference on Space Research held. – 2015. – Vol.15. – 79 p. DOI: 10.13140/RG.2.1.2801.9926.

УДК 007.51

ИССЛЕДОВАНИЯ АЛГОРИТМОВ УПРАВЛЕНИЯ ОСВЕЩЕНИЕМ МНОГОПОЛЬЗОВАТЕЛЬСКОГО ПРОСТРАНСТВА НА ОСНОВЕ АНАЛИЗА ПРОСТРАНСТВЕННОГО РАСПРЕДЕЛЕНИЯ ЯРКОСТИ

Чеботарев С.В.¹ (студент), Брусницын А.В.¹ (студент)

Научный руководитель – канд. архитектуры Быстрянцева Н.В.¹

Научный консультант – Филиппов И.М.¹

¹*Университет ИТМО*

e-mail: arseny.brusnitsyn@gmail.com

В работе рассматривается метод разработки алгоритма управления освещением на основе расчета яркости, а не освещенности, как предложено в большинстве существующих аналогов таких систем. Эти аналоги так же не затрагивают взаимодействие пользователей с системой освещения в многопользовательском пространстве.

Ключевые слова: освещение, освещенность, визуальный комфорт, яркость, система управления освещением, алгоритм оптимизации.

Исследования зависимости состояния человека от освещения принесло свои плоды еще в прошлом веке. Известно, что правильно подобранное освещение способно повысить работоспособность, а также снизить уровень стресса человека в течение рабочего дня. Существующие нормы освещения не являются универсальными, и дело не только в физических различиях между отдельными людьми, но и в том, что оценка освещения происходит не совсем корректным образом. Как правило основным критерием для оценки визуального комфорта человека, да и освещения в целом, является освещенность. Согласно ряду исследований, для оптимальной оценки уровня комфорта человека следует исследовать пространственное распределение яркости, которое включает в себя блики, яркость окон и соотношение яркости в поле зрения (контрастность) [1]. Влияние этих раздражающих зрение факторов учитывает обобщенный показатель дискомфорта UGR. Показатель UGR в мировой практике является составной частью стандартов внутреннего освещения рабочих мест, таких как CIE S 008 (CIE 2001), европейский стандарт EN12464-1 (CEN, 2011), китайский стандарт GB50034 (САВР 2014) и российский ГОСТ Р 55710-2013. Для измерения этого показателя необходимо измерять не освещенность рабочих зон, а пространственное распределение яркости. Это подтверждается исследованиями, в которых для достижения визуального комфорта вертикальная освещенность превосходит все общепринятые нормы [2].

На сегодняшний день, существующими фотокамерами возможно измерить значения яркости примерно от 10 до 50 000 кд/м² при использовании модуля камеры в качестве датчика абсолютной яркости. Помимо этого, возможно провести попиксельный анализ данных, качественных и количественных аспектов освещенности окружающей среды, получить распределение яркости, коэффициенты яркости, яркости адаптации и оценки бликов [3]. Самое главное, различные протоколы управления позволяют управлять освещением и жалюзи, в зависимости от полученных данных, а также пожеланий каждого отдельно взятого человека.

Все становится значительно сложнее, когда речь идет о многопользовательской среде. Системы освещения в таких средах реализованы в виде обычной сетки светильников, которая

часто не соответствует расположению мебели. Один осветительный прибор может освещать несколько рабочих зон, а в зависимости от расположения внутри данной среды, рабочая зона может быть освещена по-разному [4]. Используя существующие системы оценки яркости, системы управления, а также анализируя состояние каждого пользователя среды, можно создать автоматизированный алгоритм, который будет подстраиваться под пользователя, вне зависимости от времени года, естественного света и даже места посадки. Такой алгоритм позволит не только добиться максимальной работоспособности каждого отдельно взятого человека, но и предсказать и свести к минимуму вероятность конфликта между пользователями. Современные протоколы управления и автоматизации освещения, используются в основном для управления всеми источниками в помещении вместе. В таком случае не берется во внимание комфорт каждого отдельно взятого пользователя, в результате чего в многопользовательских пространствах некоторые люди, больше подверженные стрессу, утомляются сильнее других.

Разработка самообучающегося алгоритма позволит адаптировать освещение к таким изменениям, как:

- перемещение пользователей в пространстве;
- смена предпочтений пользователей относительно освещения;
- динамика естественного освещения в течение дня/года;
- перестановка мебели, перепланировка помещения;
- обеспечение комфортного микроклимата в помещении;
- снижение общего энергопотребления указанных систем.

Все еще не существует однозначного определения, как следует измерять значение яркости в помещениях с естественным освещением. Также в исследованиях часто приводятся упрощенные варианты систем либо с одним датчиком, либо с одним источником света и одним пользователем. С другой стороны, возможно использование аппаратных решений, которые были взяты в других работах. Также ценны используемые в материалах методики анализа изображений, а также корреляции субъективных оценок пользователей с измерениями освещенности и яркости рабочих мест.

План разработки:

1. Анализ факторов, влияющих на субъективное восприятие световой среды в помещении.
2. Анализ методов измерения и оценки светотехнических параметров, определяющих восприятие световой среды в помещении.
3. Анализ существующих методов управления освещением в офисных пространствах.
4. Анализ методов программной реализации алгоритма управления освещением.
5. Разработка алгоритма управления освещением многопользовательского офисного пространства.

Таким образом, становится понятна необходимость в аппаратной реализации системы управления освещением на основе разработанного алгоритма. Результаты исследования позволят получить больше информации о связи между механизмами восприятия света человеком и пространственным распределением яркости, что позволит более эффективно работать с различными индексами зрительного комфорта и использовать эти данные для создания алгоритма управления системой освещения на основе пространственного распределения яркости. Однако, понимание связи между световосприятием и значениями яркости позволит работать не только с алгоритмами, но и поможет специалистам с их разработками во многих других областях.

Литература

1. Wienold J., Christoffersen J. Evaluation methods and development of a new glare prediction model for daylight environments with the use of CCD cameras // *Energy and Buildings*. – 2006. – №38. – Pp. 743–757.
2. Van Den Wymelenberg, K., Inanici, M. A Critical Investigation of Common Lighting Design Metrics for Predicting Human Visual Comfort in Offices with Daylight // *LEUKOS: The Journal*

of the Illuminating Engineering Society of North America. – 2014. – №10:3. – Pp. 145-164. DOI: 10.1080/15502724.2014.881720.

3. Boyce P.R. Age, illuminance, visual performance and preference// Lighting Research & Technology. – 1973. – №5,3. – Pp. 125–144.
4. Despenic M., Chraibi S., Lashina T., Rosemann A. Lighting preference profiles of users in an open office environment //Building and Environment. – 2017. – № 116. – Pp. 33–43.

ТРАНСЛЯЦИОННЫЕ ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ

УДК 004.91

ИСПОЛЬЗОВАНИЕ КОРПУСНОЙ ЛИНГВИСТИКИ В СОЦИАЛЬНО-ПОЛИТИЧЕСКИХ ИССЛЕДОВАНИЯХ

Абрамовский И.С.¹ (студент)

Научный руководитель – канд. полит. наук Чугунов А.В.¹

¹*Университет ИТМО*

e-mail: abrams25@rambler.ru

Работа выполнена в рамках проекта НИР №621304 «Разработка сервиса тематической кластеризации корпуса текстов «Развитие цифрового государственного управления в Российской Федерации» на основе машинного обучения».

Статья посвящена обзору исследований социально-политической направленности, использующих методы корпусной лингвистики. В статье приводится анализ существующих в России корпусов текстов. Приводятся примеры работ, посвященных исследованию социально-политических процессов. Автор делает вывод об ограниченном использовании авторами инструментов корпусной лингвистики, и использовании уже готовых корпусов.

Ключевые слова: корпус текстов, корпусная лингвистика.

Вопросы анализа слабоформализованных текстов: новости, публикации и комментарии в социальных сетях рассматриваются в рамках корпусной лингвистики.

Корпусная лингвистика – актуальное направление современных прикладных исследований. Корпусная лингвистика – это раздел компьютерной лингвистики, занимающийся разработкой общих принципов построения и использования лингвистических корпусов с использованием компьютерных технологий [1].

В.П. Захаров определяет корпус текстов, как унифицированный, структурированный и размеченный массив языковых (речевых) данных в электронном виде, предназначенный для определенных филологических и, более широко, гуманитарных изысканий. А.М. Амиева и В.В. Филимонов выделяют несколько разработанных в России корпусов текстов:

1. Машинный фонд русского языка.
2. Национальный корпус русского языка.
3. Лаборатория общей и компьютерной лексикологии и лексикографии филологического факультета МГУ имени М. В. Ломоносова.
4. Хельсинский аннотированный корпус русских текстов.
5. Корпус русского литературного языка.
6. Корпус парафраз русского языка [2].

Ниже приведен сравнительный анализ корпусов текстов, представленных авторами (таблица). Крупнейшим является Национальный корпус русского языка, насчитывающий 283 млн. словоупотреблений. При этом отметим, что в сравнении с другими национальными корпусами российский уступает по числу словоупотреблений. Так, например, Национальный корпус Польского языка включает в себя 1 млрд. словоупотреблений, что значительно превышает число словоупотреблений в российском.

Необходимо отметить, что в основной массе публикаций по использованию корпуса текстов в социальных исследованиях затрагиваются темы: политического дискурса, категории социально-политической жизни, анализ выступлений политических акторов.

Таблица. Сравнение корпусов текстов

Наименование корпуса	Число словоупотреблений	Входящие текстовые источники
Машинный фонд русского языка	100 млн.	Художественная литература XIX–XX вв., словари, фольклор, историческая проза XIX–XX вв., газетный корпус 1990-е гг.
Национальный корпус русского языка	283 млн.	Художественная литература, мемуары, публицистика, научная, религиозная литература, повседневная печатная продукция, корпус устных текстов (публичные выступления и интервью)
Лаборатория общей и компьютерной лексикологии и лексикографии МГУ	11 млн.	Газетный корпус
Хельсинский аннотированный корпус	100 тыс.	Журнал «Итоги»
Корпус русского литературного языка	1 млн.	Художественная литература, публицистика, научная литература, драматические произведения (все тексты с 1950-х гг. до настоящего времени)

Если рассмотреть имеющиеся публикации второй группы, стоит отметить работу «Проблемы эффективности сентимент-анализа русскоязычных текстов в социальных сетях». В данной работе основное внимание уделено технической стороне проведения сентимент-анализа, или анализа тональности текста. Авторы выделяют несколько групп методов автоматического сентимент-анализа: методы, основанные на правилах и словарях, методы, использующие машинное обучение, а также гибридные методы, где лексические ресурсы (тональные словари) комбинируются с техниками машинного обучения.

Использование корпусов текстов актуально в рамках социально-политических исследований. Так, О.О. Борискина и К.М. Шилихина [3] описывают возможности корпусных исследований на материале корпуса выступлений Х. Клинтон и Д. Трампа во время предвыборной кампании 2016 года. Также рассматривается пример анализа понятия «politic» через материал корпуса NOW (News on the Web).

Авторы проводят достаточно подробный анализ употребления понятия «politic», уделяя внимание использованию его в различных вариантах английского языка: британский (GB), ирландский (IE), американский (US), канадский (CA) и т.д. Также авторы глубоко погрузились в предмет контекста употребления понятия и в различия его употребления в различных географических регионах. Результаты данного анализа могут быть использованы в дальнейших разработках.

Вторая часть, посвященная анализу предвыборных выступлений Х. Клинтон и Д. Трампа, также является серьезным испытанием, результаты которого могут быть интересны лицам, участвующих в избирательных кампаниях, работающих в сфере рекламы и маркетинга.

Исследования, проведенные в работе, хорошо себя показывают с точки зрения демонстрации возможности корпусной лингвистики в рамках междисциплинарных прикладных исследований. Однако, разбора одного понятия видится нам недостаточным. Все же стоит отметить, что продемонстрированные исследования имеют большой потенциал в рамках развития данного направления в политических исследованиях.

М.В. Рубцова и Е.А. Васильева [4] разработали методику анализа Национального корпуса текстов с использованием качественных и количественных методов. Для примера они продемонстрировали исследование корпуса текстов для категории «доверие» в российских СМИ в период с 2010 по 2011 года. В результате авторы выявили, что данная категория используется в контексте политических институтов в значении «доверие, как фактор организации политического процесса».

Выбор авторами исследования соединения в методике исследования качественных и количественных методов заслуживает позитивной оценки, поскольку позволяет комплексно изучить предмет исследования. Также стоит отметить трехуровневый разбор категории. Авторы выбрали три уровня контекста: институциональный, лексический, оценочный. Это позволило рассмотреть явление с различных точек зрения в рамках социального взаимодействия.

Text data mining используется И.А. Быковым [5] при анализе коммуникации в публичных платформах. В работе «Цифровые платформы государственного управления в системе национальных публичных коммуникаций» автор рассматривает коммуникационные площадки в разрезе публичного управления. В качестве объекта исследования были выделены публичные коммуникации в двух цифровых платформах «Наш Петербург» и «Народная экспертиза» (порталы Санкт-Петербурга и Ленинградской области). Метод интеллектуального анализа текстов используется в среде статистических вычислений.

Представленные примеры исследований отличаются использованием уже разработанных корпусов текстов. Авторы исследований обращаются к готовым разработкам, что упрощает сбор эмпирического материала, однако, снижает возможности для выбора текстовых источников. Среди них выделяется проект МГУ «Историко-политическое измерение корпуса текстов русской художественной литературы второй половины XIX в.» [6], где авторы сформировали собственный корпус текстов. Однако, примеров таких исследований еще немного.

Таким образом, корпусные исследования также применяются в области социально-политических наук, и являются источником данных о развитии общества, социально-политического дискурса.

Литература

1. Райскина В.А., Дубнякова О.А. Современные методы корпусной лингвистики при анализе текста (на примере корпуса BFM) // Актуальные вопросы современной науки. – 2015. – №40. – С. 146–154.
2. Амиева А.М., Филимонов В.В., Сергеев А.П. Инструменты корпусной лингвистики // 2я Международная конференция студентов, аспирантов и молодых ученых "Информационные технологии, телекоммуникации и системы управления": сборник докладов. Екатеринбург. УрФУ. – 2016. – С. 251–260.
3. Борискина О.О., Шилихина К.М. Корпусные исследования политического дискурса в лингвистике // Политическая наука. – 2017. – № 2. – С. 30–53
4. Рубцова М.В., Васильева Е.А. "Доверие": концептуализация и операционализация понятия в корпусной лингвистике // Социологические исследования. – 2016. – № 1(381). – С. 58–65.
5. Быков И.А. Цифровые платформы государственного управления в системе национальных публичных коммуникаций // Журнал политических исследований. 2021. – Т. 5. – № 2. – С. 45–56. DOI 10.12737/2587-6295-2021-5-2-45-56.
6. Историко-политическое измерение корпуса текстов русской художественной литературы второй половины XIX в. // Факультет Политологии МГУ им. М.В. Ломоносова. [Электронный ресурс]. – Режим доступа: <https://polit.msu.ru/science/research/18-311-00344/> (дата обращения: 17.12.2021).

УДК 004.93

АЛГОРИТМЫ ИДЕНТИФИКАЦИИ ЦИФРОВЫХ КОММУНИКАТИВНЫХ АКТОВ

Алейников С.А.¹ (студент)

Научный руководитель – д-р техн. наук, доцент Басов О.О.¹

¹Университет ИТМО

e-mail: aleynikov.sergey.a@gmail.com

Работа выполнена в рамках темы НИР №220037 «Разработка программных модулей интеллектуальной обработки данных мониторинга наземных объектов с БЛА».

В статье был проведен сравнительный анализ и разбор вербальных и невербальных методов и способов коммуникации. Были проанализированы невербальные методы и способы коммуникации с целью определить оптимальные способы для определения цифровых коммуникативных актов. Также в ходе исследования были рассмотрены существующие методы идентификации невербальных эмоциональных и произведен их сравнительный анализ с целью определения оптимальных методов.

Ключевые слова: невербальная коммуникация, вербальная коммуникация, цифровые коммуникативные акты, методы идентификации коммуникативных актов, интеллектуальная интерпретация невербальных сигналов.

В ходе межличностной коммуникации люди неосознанно используют многомодальную систему: вербальные и невербальные каналы. По вербальному каналу коммуникации осуществляется передача сообщений от одного участника к другому. Однако, регулирование процесса коммуникации, установление психологического контакта между его участниками, выражение эмоций и дополнение передаваемой информации деталями осуществляется через невербальный канал связи. Средства невербальной коммуникации могут быть переданы через акустический (паралингвистический и экстралингвистический), визуальный (передача информации посредством мимики, жестов, поз, взгляда), символический (внешний вид), тактильный (осязательный) и ольфакторный (обоняние) каналы. Фактически до 93% всей информации человек передает и получает в невербальном виде. Однако, распределение контроля между невербальным и вербальными каналами неравное. Человек может контролировать только вербальный канал. За передачу невербальных сигналов отвечает подсознание. Иными словами, собеседники могут контролировать свои слова, но полностью искренние чувства и намерения скрыть не могут. Психологи придерживаются мнения, что залогом успешного общения является верная интерпретация невербальных сигналов.

В условиях наблюдаемого сегодня роста объема и интенсивности использования технических средств межличностной коммуникации, в частности систем видеоконференцсвязи (Zoom и др.) возможности анализа невербальных сигналов оказываются ограниченными, что существенно снижает эффективность целенаправленного (например, деловых переговоров, собеседований, устных экзаменов) коммуникативного взаимодействия (в дальнейшем такое взаимодействие будем называть цифровым). Это происходит как в силу объективных причин, снижающих качество передаваемой информации (нечеткая картинка, ограниченный сектор обзора камеры, посторонние шумы в акустическом канале, искажения в условиях низкоскоростных каналов связи), так и по причине высокой степени неопределенности относительно её невербального компонента (в виду отсутствия устойчивых механизмов его

извлечения, обработки и интерпретации). Как следствие, преодоление этой проблемы требует развития когнитивных информационных технологий для восполнения невербальной составляющей поведения собеседника в ходе цифрового коммуникативного взаимодействия, исходя из объективных материалов (данных, получаемым по невербальным каналам коммуникации и характеризующих текущее состояние пользователя). Это позволит, в свою очередь, создавать инструменты поддержки принятия решений при проведении различного рода собеседований, позволяющие одному из участников цифрового коммуникативного взаимодействия оценивать истинность (ложность) высказываний (ответов на вопросы) другого собеседника, что уже сейчас востребовано в различных областях, связанных с необходимостью определения достоверности сообщаемой информации (собеседование при трудоустройстве, устный онлайн-экзамен и пр.).

Создание таких инструментов на основе уже существующих массовых интеллектуальных технологий (анализ голоса, изображений, биометрических факторов и пр.) затруднительно в силу следующих причин:

1. Отсутствия адекватных моделей коммуникативного взаимодействия, основанных на идентификации состояния собеседника и истинности его высказываний.
2. Отсутствия формального описания зависимостей между характеристиками вербальных и невербальных сигналов коммуникации и состояниями (намерениями) ее участников.
3. Ограниченного количества релевантных баз данных, содержащих синхронизированные размеченные вербальные и невербальные сигналы акустического и визуального каналов коммуникации.

Для решения перечисленных выше задач необходимо развить математический и алгоритмический аппарат извлечения и обработки таких данных и формализации получаемых знаний в форме вероятностных моделей, учитывая следующие особенности:

- трудноформализуемое описание взаимосвязи между невербальными сигналами и особенностями личности собеседников, их психофизиологическим и эмоциональным состоянием, истинностью высказываний и намерений;
- слабоструктурированный характер данных, определяющих поведение человека в процессе межличностной коммуникации;
- априорное предположение о наличии пропусков и некорректных значений в данных невербальных каналов коммуникации, связанных как с неполной наблюдаемостью среды (например, при использовании только одного коммуникационного канала), так и с неподдающимся контролю поведением коммуникантов.

Совокупно учитывая методическую, алгоритмическую и вычислительную сложность проблемы, ее возможное решение видится на стыке трех направлений исследований:

- а) развития методов извлечения невербальных сигналов акустического и визуального каналов коммуникации,
- б) разработке алгоритмов идентификации вероятностных моделей цифрового межличностного взаимодействия на основе интеллектуальной интерпретации невербальных сигналов,
- в) создания информационной технологии, обеспечивающей эффективную реализацию методов вероятностной оценки истинности высказываний собеседника в процессе цифрового коммуникативного взаимодействия в различных сферах деятельности человека, направленных на установление достоверности сообщаемой информации.

Модели цифровой коммуникации разделяются на следующие виды:

- синхронная коммуникация;
- асинхронная коммуникация;
- вербальная коммуникация;
- невербальная коммуникация.

Методы цифровой коммуникации:

- система видеоконференцсвязи;
- система аудиоконференцсвязи;
- текстовый мессенджер;

- система управления проектами;
- видеозаписи.

Синхронная коммуникация подразумевает передачу информации одним из участников коммуникативного акта и одновременной интерпретации передаваемой информации другими участниками коммуникативного акта.

Преимущества синхронной коммуникации:

- высокая скорость получения информации;
- быстрое получение обратной связи.

Минусы синхронной коммуникации:

- необходимость присутствия всех участников;
- периодически возникают сложности и недопонимания в коммуникации;
- невозможность параллельного вещания нескольких участников.

Асинхронная коммуникация подразумевает передачу информации одним из участников таким способом, чтобы остальные участники могли ознакомиться с информацией в удобное время после передачи.

Преимущества асинхронной коммуникации:

- отсутствие затрат на синхронизацию;
- больше времени на интерпретацию информации, по сравнению с синхронной коммуникацией.

Минусы синхронной коммуникации:

- меньше вербальной информации;
- потеря контекста коммуникации;
- получение обратной связи отложено во времени, относительно момента передачи информации.

Методы цифровой коммуникации:

- система видеоконференцсвязи;
- система аудиоконференцсвязи;
- текстовый мессенджер;
- система управления проектами;
- видеозаписи.

Разберем подробнее методы цифровой коммуникации:

1. Вербальная коммуникация – коммуникативный акт, в котором кроме непосредственно информации об объекте коммуникации, не передается дополнительной информации.
2. Невербальная коммуникация – коммуникативный акт, в котором в дополнение к информации об объекте коммуникаций, передается невербальная информация.
3. Система видеоконференцсвязи – набор технических решений для проведения цифрового коммуникативного акта, который содержит как видеoinформацию, так и аудио. Данная система позволяет осуществлять передачу вербальной и невербальной информации, рассчитана на синхронное взаимодействие.
4. Система аудиоконференцсвязи – набор технических решений для проведения цифрового коммуникативного акта, который содержит только аудиoinформацию. Данная система позволяет осуществлять передачу вербальной и невербальной информации, рассчитана на синхронное взаимодействие. Невербальной информации передается в значительно меньших объемах, чем посредством систем видеоконференцсвязи.
5. Текстовый мессенджер – набор технических решений для проведения цифрового коммуникативного акта, который содержит только текстовую информацию. Данная система позволяет осуществлять передачу только вербальной, рассчитана на синхронное и асинхронное взаимодействие. Невербальной информации практически не передается.
6. Система управления проектами – набор технических решений для проведения цифрового коммуникативного акта, который содержит только текстовую информацию. Данная система позволяет осуществлять передачу только вербальной, рассчитана на асинхронное взаимодействие. Невербальной информации не практически не передается.

7. Видеозаписи – заранее записанное выступление участника цифрового коммуникативного акту, рассчитанное на тиражирование среди участников асинхронного коммуникативного акта. Данный способ позволяет осуществлять передачу вербальной и невербальной информации, но без обратной связи.

Описанные выше модели и методы коммуникативного акта представлены в таблице 1, при помощи которой может быть осуществлены описанные выше методы коммуникативного акта.

Таблица 1. Сравнение моделей и методов коммуникации

	Система видеоконференцсвязи	Система аудиоконференцсвязи	Текстовый мессенджер	Система управления проектами	Видеозаписи
Асинхронная коммуникация	–	–	+	+	+
Синхронная коммуникация	+	+	+	–	–
Вербальная коммуникация	+	+	+	+	+
Невербальная коммуникация	+	+	–	–	+

Среди перечисленных выше методов и моделей наиболее подходящими для дальнейшей работы подходит синхронный метод цифровой коммуникации посредством видеоконференцсвязи, т.к. в данном способе содержится больше всего невербальной информации, которую можно использовать для устранения недостатков выбранного способа.

Далее рассмотрим методы и работы для идентификации невербальных сигналов. Список отобранных методов идентификации невербальных сигналов:

1. Deep multi-path convolutional neural network joint with salient region attention for facial expression recognition [1].
2. Spatio-Temporal Convolutional Features with Nested LSTM for Facial Expression Recognition [2].
3. Unsupervised adaptation of a person-specific manifold of facial expressions [3].
4. Peak-Piloted Deep Network for Facial Expression Recognition [4].
5. Deep generative-contrastive networks for facial expression recognition [5].

Таблица 2. Сравнение методов идентификации невербальных сигналов

№ работы	Результаты и наборы данных	Возможные улучшения
1	CK+ (49,10%), JAFFE (43,38%), TFEID (52,75%), BAUM-2i (67,92%), FER2013 (81,53%), SFEW (42,30%)	Совместить две модели (SERD и MPVS-Net) в одну
2	CK+ (99,8%), MMI (84,53%), Oulu-CASIA (93,35%), BP4D (0,58% fl)	–
3	–	Построение метрик и оценка качества модели Адаптация метода для временного распознавания эмоций Тестирование с другими наборами эмоций
4	CK+ (99,3%), Oulu-CASIA (84,59%)	–
5	CK+ (97,28%), MMI (81,53%), Oulu-CASIA (86,11%)	–

В рассматриваемых методах идентификации невербальных сигналов использовались следующие наборы данных (табл. 2):

- CK+ – набор данных состоит из 593 видеозаписей от 123 различных людей в возрасте от 18 до 50 лет;

- JAFFE – набор данных состоит из 213 изображений различных выражений лица;
- TFEID – набор лицевых эмоций;
- BAUM-2i – набор 1047 видео, на которых запечатлены лицевые эмоции;
- FER2013 – примерно 30 000 изображений лица с различными выражениями, размер которых ограничен 48×48 пикселей;
- SFEW (Static Facial Expressions in the Wild) – набор данных для распознавания выражений лица;
- Oulu-CASIA – набор данных выражений лица состоит из шести выражений (удивление, счастье, печаль, гнев, страх и отвращение) от 80 человек в возрасте от 23 до 58 лет;
- MMI – база данных выражений лица;
- BP4D – база данных 3D-видео о спонтанных выражениях лица от разнообразных групп людей;
- MUG – база данных выражений лица;
- RECOLA – база данных выражений лица;
- FAST (Faceted Application of Subject Terminology) – база данных выражений лица.

В таблице 2 приведен разбор результатов рассматриваемых методов и предложены возможные улучшения. В первом столбце указан номер рассматриваемой работы, во втором – использованные наборы данных и оценка точности, в третьем – возможные улучшения рассмотренных работ. Первая работа показала точность немногим лучше случайного ответа, поэтому предлагается совместить две модели для повышения точности. Также возможные улучшения предложены для третьей работы, т.к. для нее отсутствуют замеры точности. Все остальные работы показали высокую точность, поэтому возможных улучшений для них в рамках текущих исследований не предложено.

Из таблицы можно увидеть, что ни одна работа не имеет оценки скорости работы представленных методов и используются различные наборы данных для оценки качества методов. В качестве дальнейшей работы предполагается провести сравнение скорости работы методов из рассмотренных выше статей и качества на одинаковых наборах данных.

Литература

1. Siyue Xie, Haifeng Hu, Yongbo W. Deep multi-path convolutional neural network joint with salient region attention for facial expression recognition // *Pattern Recognition*. – 2019. – V.92. – Pp. 177–191.
2. Zhenbo Yu, Guangcan Liu, Qingshan Liu, Jiankang Deng. Spatio-Temporal Convolutional Features with Nested LSTM for Facial Expression Recognition // *Neurocomputing*. – 2018. – V.317. – Pp. 50–57.
3. Raphaël Weber, Vincent Barrielle, Catherine Soladié, Renaud Segurier. Unsupervised adaptation of a person-specific manifold of facial expressions // *IEEE Transactions on Affective Computing*. – 2018. – V.11. – Pp. 419–423.
4. Rouast P.V., Adam M.T.P., Chiong R. Deep Learning for Human Affect Recognition: Insights and New Developments // *IEEE Transactions on Affective Computing*. – 2019. – V.12. – Pp. 524–543.
5. Kim Y., Yoo B., Kwak Y., Choi C., Kim J. Deep generative-contrastive networks for facial expression recognition // *arXiv*. – 2019. arXiv:1703.07140.

УДК 004.514

ЭФФЕКТИВНОСТЬ РАСПОЛОЖЕНИЯ БУРГЕР-МЕНЮ НА МОБИЛЬНЫХ ВЕРСИЯХ САЙТОВ

Баглык Е.В.¹ (студент)

Научный руководитель – канд. филос. наук Смолин А.А.¹

¹Университет ИТМО

e-mail: liza.baglyk@mail.ru

В работе исследуется является ли расположение бургер-меню в левом нижнем углу более эффективным и удобным для пользователей, чем остальные расположения на устройствах, размер которых превышает 4,7 дюйма. Исследование проводилось с помощью тестирования четырех прототипов сайта, различающихся только расположением меню. В результате было выявлено, что расположение бургер-меню в левом нижнем наиболее эффективно и предпочтительно пользователями, чем другие варианты расположения.

Ключевые слова: бургер-меню, навигация, мобильные версии сайтов, гамбургер-меню, мобильные устройства.

С момента появления графического интерфейса он постепенно менялся, что в первую очередь было обусловлено развитием технологий и трансформацией процесса проектирования. Сейчас 4,54 миллиарда человек в мире пользуются интернетом на мобильных устройствах, из них 118 миллионов живут в России. На мобильные телефоны приходится больше половины времени, проведенного в интернете — 50,1% [1], причем большинство людей предпочитают использовать смартфоны с диагональю экрана от 4,7 дюйма. Согласно данным компании Devise Atlas на 2019 год, около 95% пользователей в 36 странах мира пользуются устройствами такого размера, в России их – 91% [2].

Увеличение устройств должно было повлечь за собой изменение интерфейса и расположения элементов в нём. Знание характеристик экрана смартфона позволяет производителям лучше адаптировать и формировать интерфейс для пользователей. Но, помимо этого, необходимо учитывать ещё множество факторов, например, каким именно образом люди взаимодействуют с устройством. В контексте использования смартфонов необходимо знать, как именно люди его держат, когда прикасаются к экрану. В исследовании Стива Хубера, проведённом во время наблюдения за 1333 людьми, выяснилось, что во время нажатия на экран, чаще всего люди держат смартфон тремя способами: одной рукой, поддерживая второй рукой и двумя руками. 49% задействовало только одну руку при поддержке телефона и касании экрана, а большой палец – 85% [3]. Большой палец может дотянуться до большей части поверхности экрана, но только на трети экрана можно взаимодействовать с устройством комфортно: внизу с противоположной от пальца стороны. Эта территория получила название: зона большого пальца [4]. Для большего удобства пользователей и лучшей эффективности самые главные и часто используемые элементы интерфейса стоит размещать именно в этой зоне.

Одним из таких элементов является иконка бургер-меню, которая скрывает под собой навигацию. Обычно она располагается в левом верхнем углу экрана. Бургер-меню быстро стало один из наиболее часто используемых шаблонов навигации на мобильных устройствах, а также на веб-сайтах. Если учесть, что большинство людей пользуется устройствами с большой диагональю экрана, и посмотреть на зону большого пальца, можно увидеть, что

данный элемент находится в самом неудачном и неудобном для пользователя месте. Необходимо адаптировать привычный дизайн под изменяющийся размер мобильных устройств и найти более выгодное расположение такого элемента, как бургер-меню.

В первую очередь для определения уровня эффективности человеко-компьютерного взаимодействия при различных расположениях бургер-меню было спроектировано четыре макета интерфейса мобильного приложения, они представлены на рисунке. В первом бургер-меню располагается в левом верхнем углу, во втором – в левом нижнем, в третьем – в правом нижнем и в четвертом – в правом верхнем углу. Дизайн всех остальных элементов в обоих вариантах интерфейса абсолютно идентичен. Проектирование осуществлялось с использованием программы Figma.



Рисунок. Макеты, используемые в исследовании

Генеральная совокупность: мужчины и женщины, правши, в возрасте от 20 до 26 лет, пользующиеся мобильным телефоном, диагональ экрана которого равна или больше 4,7 дюйма. Для тестирования всех макетов и получения максимально достоверных данных, респонденты были поделены на 4 независимых выборки.

Гипотезы эксперимента:

1. Расположение бургер-меню в нижнем левом углу экрана более эффективно при взаимодействии пользователей с интерфейсом. Нулевая гипотеза – расположение бургер-меню не влияет на эффективность взаимодействия пользователей с интерфейсом. Данная гипотеза является главной.
2. Существует корреляционная связь между оценкой удобства использования интерфейса и временем, потраченным на взаимодействие с ним. Нулевая гипотеза – количество времени, потраченного на взаимодействие с интерфейсом, не влияет на оценку удобства использования данного интерфейса.

Перед проведением основного эксперимента был проведен пилотный, который позволил скорректировать процесс проведения и рассчитать размер выборки.

Необходимый размер выборки для первой гипотезы (табл. 1) был рассчитан с помощью точного теста Фишера по формуле:

$$n = z^2 \times \frac{p_0 \times (1 - p_0)}{2 \times (p_1 - p_0)^2},$$

где $z = z_\alpha + z_\beta$, т.к. проверка левосторонняя.

Таблица 1. Значения для расчета контрольной выборки для первой гипотезы

Желаемая статистическая значимость	Желаемая мощность	z	p_0	p_1	n	Размер выборки
0,01	0,9	-3,6	0,45	0,18	21	42

Для второй гипотезы (табл. 2) при расчете необходимого размера выборки также использовалось преобразование Фишера с использованием коэффициента корреляции Спирмена.

Таблица 2. Значения для расчета контрольной выборки для второй гипотезы

Желаемая статистическая значимость	Желаемая мощность	Кол-во сторон проверки	z	r _s	Размер выборки
0,05	0,75	2	-2,63	-0,378	49

Опираясь на расчеты, в основном эксперименте участвовал 91 респондент, выборка была разделена на четыре независимые группы. Процесс тестирования осуществлялся с мобильного устройства. Вначале, для того, чтобы обеспечить репрезентативность выборки, респонденты отвечали на вопрос: какая диагональ у экрана их смартфона. Если она была меньше 4,7 дюйма, то опрос заканчивался, если больше или равна – участник переходил к следующему этапу. Для того, чтобы выборка была случайной, респонденты самостоятельно выбирали макет, на котором они будут проходить тестирование, выбрав в опросе цифру от одного до четырех.

Далее им нужно было заполнить графы пол и возраст и ответить на вопрос, как они чаще всего держат мобильное устройство при взаимодействии с ним. Ответ на данный вопрос включает в себя следующие значения: ”в одной руке, используя только большой палец”, ”двумя руками, используя большие пальцы обеих рук”, ”двумя руками, используя большой палец одной руки”, ”в одной руке, используя один палец другой руки (не большой)”, ”двумя руками, используя один палец (не большой)”. После этого респонденты получали задание и переходили по ссылке на макет, чтобы его выполнить. Им необходимо было выполнить целевое действие - перейти на страницу ”Связаться с нами”. Время выполнения данного действия засекалось для оценки эффективности тестируемого расположения бургер-меню. После тестирования макета участники возвращались к опросу, чтобы измерить субъективные показатели эффективности. Им нужно было оценить удобство использования интерфейса по шкале от 1 до 5, где 1 – ”очень неудобно”, 5 – ”очень удобно”. и указать, пришлось ли им менять положение руки во время нажатия на иконку бургер-меню.

По окончании эксперимента, в первой группе, тестирующей макет, в котором иконка бургер-меню располагалась в левом верхнем углу, оказалось 25 респондентов, во второй группе (иконка в левом нижнем углу) – 19, в третьей группе (иконка в правом нижнем углу) – 24, в четвертой (иконка в правом верхнем углу) – 22. Усредненные данные, полученные в основном эксперименте, можно увидеть в таблице 3.

Таблица 3. Данные, полученные в основном эксперименте

	Возраст	Диагональ (дюймы)	Время	Удобство	Изменение положения руки
Макет 1	22,36	5,492	6,867	2,8	1,28
Макет 2	22,53	5,526	5,509	4,37	1,79
Макет 3	21,75	5,47	6,41	3,8	1,5
Макет 4	21,9	5,6	6,24	3,69	1,37

Для расчета статистической значимости первой гипотезы использовался точный тест Фишера по формулам:

$$p = \frac{(A+B)! \times (C+D)! \times (A+C)! \times (B+D)!}{A!B!C!D!N!}.$$

Для подтверждения гипотезы данные макетов 1, 3 и 4 сравнивались с макетом 2. Успехом считался результат, меньший, чем медиана, рассчитанная по результатам пилотного эксперимента (6,37). Зависимой переменной выступает время выполнения целевого действия на макете сайта. Данные, необходимые для расчёта представлены в таблице 4.

Получен статистически значимый результат (таблица 5), значит поставленная гипотеза подтверждена. По полученным данным видно, что во время тестирования второго макета успешных результатов намного больше, чем при тестировании первого. Значит, расположение бургер-меню в нижнем левом углу сокращает время, необходимое на нажатие иконки и увеличивает эффективность человеко-компьютерного взаимодействия.

Таблица 4. Данные, необходимые для расчета статистической значимости для первой гипотезы

	Кол-во успехов	Кол-во респондентов
Макет 1	6	25
Макет 2	17	19
Макет 3	15	24
Макет 4	16	22

Таблица 5. Статистическая значимость для первой гипотезы

Макеты	Кол-во успехов
1 и 2	0,000015
3 и 2	0,006
4 и 2	0,0009

Для расчёта статистической значимости для второй гипотезы использовалось преобразования Фишера от коэффициента Спирмена по формулам:

$$r_s = \frac{\sum_{i=1}^n (R[x_i] - \bar{R})(R[y_i] - \bar{R})}{\sqrt{\sum_{i=1}^n (R[x_i] - \bar{R})^2 \times \sum_{i=1}^n (R[y_i] - \bar{R})^2}}$$

$$z = \frac{1}{2} \ln\left(\frac{1+r_s}{1-r_s}\right) \times \sqrt{\frac{n-3}{1.06}}, \text{ учитывая поправку на непрерывность: } \frac{6}{n \times (n^2 - 1)}.$$

Для второй гипотезы результат получился статистически значим, значит существует корреляция между измеряемыми показателями (табл. 6).

Таблица 6. Данные, используемые для расчета статистической значимости для второй гипотезы

r_s	z	n	p-value
-0,58	4,1	91	0,0002

При анализе данных, полученных во время проведения эксперимента, подтвердились все поставленные гипотезы. Следовательно расположение бургер-меню в нижнем левом углу оказалось эффективнее и удобнее для пользователей, чем его привычное расположение в верхнем левом углу. На оценку респондентами удобства использования интерфейса влияют такие показатели, как время, потраченное на совершение необходимого действия и необходимость при этом менять положение руки

Выводы данного исследования могут быть применены в других исследованиях, связанных с интерфейсами мобильных устройств. Также они могут быть полезны для разработчиков при создании интерфейсов. Исследование влияния различных расположений бургер-меню на эффективность человеко-компьютерного взаимодействия необходимо, поскольку позволит научным образом обосновать выдвигаемые гипотезы, а также поможет разработчикам сделать их мобильные приложения более эффективными.

Литература

1. Kemp S. Digital 2020: global digital overview, 2020. [Электронный ресурс]. – Режим доступа: <https://datareportal.com/reports/digital-2020-global-digital-overview>. (дата обращения: 17.10.2021).
2. The most popular smartphones in 2019 – DeviceAtlas. [Электронный ресурс]. – Режим доступа: <https://deviceatlas.com/blog/most-popular-smartphones>. (дата обращения: 03.12.2021).
3. Hooper S. How Do Users Really Hold Mobile Devices? [Электронный ресурс]. – Режим доступа: <https://www.uxmatters.com/mt/archives/2013/02/how-do-users-really-hold-mobile-devices.php>. (дата обращения: 23.11.2021).
4. Berkman E., Hooper S. Designing Mobile Interfaces: Patterns for Interaction Design. – O'Reilly Media 2011. – 584 p.

УДК 004.8

ПРИМЕНЕНИЕ ТЕХНОЛОГИЙ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА В ТЕМАТИЧЕСКОЙ КЛАСТЕРИЗАЦИИ ТЕКСТОВ

Беген П.Н.¹ (аспирант)

Научный руководитель – канд. техн. наук Митягин С.А.¹

¹Университет ИТМО

e-mail: petyabegen@mail.ru

Работа выполнена в рамках темы НИР №621304 «Разработка сервиса тематической кластеризации корпуса текстов «Развитие цифрового государственного управления в Российской Федерации» на основе машинного обучения».

В работе рассмотрено применение подхода к задаче распознавания именованных сущностей в текстах по тематике цифровой трансформации общества на основе современных технологий искусственного интеллекта. В работе представлен обзор технологий, а также подготовлен список сущностей, готовых для выявления в текстах о цифровой трансформации в России.

Ключевые слова: цифровое развитие, тематическое моделирование, кластеризация, искусственный интеллект, распознавание именованных сущностей, трансформеры.

Работа направлена на решение проблемы кластеризации и классификации специализированных текстов по тематике цифровой трансформации общества. На данный момент существует проблема распознавания сущностей в таких текстах, поскольку имеющиеся наборы размеченных русскоязычных корпусов текстов лежат вне области цифровой трансформации и существенно сужают корректность и точность тематической кластеризации.

В исследовании используется массив текстов новостных сообщений «Развитие цифрового государственного управления в Российской Федерации», который начал формироваться в процессе выпуска еженедельного информационного дайджеста Центра технологий электронного правительства с 2010 г. На данный момент массив насчитывает более 8,5 тыс. единиц текста.

Основными задачами проекта тематической кластеризации массива текстов о цифровом развитии в Российской Федерации являются:

1. Разработка комплексного метода тематической кластеризации корпуса текстов о цифровой трансформации на основе машинного обучения.
2. Представление сравнительного анализа работы алгоритма на основе новой комбинации методов по сравнению с традиционными.

Данная работа направлена на выявление основных методов и технологий искусственного интеллекта, используемых исследователями в подобного рода задачах. Под технологиями в контексте данной работы понимаются методы машинного обучения, методы глубокого обучения (deep learning) и методы обработки естественного языка (Natural Language Processing – NLP).

Далее представлены основные методы и подходы:

1. Трансформеры (Transformers).
 2. MCDM (multiple-criteria decision-making).
 3. DANP (DEMATEL-based analytic network process).
 4. LDA (Latent Dirichlet allocation).
 5. Random Forest.
-

6. Recurrent Neural Networks (RNN), LSTM.

7. Word2Vec.

Данные методы используются в широком спектре задач, например:

- Topic modeling (Тематическое моделирование);
- Named Entity Recognition (Распознавание именованных сущностей);
- Sentiment analysis (Анализ настроений/тональности);
- Clustering task (Задача кластеризации);
- Classification task (Задача классификации);
- Trending (построение трендов);
- Forecasting (прогнозирование).

В данной работе основной фокус направлен на тематику распознавания именованных сущностей (сокр. NER), являющихся специфичными в области цифрового развития государственного управления.

В текстах о цифровой трансформации общества были выделены следующие сущности для распознавания:

- уровень;
- отрасль;
- технология;
- статус;
- организация;
- участники;
- география.

Для сущностей Уровень, Отрасль и Статус задан определенный список значений, например для сущности «Уровень» в списке выделены уровни:

- 1) муниципальный (т.е. местного значения);
- 2) региональный;
- 3) федеральный;
- 4) международный.

При изучении и анализе массивов текстов в рамках задачи NER исследователи [1, 2] зачастую используют рекуррентные нейронные сети с LSTM-блоком либо традиционные методы машинного обучения, как Наивный Байесовский классификатор, SVM, деревья решений и др. С 2017 года начался период активного перехода к усовершенствованным моделям типа трансформер (Transformer) [3], которые начали показывать более качественные результаты распознавания и прогнозирования.

Архитектура трансформера состоит из кодировщика и декодировщика.

Каждый кодировщик состоит из механизма self-attention, самовнимания, (вход из предыдущего слоя) и нейронной сети с прямой связью (вход из механизма самовнимания).

Каждый декодировщик состоит из механизма самовнимания (вход из предыдущего слоя), механизма внимания к результатам кодирования (вход из механизма самовнимания и кодировщика) и нейронной сети с прямой связью (вход из механизма внимания). Популярные модели (BERT, GPT, Electra и др.) часто используют в машинном переводе в сервисах Яндекс и Google, в голосовых помощниках, обработке изображений и видео.

Помимо использования трансформеров применяется подход MCDM (multiple-criteria decision-making) для выбора наилучших стратегий для выявления сущностей.

MCDM – активная область исследований с 1960-х годов: разработано более 20 методов и множество различных модификаций, включая Fuzzy (нечёткие) модификации.

MCDM изучает проблему выбора наилучшей альтернативы среди ряда конкурирующих альтернатив.

Существует ряд принципов данного подхода:

- каждая альтернатива описывается в терминах набора оценочных критериев;
- часть критериев противоречива;

- несоразмерность: критерии могут иметь разные единицы измерения;
- критерии связаны с важными весами.

Проблема MCDM обычно решается в два этапа:

1. Рейтинг, то есть агрегирование значений критериев для каждой альтернативы.
2. Ранжирование или упорядочение между альтернативами в отношении глобальной степени удовлетворения.

Математические модели на основе данного подхода активно применяют в экономике и финансах, логистике, бизнесе, управлении рисками. Однако, использование в интеллектуальных системах, в частности в работе с текстами, появилось совсем недавно, в 2020 г. [4].

В качестве метода из подхода MCDM используется DANP, который является гибридной формой методов DEMATEL и ANP.

DEMATEL – это комплексный метод построения и анализа структурной модели, включающей причинно-следственные связи между сложными факторами. Основано на теории графов и дискретной математике. Этот метод может анализировать связи между факторами и преобразовывать эти связи в сравнительную структуру, которая может продемонстрировать группы причинных и группы следственных факторов и показать численную силу влияния.

Метод аналитического сетевого процесса (ANP) является более общей формой процесса аналитической иерархии (AHP), используемого при анализе многокритериальных решений. Группа экспертов представляет проблему в виде сетевой структуры (сети или орграфа), где однородные критерии сгруппированы в группы таким образом, что их можно оценить по влиянию относительных критериев в смежных группах.

DANP (DEMATEL+ANP) упрощает аналитическую процедуру методов, основанных на ANP, и учитывает все взаимосвязи влияния при выводе весов влияния.

Предполагается, что использование трансформеров с использованием математических подходов на основе MCDM (DANP) в текстах о цифровой трансформации общества позволит выявлять сущности с гораздо более высокой точностью, по сравнению с традиционными методами машинного обучения и классическими подходами.

Литература

1. Dieng A.B., Ruiz F.J.R., Blei D.M. Topic Modeling in Embedding Spaces // Transactions of the Association for Computational Linguistics. – 2020. – №8. – Pp. 439–453. DOI: 10.1162/tacl_a_00325.
2. Yufan Jiang, Chi Hu, Tong Xiao, Chunliang Zhang, Jingbo Zhu. Improved differentiable architecture search for language modeling and named entity recognition // Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP). – 2019. – Pp. 3585–3590. DOI: 10.18653/v1/D19-1367.
3. Wolf T., Debut L., Sanh V. etc. Transformers: State-of-the-Art Natural Language Processing // Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing: System Demonstrations. – 2020. – Pp. 38–45.
4. Huang C-Y, Yang C-L, Hsiao Y-H. A Novel Framework for Mining Social Media Data Based on Text Mining, Topic Modeling, Random Forest, and DANP Methods // Mathematics. – 2021. – 9(17):2041. – DOI: 10.3390/math9172041.

УДК 004.6

ИССЛЕДОВАНИЕ ТРЕБОВАНИЙ, ПРЕДЪЯВЛЯЕМЫХ МЕТОДАМИ АНАЛИЗА БОЛЬШИХ ДАННЫХ, К ОРГАНИЗАЦИИ ВЫЧИСЛИТЕЛЬНОЙ СЕТИ ИНФОКОММУНИКАЦИОННОЙ СИСТЕМЫ

Бейлин М.Т.¹ (студент)

Научный руководитель – канд. техн. наук, доцент Осипов Н.А.¹

¹*Университет ИТМО*

e-mail: icevenom98@gmail.com, nikita@ifmo.spb.ru

В работе проведен анализ наиболее распространенных методов обработки больших данных, приведены преимущества и недостатки каждого метода, а также исследованы требования к вычислительным сетям для реализации методов. Приведены показатели оценки эффективности вычислительных сетей. На основе существующих алгоритмов анализа данных предложены и реализованы в виде прототипов модели вычислительных сетей для анализа больших данных.

Ключевые слова: вычислительная система, большие данные, анализ данных, распределенные вычисления, эффективность обработки.

Большие данные – это понятие, без которого в настоящее время не обходится практически ни одна отрасль науки или промышленности. И на увеличение эффективности анализа больших данных направлены многие исследования и разработки. Все работы в этом направлении можно разделить на два направления: разработка более совершенных устройств и вычислительных мощностей, которые позволят физически ускорить вычисления над большими данными, и усовершенствование самих алгоритмов обработки больших данных, подразумевающее сокращения числа или упрощение операций обработки данных.

Для анализа методов кластеризации информации разработаны метрики точности для измерения качества кластеризации, которые делятся на внутренние, сравнивающие данные внутри кластеров и различия кластеров между собой, и внешние, определяющие соответствие практических кластеров эталонным, то есть заранее известным. В зависимости от устанавливаемой задачи по анализу больших данных может применяться тот или иной из описанных методов кластеризации [1, 2].

При проведении анализа эффективности алгоритмов и вычислительных сетей в первую очередь внимание обращается на размер оперативной памяти на ядро и на скорость работы процессора [3].

Основными требованиями к программам и системам подобного рода являются: возможность масштабируемости инфраструктуры; оперативность обработки данных; обеспечение требований безопасности. Соответственно, при разработке вычислительной сети для анализа больших данных необходимо учитывать физические и логические требования к структуре сети, которые гарантируют оптимальную работу алгоритмов.

При определении средств анализа данных необходимо четко представлять цели задач и ожидаемые результаты. «Алгоритм обработки должен иметь минимальную зависимость от объема поступающих данных и должен позволять оперативно обрабатывать постоянно увеличивающиеся объемы данных, либо должен подразумевать возможность масштабирования в зависимости от желания пользователя» [4].

Многие исследования направлены на сравнение алгоритмов анализа больших данных. Особенно внимание акцентируется на такие алгоритмы как GraphLab, Galois, GraphX. В таблице 1 для них приводятся сведения о требовательности и гибкости при реализации.

Таблица 1. Сравнение программных средств работы с большими сетями [5, 6]

Программа	GraphLab	Galois	GraphX
Каркас	MapReduce	MapReduce	ApacheSpark
Производительность	Высокая при числе вершин до 100000, далее значительно падает	Высокая при числе вершин до 100000, далее значительно падает	Высокая
Требовательность к ресурсам	Не требовательна	Не требовательна	Очень требовательна
Реализация алгоритмов	Гибкая настройка, поддержка разных языков программирования	Использует только собственные структуры данных	Гибкая настройка, поддержка разных языков программирования

Отказоустойчивой архитектуре графов (ОАГ) накладывает ограничения на структуру вычислительной сети. Согласно ей, узлы сети располагаются в определенной последовательности и связаны не только с общим хранилищем данных, но и друг с другом, что позволяет производить вычисления с учетом зависимостей данных. Такие алгоритмы позволяют создавать отказоустойчивую архитектуру, но требовательную к характеристикам скорости чтения-записи данных в устройствах сети. Иной подход – потоковые архитектуры графов (ПАГ) – призваны повысить скорость обработки данных за счет обработки в режиме реального времени. Однако, такой подход затрудняет обеспечение отказоустойчивости, так как из-за жесткой связи узлов друг с другом выбывание из строя как минимум одного из них приводит к нарушению всех вычислений, которые необходимо выполнять сначала. Таким образом, данный алгоритм предназначен только для надёжных сетей [7].

Построение граф-ориентированной сети подразумевает разработку матрицы переходов, пример которой приведен в таблице 2.

Таблица 2. Пример взаимосвязей узлов сети

	Узел №1	Узел №2	...	Узел №N
Узел №1		+(0.1)		+(0.05)
Узел №2			+(0.9)	
...				
Узел №N	+(1)	+(0.8)	+(0.01)	

На основе приведенной информации была построена таблице 3, содержащая сравнение методов анализа больших данных, в которую были сведены данные об их применимости и о требованиях к вычислительным сетям.

Для оценки эффективности вычислительных сетей применяются такие параметры, как:

1. Общая производительность смоделированной вычислительной сети.
2. Время проведения анализа в разной конфигурации модели – относительный показатель, при котором сравнивается время анализа при разных конфигурациях модели с целью выявления оптимальной.
3. Уровень использования компонентов сети на протяжении выполнения эксперимента.
4. Отношение используемых и простаивающих вычислительных узлов.
5. Отношение выполняющихся и ожидающих заданий – рассчитывается на основе объемов данных, находящихся внутри сети во время их обработки и во время ожидания очереди на обработку [8].

В рамках текущей работы под общей производительностью понимается характеристика скорости обработки данных сетью. Для расчета скорости необходимо замерить время эксперимента и объем данных, проходящий через вычислительную сеть, что легко реализуемо как при проверке моделей, так и на практике.

Таблица 3. Основные алгоритмы анализа данных

Название	Описание	Применимость	Требования к сети
Hadoop (MapReduce)	Классический алгоритм, состоящий из кластеризации и параллельной обработки данных	Для больших объемов данных, не обладающих высокой связностью	Кластера вычислительных машин, объединенных в единую сеть
Spark	Распределенных вычисления, использующие общий распределенный набор данных	Для ограниченных наборов данных, не обладающих высокой связностью	Кластер вычислительных машин, связанных быстрой сетью, большие объемы оперативной памяти
ОАГ GraphLab	Архитектура графов, обработка данных со сложными вычислительными зависимостями	Обработка данных, требующих совокупного анализа	Высокая скорость чтения и записи информации, надежный доступ к общей системе хранения данных
ПАГ	Архитектуры графов, работа с зависимыми данными в режиме потоковой обработки	Обработка больших данных с высокой скоростью и в больших объемах	Надежная вычислительная сеть, надежная одноранговая сеть связи, высокая отказоустойчивость

Второстепенным показателем эффективности будет принят коэффициент использования или простоя ресурсов. Такой показатель позволит оценить насколько целесообразно усложнение и расширение вычислительной сети при определенных нагрузках на нее. В общем виде коэффициент использования ресурсов рассчитывается как среднее арифметическое коэффициентов использования каждого из узлов сети по формуле (1).

$$K = \frac{\sum k_n}{n}, \quad (1)$$

где k_n – коэффициент использования n -го узла,
 n – число узлов сети.

k_n в зависимости от ситуации рассчитывается как временной или ресурсный показатель по формуле (2).

$$k_n(V) = \frac{V_n}{V_{nmax}}; k_n(t) = \frac{t_n}{t_{max}}, \quad (2)$$

где V_n – используемый объем ресурсов узла,
 V_{nmax} – весь ресурс узла,
 t_n – время работы узла,
 t_{max} – все время работы сети.

Для реализации выбранных алгоритмов используется ПО AnyLogic, общий вид модели представлен на рисунке 1.

Вводными параметрами являются:

1. Количество узлов.
2. Объем данных для анализа.
3. Параметры узлов:
 - скорость обработки;
 - объем загружаемых данных;
 - надежность узла (1-р, где р – вероятность отказа).

Главными выходными значениями являются скорость обработки и объем обработанных данных на момент времени t . Расчет времени обработки данных t проводится по формуле (3).

$$t = \left(\frac{D_n}{V_n}\right), \quad (3)$$

где D_n – объем данных на одном узле,
 V_n – скорость обработки на одном узле.

В общем случае объем данных на одном узле D_n рассчитывается по формуле (4).

$$D_n = \frac{D}{N} \quad (4)$$

где D – значение общего объема данных в сети.

Однако, возможны доработки алгоритма, позволяющие нагружать узлы пропорционально их производительности. Скорость обработки V_n зависит от технических характеристик каждого из узлов.

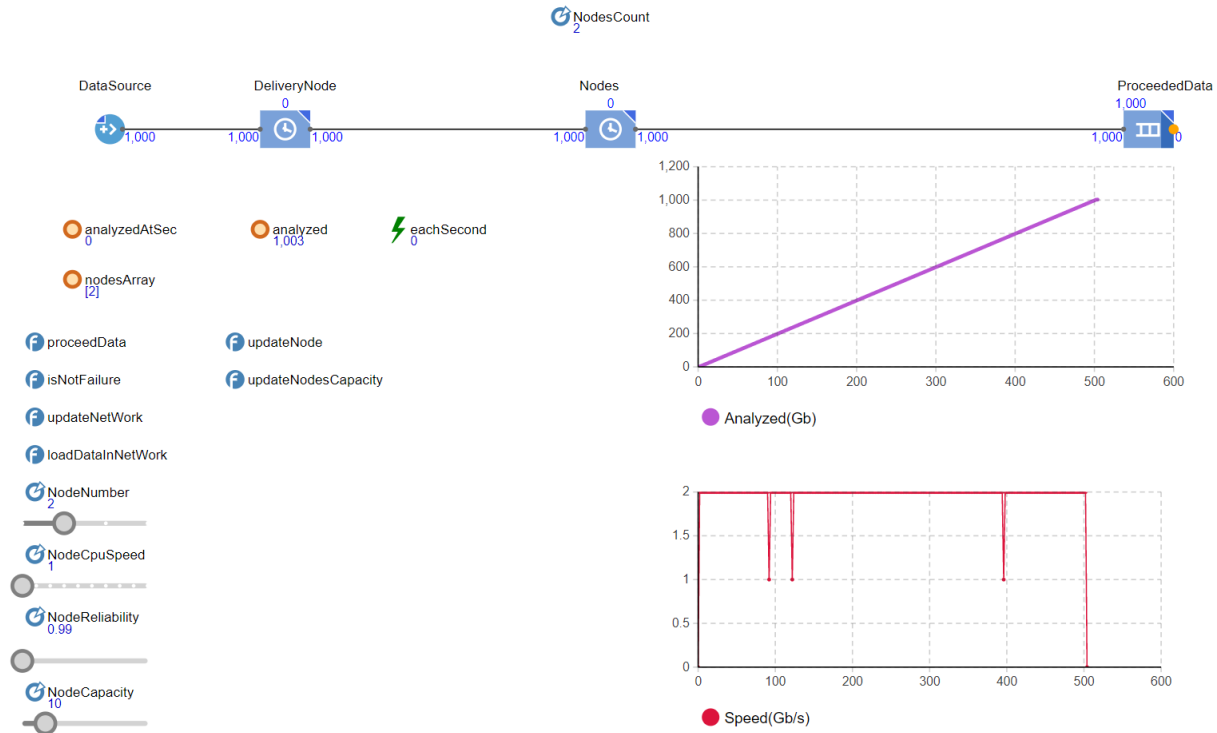


Рис. 1. Модель вычислительной сети, реализующей алгоритм MapReduce

Для моделирования режима ApacheSpark при расчете времени работы каждого узла t_n вводится параметр объема оперативной памяти M . Таким образом, время обработки данных рассчитывается по формуле (5).

$$\{M \geq D \Rightarrow t_n = \frac{D}{V} \quad M < D \Rightarrow t_n = \frac{M}{V} + t_n(D = D - M), \quad (5)$$

где t_n – время работы узла,

V – скорость обработки данных,

D – объем данных.

С целью обеспечить задержку всех агентов (пакетов данных) перед их распределением по узлам сети в начале каждого цикла обработки данных производится перерасчет задержки для всех агентов в очереди. Затем для каждого из узлов в цикле происходит расчет объема обработанных данных на основе введенных параметров.

Модель для граф-ориентированных алгоритмов отличается введенным в нее понятием матрицы переходов (таблица 2), которая моделируется двумерным массивом, где значения с индексом m, n – это значения перехода данных из вычислительного узла № m в узел № n . Данная матрица применяется при расчете объемов информации на каждом узле в текущий момент времени и позволяет моделировать поток данных в вычислительной сети в целом. Время работы узлов в таком случае рассчитываются по следующей формуле:

$$t_{n,i} = (t_{1...N,i-1}) + \frac{D_{n,i}}{V_n}, \quad (6)$$

где $t_{1...N,i-1}$ – время обработки данных на каждом предыдущем узле,

$D_{n,i}$ – объем данных на входе данного узла,

V_n – скорость обработки данных на узле.

Объем данных – это сумма всех данных, поступающих с предыдущих узлов на данной итерации.

Пример работы модели и процесса анализа эффективности граф-ориентированной сети (по показателю загруженности узлов) представлен на рисунке 2.

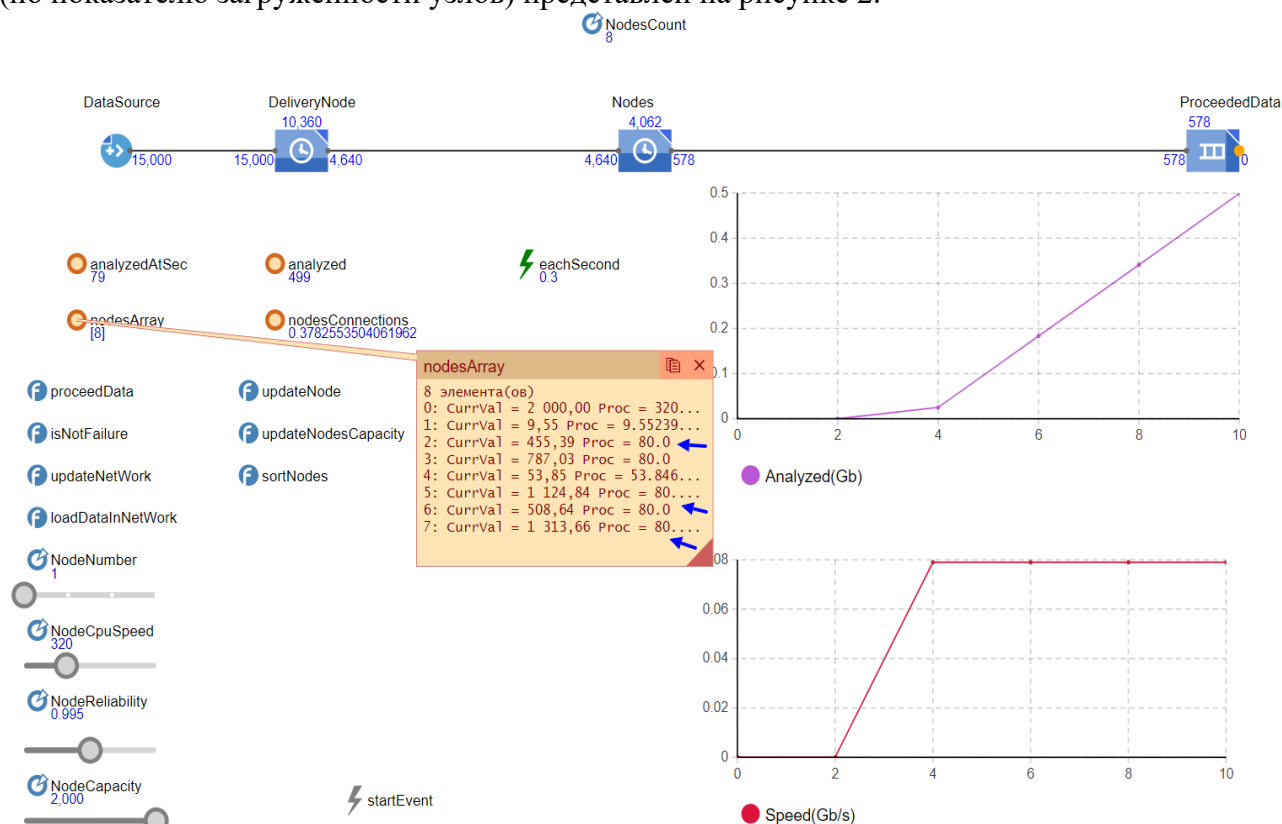


Рис. 2. Анализ работы модели граф-ориентированной сети

Описанные в работы модели сети могут быть оценены с точки зрения эффективности по перечисленным методам. Применение методов моделирования в совокупности с оценкой эффективности получаемых результатов позволят на практике добиться оптимальных конфигураций сети, используемой для решения поставленной задачи.

Дальнейшие исследования в этом направлении должны быть направлены на расширение описательной базы моделей с целью их более точной и гибкой настройки. Особенное внимание должно быть уделено проектированию граф-ориентированных архитектур вычислительных сетей, так как от алгоритма построения сети могут зависеть её качественные характеристики – надежность, безопасность и эффективность её работы.

Литература

1. Суй Д., Тянь Ю. Всесторонний обзор алгоритмов кластеризации // *Анналы науки о данных*. – 2015. – Т.2. – №2. – С. 165–193.
2. Лутов А., Хаяти М., Кудре-Мору П. Оценка точности алгоритмов кластеризации с перекрытием и множественным разрешением на больших наборах данных // *Международная конференция IEEE 2019 по большим данным и интеллектуальным вычислениям (BigComp)*. – 2019. – С. 1–8.
3. Кушнеревич А.Г. Оценка эффективности платформы распределенного анализа данных hadoop в сравнении с платформой SPARK для целей сетевой безопасности // *Региональная информатика и информационная безопасность*. – 2017. – С. 228–230.
4. Мартышкин А.И., Коростелев А.В. Алгоритм безопасной обработки больших данных облачными сервисами с помощью технологий Big Data // *XXI век: итоги прошлого и проблемы настоящего плюс*. – 2020. – Т.9. – №2. – С. 124–131.
5. Юдина М.Н., Дёмина А.Р. Пакеты анализа сетей на базе HADOOP // *Информационные технологии и автоматизация управления*. – 2017. – С. 144–148.

6. Пролетарский А.В. и др. Методы ситуационного анализа и графической визуализации потоков больших данных // Вестник Московского государственного технического университета им. НЭ Баумана. Серия «Приборостроение». – 2018. – №2(119). – С. 98-99.
7. 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.
8. Астриков Д.Ю., Кузьмин Д.А. Исследование методов оценки эффективности вычислительных систем на основе функциональных моделей // Сибирский журнал науки и технологий. – 2016. – Т.17. – №2. – С. 295–300.

УДК 004.02

ИНСТРУМЕНТЫ АВТОМАТИЧЕСКОГО ПОЛУЧЕНИЯ ОНТОЛОГИИ ИЗ РЕЛЯЦИОННЫХ БАЗ ДАННЫХ

Волков А.А.¹ (студент)

Научный руководитель – канд. техн. наук Тесля Н.Н.¹

¹Университет ИТМО

e-mail: artvolkov96@yandex.ru, nnteslya@itmo.ru

В работе рассмотрены существующие инструменты по получению графов знаний из реляционных баз данных. Для этого используются медиаторы, которые преобразуют данные из одного формата в другой по определенным правилам. Файл, содержащий набор таких правил, называют маппингом. Для маппинга онтологий и реляционных баз данных используется стандарт R2RML, который связывает RDF и SQL запросы. Получение и использование маппингов позволяет использовать источники разных форматов для получения графов знаний. В статье представлено описание наиболее популярных инструментов, приведено их сравнение и представлены примеры маппингов R2RML.

Ключевые слова: открытые данные, онтология, маппинг данных, реляционные базы данных, R2RML, SPARQL.

Чаще всего для хранения данных используются реляционные базы данных. Многие открытые источники данных также публикуют свои данные в формате CSV, которые, по сути, представляют собой выгрузку из реляционных баз данных. Большой процент использования реляционных баз данных не означает что данные хранятся оптимально своей структуре. В некоторых случаях лучше использовать другие способы хранения данных.

Это становится проблемой включения всех источников в единую базу знаний, так как требуется скопировать все данные из одного хранилища в другое. При этом если данные часто изменяются, то процесс копирования необходимо повторить. Такие данные нужно преобразовывать перед использованием в графе знаний, что приводит к дублированию информации и лишним операциям по преобразованию данных. Если хранить данные в исходном виде, используя медиатор для доступа к таким данным, то это позволит значительно сократить работу.

Для получения данных из реляционной базы данных с использованием онтологии был разработан стандарт R2RML (RDB to RDF Mapping Language) [1]. Маппинг – это определение соответствия данных между потенциально различными семантиками разных объектов(систем). Стандарт описывает каким образом описывать файл маппинга между онтологией и базой данных. Доступ к таким данным организован через язык запросов SQL. Напрямую невозможно работать с такими данным как с онтологией и трансформировать запросы SPARQL в SQL [2]. Для такой возможности требуется разработать медиатор, который будет заниматься трансляцией запросов из SPARQL в SQL. Также данные полученные в SQL необходимо преобразовывать в граф знаний. Пример R2RML маппинга представлен на рисунке 1.

Рассмотрим работу медиатора R2RML, который предоставляет точку запросов SPARQL для пользователей. Пользователь отправляет в медиатор SPARQL запрос. Медиатор, используя маппинг R2RML осуществляет трансляцию SPARQL запроса в SQL запрос и отправляет в реляционную базу данных. База данных возвращает ответ к медиатору в формате табличных данных. Медиатор, используя маппинг, трансформирует табличные данные в граф знаний и отправляет пользователю [3]. Структурная схема представлена на рисунке 2, в скобках указана последовательность работы системы.

```

Example R2RML mapping
<#TriplesMap1>
  rr:logicalTable [ rr:sqlQuery """
    SELECT EMP.*, (CASE JOB
      WHEN 'CLERK' THEN 'general-office'
      WHEN 'NIGHTGUARD' THEN 'security'
      WHEN 'ENGINEER' THEN 'engineering'
    END) ROLE FROM EMP
    """ ];
  rr:subjectMap [
    rr:template "http://data.example.com/employee/{EMPNO}";
  ];
  rr:predicateObjectMap [
    rr:predicate ex:role;
    rr:objectMap [ rr:template "http://data.example.com/roles/{ROLE}" ];
  ].

```

Рис. 1. Пример маппинга

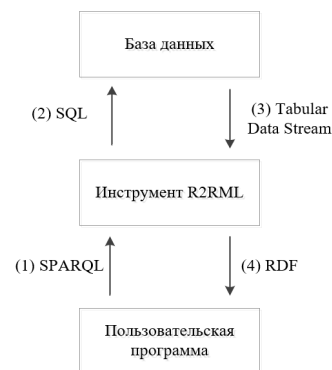


Рис. 2. Структурная схема медиатора R2RML

Тестирование работы инструментов заключалось в получении онтологии для данных карточек ДТП, представленных в реляционной базе данных PostgreSQL. Для этого были сформированы следующие задачи: рассмотреть стандарты для онтологии реляционных данных, провести обзор инструментов, получить онтологию для данных карточек ДТП и проверить работу SPARQL запросов.

Сложность задачи использования стандарта R2RML заключается в том, что маппинг данных необходимо написать вручную. Онтология также разрабатывается вручную специалистами предметной области. В случае реляционной базы данных информацию о семантике данных можно получить из схемы данных. Поэтому является логичным воспользоваться информацией в схеме и на основе ее создать, сгенерировать, маппинг.

В маппинге будет как описание данных в виде онтологии, так и запросы SQL для получения данных. Поэтому любой инструмент по работе с онтологией не подойдет, так как не все из них могут использовать внешние источники данных. Одним из решений проблемы является далее с помощью маппинга получить все данные в формате RDF и загрузить уже в специализированное хранилище. Но это не всегда возможно, особенно когда данные часто меняются и их большой объем.

Некоторые инструменты позволяют автоматически сгенерировать маппинг и онтологию, используя подключение к базе данных. Но воспользоваться существующими обзорами инструментов по работе с R2RML для выбора наиболее подходящего не всегда возможно, так как не всегда можно получить объективную картину. Это связано с тем, что такие обзоры могут приводиться для аргументирования разработки собственного инструмента.

Основной проблемой при использовании данных инструментов является их большое количество и одновременно с этим малый выбор подходящих вариантов. Нередко, когда разрабатывается инструмент для проведения конкретного исследования после чего разработка и поддержка инструмента прекращается. В качестве примера такого инструмента можно привести Map-On [4]. При его разработке приводились конкурентные преимущества и способ работы. Поскольку код открыт и распространяется на Github, то можно увидеть, что код не обновлялся уже три года. При этом документации, кроме как описания работы в статье, почти нет. Есть рабочий демо-сайт работы инструмента, но не получается зарегистрироваться и провести тестирование работы.

В ходе исследования была протестирована Python библиотека pyrd2rdf, которая переводит данные из табличного формата в RDF. Не получилось настроить и запустить тестовую программу, так как актуальные версии Python несовместимы с библиотекой. Последний раз исходный код изменялся 7 лет назад, поэтому при выборе инструмента важно определить его актуальность и популярность. Тестовая программа представлена на рисунке 3.

```

import rdflib as _rdf
import sqlalchemy as _sqla

db_str = 'postgresql:'
db = _sqla.create_engine(db_str)
graph = _rdf.Graph('rd2rdf_dm')
graph.open(db)

print(graph.serialize(format='nt'))

```

Рис. 3. Пример использования PyRDB2RDF

Далее рассмотрим инструменты, которые были использованы для получения маппинга базы данных PostgreSQL.

В базе данных были сохранены открытые данные по карточкам ДТП г. Санкт-Петербург.

На текущий момент, наибольшей распространенностью пользуется Ontop [3]. Ontop достаточно проработан в плане оптимизации SQL запросов для повышения производительности запросов в онтологию. На текущий момент этот инструмент может быть использован только в качестве независимого приложения, консольной утилиты. Работа через API объявлена устаревшей. Позволяет генерировать маппинг в форматах OBDA и R2RML. Может использовать сгенерированный маппинг для создания SPARQL endpoint. Архитектура Ontop представлена на рисунке 4.

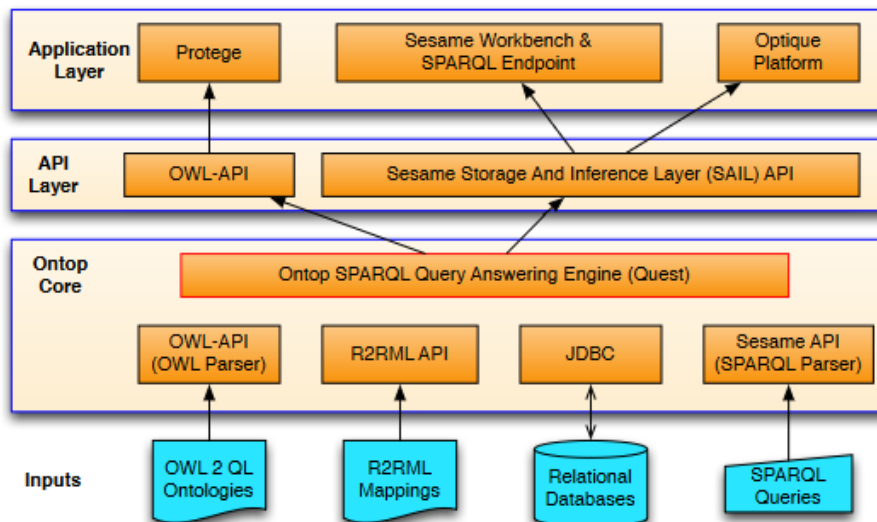


Рис. 4. Архитектура Ontop [3]

Одним из протестированных инструментов была Java программа Morph-RDB, которая была описана в статье от Политехнического Университета Мадрида [5]. При изучении исходного кода и документации были изучены и другие разработанные инструменты университета. В итоге была найдена Python библиотека Morph-KGC, которая повторяет архитектуру Morph-RDB и при этом обладает более богатым функционалом. Библиотека позволяет программно добавить дополнительную логику в работу SPARQL точки запросов. Для проверки работы тестовой программы с использованием Morph-KGC к ней был подключен уже сгенерированный маппинг Ontop, так как сам Morph-KGC не генерирует маппинг.

По итогу тестирования было получено, что чтение файла с маппингом занимает в среднем 21 секунду. Далее программа для выполнения запроса SPARQL проверяет каждую строку маппинга, что занимает по одной секунде. Это показывает низкую производительность на сгенерированном Ontop маппинге. Так как таких маппингов более 700 и была получена гораздо выше производительность у Ontop, то далее работа программы не рассматривалась. Пример настройки Morph-KGC представлен на рисунке 5. Тестовая программа для проверки работы библиотеки представлена на рисунке 6.

```

[CONFIGURATION]
output_dir=./kernel/output
output_file=result

[DataSource1]
mappings=./ontology_mapping/postgres_r2rml.ttl
db_url=postgresql:
  
```

Рис. 5. Настройка Morph-KGC (файл kernel.ini)

```

graph = morph_kgc.materialize('./kernel/configs/kernel.ini')
print(graph.query("""
BASE <http://example.edu/smart-city/>

SELECT ?iid ?road_importance ?house

WHERE {
?cid <stat_gibdd_accident_cards#road_importance> ?id .
?iid <stat_gibdd_road_importances#id> ?id .

?cid <stat_gibdd_accident_cards#house> ?house .
?iid <stat_gibdd_road_importances#road_importance> ?road_importance .

FILTER (?house = "143"^^xsd:string) .
}
"""))
  
```

Рис. 6. Использование Morph-KGC

По итогу обзора инструментов была составлена сравнительная таблица. Ontop лидирует по всем типам функциональности, которая требовалась для решения поставленной тестовой задачи.

Таблица. Сравнение методов работы с открытыми данными

Функциональность	Ontop	Morph-KGC	PyRDB2RDF
Предоставление SPARQL endpoint	Присутствует	Присутствует	Отсутствует
Обработка неразмеченных источников	Присутствует	Отсутствует	Отсутствует
Получение потоковых данных	Присутствует	Присутствует	Присутствует
Гибридный подход	Высокая	Низкая	Отсутствует

На текущий момент Ontop можно считать наиболее продвинутым инструментом для работы с R2RML маппингами. Существует также большое количество статей с обзором других решений, в которых указываются преимущества работы в сравнении с Ontop. Но на практике чаще всего оказывается, что инструмент уже устаревший. Использование R2RML маппингов позволяет задействовать существующие источники данных. Из-за этого отсутствует дублирование информации и не нужен шаг с трансформированием данных в RDF. Схему данных в реляционной базе данных можно использовать для генерации R2RML маппинга. Также это позволяет убрать необходимость ручного написания онтологии.

Литература

1. R2RML: RDB to RDF Mapping Language. [Электронный ресурс]. – Режим доступа: <https://www.w3.org/TR/r2rml> (дата обращения: 01.03.2022).
2. SPARQL query language for RDF. [Электронный ресурс]. – Режим доступа: <https://www.w3.org/TR/rdf-sparqlquery/> (дата обращения: 01.03.2022).
3. Calvanese D., Cogrel B., Komla-Ebri S., Kontchakov R., Lanti D., Rezk M., Rodriguez-Muro M., Xiao G. Ontop: Answering sparql queries over relational databases // *Semantic Web*. – 2016. – Pp. 2–8.
4. Sicilia A., Nemirovski G., Nolle A. Map-On: A web-based editor for visual ontology mapping // *Semantic Web – Interoperability, Usability, Applicability*. – 2017. – Pp. 969–980.
5. Priyatna F., Alonso-Calvo R., Paraiso-Medina S., Padron-Sanchez G., Corcho O. R2RML-based access and querying to relational clinical data with Morph-RDB. – 2015. DOI:10.3233/978-1-61499-289-9-971.

УДК 004.02

ОБЗОР МЕТОДОВ ДОСТУПА К ОТКРЫТЫМ ДАННЫМ В СИСТЕМЕ УМНОГО ГОРОДА

Волков А.А.¹ (студент)

Научный руководитель – канд. техн. наук Тесля Н.Н.¹

¹Университет ИТМО

e-mail: artvolkov96@yandex.ru, nnteslya@itmo.ru

В работе рассмотрены существующие методы доступа к открытым данным для систем умного города. Сформирован список параметров, с помощью которого можно явно разграничить методы, и составлена сравнительная таблица методов. Описан гибридный подход, который является наиболее сбалансированным в контексте выбранных параметров.

Ключевые слова: система умного города, открытые данные, обработка данных, парсинг, потоковые данные, программные интерфейсы.

Система умного города направлена на соединение гетерогенных частей городской инфраструктуры и ее оптимизации. Эффективность работы системы умного города напрямую связана с качеством и количеством входных данных. Одним из простых способов добавления данные в систему — это использование открытых источников данных.

Этот способ позволяет не заниматься пересозданием уже существующей справочной информации и помогает дать быстрый старт проекту [1]. Однако, при использовании открытых данных важно выбрать оптимальный способ работы с ними. Проблема заключается в том, что методы доступа к открытым данным, как и их формат, не до конца стандартизированы и этап предварительной настройки может занимать длительное время [2]. Также стоит отметить, что не все методы доступа могут быть одинаково эффективны для конкретной задачи.

Задачи системы умного города можно поделить в зависимости от требуемых источников данных [3]. Например, задачи, связанные с визуализацией и мониторингом, требуют доступа к данным в реальном времени. Для оставшихся задач, связанных с аналитикой и прогнозами, лучше подходят исторические данные. Исходя из этой логики, можно разделить источники данных на два типа: исторические и реального времени, а их обработка должна проводиться по отдельности.

Цель работы состоит в рассмотрении методов доступа к открытым данным и сравнении методов в контексте применения в системе умного города. Исходя из цели работы были поставлены следующие задачи: сформировать список параметров для методов, рассмотреть существующие методы работы с открытыми данными, определить их преимущества и недостатки и составить сводную сравнительную таблицу для всех методов.

Данные чаще всего можно охарактеризовать объемом и актуальностью. Соответственно, для оценки методов доступа необходимо, в первую очередь, опираться на то, как это влияет на данные. Исходя из этого можно сформулировать следующие параметры оценки методов: доступность всего набора данных, возможность получить срез данных и частота обновления данных.

В классификации не указаны параметры, которые являются общими для всех форматов данных, а значит не вносят вклад в проведение границ между методами. К таким параметрам можно отнести: охватываемый временной диапазон, ограничение на количество запросов и привязка к пространству.

Метод прямого доступа к размеченному источнику заключается в работе с исходными данными без предобработки. Под отсутствием предобработки подразумевается отсутствие слоя обработки данных на стороне источника перед отправкой, данные отдаются в первоначальном виде. Этот метод также подразумевает, что основная цель функционирования источника связана именно с передачей данных в удобном формате для дальнейшей обработки. Такое уточнение позволит более явно провести разделительную черту с методом, описанным ниже.



Рис. 1. Схема прямого доступа к размеченному источнику

В терминах программной реализации это означает, что существует всего два блока в системе. В одном блоке находятся данные, а в другом - алгоритмы. Графически метод представлен на рисунке 1. Стоит обратить внимание, что стрелка двунаправленная.

Это связано с тем, что можно как конкретизировать, так и расширить запрос.

Примерами источников являются статичные таблицы и программный интерфейс (API). Под статичными таблицами подразумеваются данные, которые выложены в качестве единого файла. Формат файла обычно относится к CSV (comma separated values, разделенные запятыми значения), JSON (JavaScript object notation, объектная нотация JavaScript) или XML (extensible markup language, расширяемый язык разметки). Это самый простой способ выгрузки данных, который используется на многих государственных порталах открытых данных.

Под API понимается сервис, с которым можно организовать взаимодействие в формате запрос-ответ. Такой способ предоставляют многие порталы открытых данных.

К преимуществам метода можно отнести строгую формализацию данных и высокую скорость работы. Воспроизводство результатов при повторном запросе помогает упростить логику обработки данных. Также вариативность запроса данных позволяет получить только нужную часть данных. Это актуально для больших исторических данных, которые необходимо предварительно обрабатывать перед использованием.

Кроме преимуществ есть и недостатки. Первый — сложно оценить валидность данных. Ответственным за выгрузку данных является только их поставщик. Выкладываемые данные могут не проверяться на наличие синтаксических ошибок, что достаточно трудно отследить и может приводить к ошибкам обработки. Иногда, в таких случаях в методе предусматривают этап предобработки данных, если наблюдается закономерность в ошибках. Также с течением времени возможно изменение схемы данных, что может потребовать доработки инструмента анализа.

Другим недостатком является низкая частота обновления. Требования к обновлению таких данных чаще всего отсутствуют, а значит актуальность данных всегда будет под вопросом.

На рисунках 2 и 3 показаны отрывки из документации по использованию API двух государственных порталов открытых данных (Санкт-Петербурга и Москвы). На них можно заметить, что получение полного набора данных чаще всего невозможно. Такой формат данных лучше всего подходит для частичной выгрузки.

Содержимое набора данных

```
Запрос к ресурсу /datasets/{id}/versions/{num_id}/data/ возвращает содержимое набора.
Запрос:
GET http://data.gov.spb.ru/api/v1/datasets/18/versions/1/data/
Параметры запроса:
page - Номер страницы, необязательно.
per_page - Количество объектов на странице, по умолчанию 10, максимум 100, необязательно.

Ответ:
[
  {
    "num_id": 1,
    "row": {
      "id": "1",
      "name": "Элемент 1"
    }
  },
]
```

Рис. 2. Часть инструкции к использованию портала открытых данных Санкт-Петербурга

Содержимое набора данных

Запрос к ресурсу `/datasets/{id}/rows` возвращает список строк указанного набора данных.

Внимание: при запросе датасетов с количеством записей более 10000шт., в ответе будет передан статус 413. Получение таких датасетов возможно с применением описанных ниже параметров `$top` (максимальное допустимое значение - 500шт.) и `$skip`. Для получения количества записей в датасете можно воспользоваться запросом `/datasets/{id}/count`.

Рис. 3. Часть инструкции к использованию портала открытых данных Москвы [4]

В методе обработки неразмеченных источников вручную определяется зона поиска данных в сети и проводится автоматизированный сбор и обработка. Данные при этом, чаще всего, собираются с нескольких источников, где они могут быть представлены в разном виде. Из-за этого требуется периодическое участие человека в работе программы, так как схема доступа периодически меняется. Основная цель данного источника данных заключается в предоставлении информации в удобном виде для человека.



Рис. 4. Схема обработки неразмеченных источников

В большинстве случаев подразумевается, что открытые данные могут располагаться только на специализированных порталах, что не совсем корректно. Любые открытые ресурсы можно использовать в качестве источника открытых данных. Примерами подобных источников являются веб-страницы и новостные ленты. Эти источники всегда в свободном доступе и поэтому сбор данных с них легко автоматизировать.

Первое преимущество данного подхода – это высокая частота обновления. Скорость обновления таких данных очень высока, так как в этом есть прямая заинтересованность источников. Особенно это актуально для новостных источников и биржевых котировок. Второе преимущество - вариативность доступа к данным. Хотя данные чаще всего и представлены как HTML страницы, парсинг их представляет собой нетривиальную задачу. На одной странице могут быть данные, относящиеся к разным предметным областям и отличающиеся по уровню детализации. Это требует большей конкретизации запроса.

К недостаткам такого метода относится необходимость предварительной обработки. Некоторые порталы дополнительно содержат в себе защиту от доступа автоматических инструментов. Все это повышает временные затраты на подготовку и не всегда рационально. Схема данных может значительно отличаться от источника к источнику, поэтому требуется каким-то образом сопоставлять различные описания. И в зависимости от методов обработки данных можно получать различающиеся результаты из одного и того же источника.

Потоковая обработка данных – это набирающий популярность подход в высоконагруженных распределенных системах.

Это подход пересекается с понятием событийно-ориентированной системы. В данном подходе программа-клиент подписывается на новые данные и в момент их появления получает уведомление. Графически метод представлен на рисунке 5.



Рис. 5. Схема работы с потоковыми данными

Обмен данными с источником двусторонний, так как подписка на нужные обновления происходит в начальный момент времени. Но этим можно пренебречь, чтобы показать основное отличие от остальных подходов.

Примером таких данных являются открытые данные в реальном времени. Это данные о местоположении общественного транспорта, текущая погода, данные о свободном паркинге, биржевые котировки. Эти типы данных объединяет высокая ценность в конкретный момент времени. Большинству людей важнее какая погода на сегодня, чем какая погода была два месяца назад. На рисунке 6 показан пример доступа через HTTP-запрос к системе проката велосипедов в Берлине.

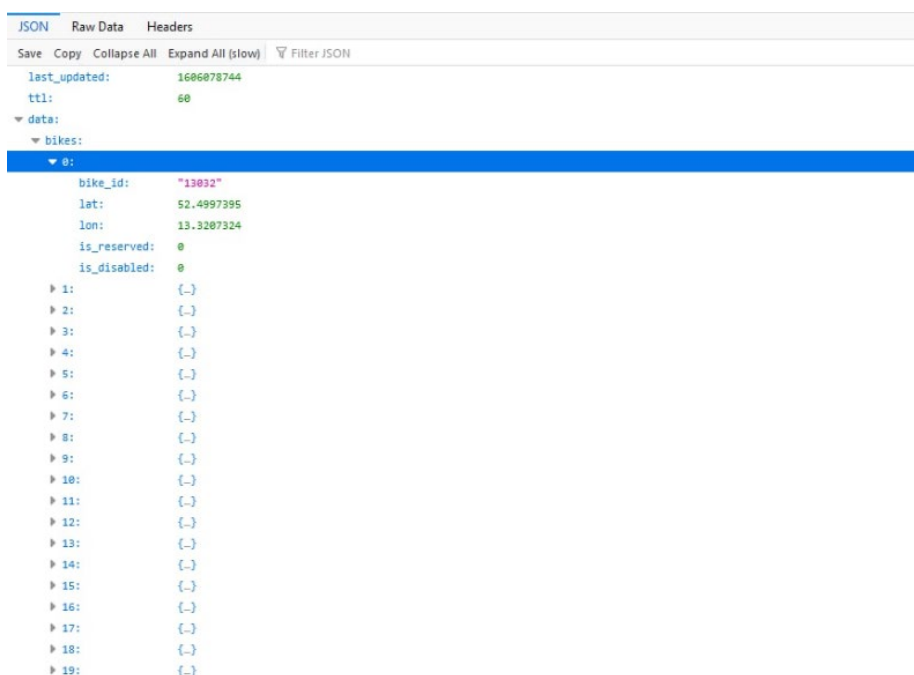


Рис. 6. Получение списка геоточек со свободными велосипедами из открытого источника данных

В случае данных, обновляемых в реальном времени, частота обновления будет максимальной. Задержки в данном случае будут минимальны по отношению ко всем другим подходам. Из этого следует преимущество получения актуальных значений и удобство визуализации. А также с использованием этого подхода можно настроить мониторинг аномалий и проактивное реагирование на проблемы. Но это больше относится к задачам, которые лучше всего решаются с использованием потоковых данных.

Недостатки подхода происходят от его преимуществ. Чаще всего, используя потоковые данные, можно получить часть данных, но не весь объем. Это связано с тем, что исторические данные могут быть представлены в другом формате. Тогда обработка таких данных должна происходить отдельно, что влечет за собой дополнительные затраты.

Вышеописанные подходы с прямым доступом к размеченному источнику и с потоковыми данными противоположны и независимы по своей сути. Однако, их можно совместить в рамках единого метода. В таком случае получить весь объем данных и построить по нему срез можно с помощью статических таблиц [5]. А для визуализации и получения актуальных значений лучше будет использовать данные в реальном времени. Подход далее можно усовершенствовать, записывая данные в реальном времени в отдельную таблицу. Графически подход представлен на рисунке 7.



Рис. 7. Схема гибридного подхода

По итогу обзора каждого из методов доступа к открытым данным было составлено сравнение, представленное в таблице.

В ходе обзора были рассмотрены основные методы доступа к открытым данным, приведены их преимущества и недостатки. Метод доступа влияет на обработку и хранение данных, а также на возможный спектр решаемых задач. Поэтому перед выбором метода доступа к данным важно определить какую задачу необходимо решить. Затем – насколько актуальные данные требуются и в каком объеме. Немаловажную роль играет и частота обновления данных. Это все относится к требованиям для решения исходной задачи.

Таблица. Сравнение методов работы с открытыми данными

Метод работы	Доступность всего набора данных	Возможность получить срез данных	Частота обновления данных
Прямой доступ к размеченному источнику	Полный доступ	Есть	Редко
Обработка неразмеченных источников	Частично	Нет	Часто
Получение потоковых данных	Частично	Частично	В реальном времени
Гибридный подход	Полный доступ	Есть	В реальном времени

Без проработки вопросов итогов результата в середине реализации может выясниться, что выбранный метод не подходит. Из схем работы методов можно заметить, что выбор конкретного метода влияет на архитектуру системы. А значит смена метода влечет в себе большие затраты начиная с первого шага реализации. Некоторые порталы открытых данных предоставляют возможность использовать гибридный подход в работе с данными. Это является предпочтительным вариантом в большинстве задач, так как позволяет упростить изменить решения.

Литература

1. Ojo A., Curry E., and Zeleti F.A. A tale of open data innovations in fivesmart cities // 48th Hawaii International Conference on System Sciences. – 2015. – Pp. 2326–2335.
2. Malik K., Sam Y., Hussain M., Abuarqoub A. A methodology for real-time data sustainability in smart city: Towards inferencing and analytics for big-data // Sustainable Cities and Society. – 2017. – 39 p.
3. Haslam P.A. Bigger data and quantitative methods in the study of socio-environmental conflicts // Sustainability. – 2020. – Pp. 12–18.
4. Портал открытых данных города Москвы. [Электронный ресурс]. Режим доступа: <https://apidata.mos.ru/Docs>. (дата обращения: 31.01.2021).
5. Schauppenlehner T., Muhar A. Theoretical availability versus practical accessibility: The critical role of metadata management in open data portals // Sustainability. – 2018. – Pp. 10–12.

УДК 004.934.2

РАСПОЗНАВАНИЕ КАШЛЯ С ПОМОЩЬЮ АНАЛИЗА СПЕКТРОГРАММ

Двойникова А.А.^{1,2} (аспирант)

Научный руководитель – д-р техн. наук, доцент Карпов А.А.^{1,2}

¹*Санкт-Петербургский Федеральный исследовательский центр Российской академии наук*

²*Университет ИТМО*

e-mail: nastya.yal@yandex.ru, karpov@iias.spb.su

Работа выполнена в рамках проекта Российского Фонда Фундаментальных Исследований № 20-04-60529 и гранта Совета по грантам Президента РФ № НШ-17.2022.1.6.

В работе рассматривается метод автоматического распознавания кашля с использованием анализа спектрограмм и сверточной нейронной сети. Описывается методика сбора базы данных, содержащей экстралингвистические события, такие как кашель и другие звуковые события, и ее описание. Проводятся экспериментальные исследования с различными методами аугментации изображений спектрограмм, такие как угол сдвига, изменение масштаба, сдвиг по ширине, изменение яркости и горизонтальный поворот.

Ключевые слова: спектрограмма, сверточная нейронная сеть, аугментация изображений спектрограмм, распознавание кашля, экстралингвистические события.

В настоящее время набирает популярность область телемедицины, преимущество которой в том, что больной может, не контактируя с другими людьми, узнать свой диагноз, а врач может удаленно отслеживать состояние своего пациента. Основываясь на этом, с каждым годом все больше и больше развиваются автоматические системы анализа здоровья людей. Такие системы позволяют мониторить физическое состояние сотрудников на рабочем месте или людей с хроническими заболеваниями, для которых необходимо регулярное наблюдение состояния здоровья. Автоматическое распознавание кашля – одна из самых востребованных задач в области телемедицины, т.к. кашель может способствовать передаче от человека к человеку инфекции, опасной для дыхательных путей. Поэтому системы распознавания кашля могут помочь в предотвращении распространения опасных заболеваний, таких как туберкулез, грипп или коронавирусная инфекция, а также в мониторинге состояния здоровья работников для формирования здорового окружения в коллективе.

Анализ предыдущих исследований. На сегодняшний день работ в области автоматического распознавания кашля с применением алгоритмов машинного обучения достаточно мало. Однако, можно выделить основные направления исследований. Для автоматического распознавания кашля в статье [1] используются мел-частотные кепстральные коэффициенты (MFCC), которые потом анализируются при помощи градиентного бустинга. В работе [2] анализируется набор неречевых акустических событий, в том числе и кашель, при помощи скрытых марковских моделей. Однако, большинство исследователей при анализе кашля используют спектрограммы аудиосигнала, т.к. кашель относится к неречевым акустическим событиям, а их спектральные характеристики являются наиболее репрезентативными для автоматического анализа. В работе [3] для анализа кашля из акустических данных извлекались спектрограммы, которые подавались на вход математическому алгоритму на основе правил. С развитием глубоких нейронных сетей

изменился и подход в области распознавания кашля. В работе [4] используются сети прямого распространения для автоматического распознавания кашля. Как показывает анализ предыдущих исследований, наиболее эффективным считается подход, основанный на применении спектрограмм и нейронных сетей.

База данных. Для экспериментальных исследований в области распознавания кашля вручную собирался корпус данных, содержащий в себе аудиозаписи антропофонов и речи. Антропофоны – языковое представление неречевых звуков, издаваемых человеком [5] (то же, что экстралингвистические составляющие речи). В данной работе будут использоваться такие антропофоны, которые передают физиологическое, такие как кашель, хрип, сморкание, икота, отрыжка, зевота, першение, сопение, чихание, храп, и эмоциональное состояние человека, например смех и плач, и звуков, передающих движение человека, например жевание. Кашель собирался из нескольких существующих корпусов и дополнялся аудиозаписями, собранными вручную из интернет-источников [6–8]. Так из корпуса Coswara [9] было использовано 421 аудиофайла, содержащих звуки кашля, из корпуса Cambridge COVID-19 Sound [10], который был использован на соревнованиях ComParE 2021 [11] в рамках конференции INTERSPEECH 2021, было использовано 725 и 364 аудио, содержавших звуки кашля и речи, соответственно. Каждая аудиозапись содержала только одно акустическое событие. Все антропофоны были сгруппированы в 2 класса: кашель и не кашель. На рисунке 1 показано распределение данных по двум классам, а также количество экземпляров аудиоданных для каждого антропофона.

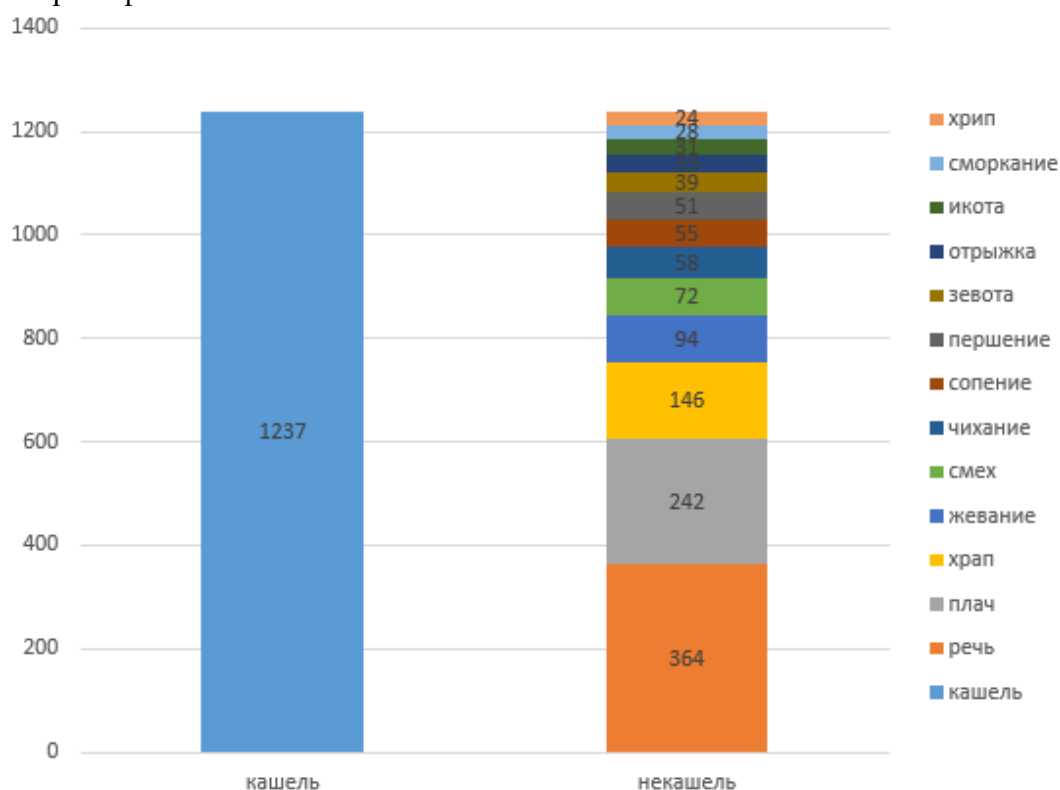


Рис. 1. Распределение экземпляров классов кашель и не кашель

Как видно из рисунка 1 общий объем базы данных составляет 2437 экземпляров аудиофайлов, причем классы кашель и не кашель являются сбалансированными. Затем все данные делились на три множества: обучающее, валидационное и тестовое, в процентном соотношении 70, 15 и 15, соответственно. В итоге объем обучающего множества составил 1730 аудиофайлов, валидационного – 370 и тестового – 374.

Предлагаемый подход. Для автоматического распознавания кашля в работе используется подход, основанный на извлечении спектрограмм, которые потом подаются на вход сверточной нейронной сети (англ. Convolution Neural Network, CNN). Архитектура используемой CNN состоит из 3 слоев свертки с пулингом, она показана на рисунке 2.

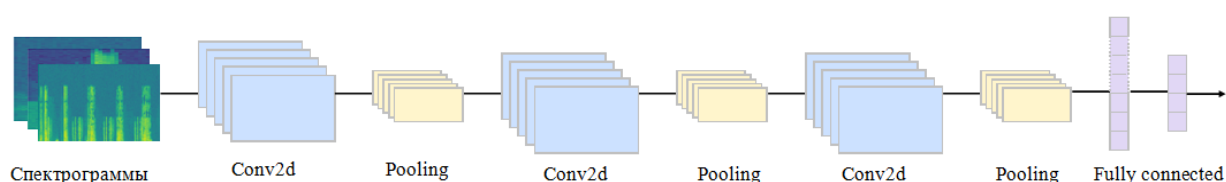
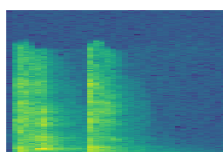


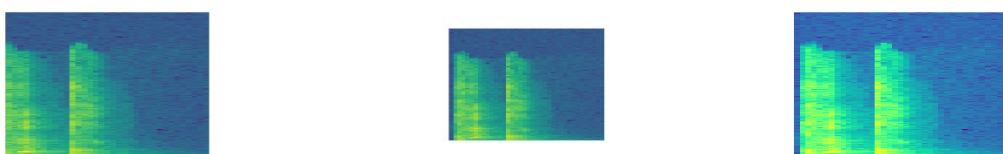
Рис. 2. Архитектура сверточной нейронной сети

Перед обучением нейронной сети все аудиозаписи приводились к одной длине – 5 секунд, более короткие аудио дополнялись нулями в конце, а более длинные – обрезались. Затем из всех аудиосигналов извлекались изображения спектрограмм с разрешением 433x288 пикселей. Для обучения сверточной нейронной сети использовались следующие параметры: размер батча (`batch_size`) = 128, скорость обучение (`learning rate`) = 0,02, оптимизатор – `sgd`, функция потерь – категориальная перекрестная энтропия, количество эпох – 50, причем сохранялась наилучшая модель (наиболее низкое значение функции потерь на валидационном множестве) обучения, которая использовалась потом при проверке эффективности модели на тестовом множестве.

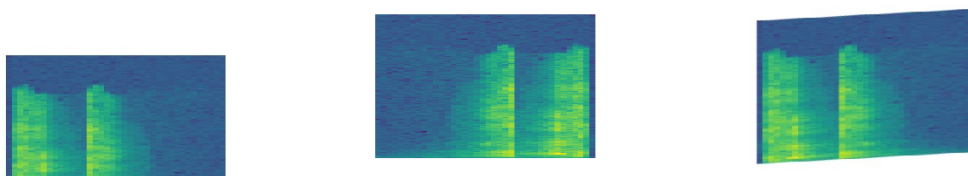
Так как объем обучающих данных небольшой, то к изображениям спектрограмм из обучающего набора применялись различные методы аугментации. В библиотеке `keras` [12] языка программирования Python встроена функция для автоматической аугментации онлайн, которая позволяет видоизменять изображения, в данном случае спектрограммы. Для экспериментальных исследований использовались следующие методы видоизменения: поворот по часовой стрелки, изменение масштаба, сдвиг по ширине и высоте, изменение яркости и горизонтальное отражение. На рисунке 3 показаны примеры видоизменений спектрограмм, для наглядности производилось искажение сильнее, чем использовалось при обучении.



а) изображение исходной спектрограммы



б) слева направо: сдвиг по ширине, изменение масштаба, изменение яркости



в) слева направо: сдвиг по вертикали, горизонтальное отражение, угол наклона

Рис. 3. Методы видоизменения изображений спектрограмм

Экспериментальные исследования проводились с каждым методом видоизменения отдельно. Для каждого метода устанавливался определенный диапазон значений, в рамках которого случайным образом при онлайн аугментации выбиралось значение и применялось для каждой спектрограммы.

Результаты экспериментов. Сдвиг по ширине изображения спектрограммы имеет физический смысл и обозначает уменьшение длины аудиосигнала путем обрезания какого-либо участка аудио, который выбирался случайным образом при онлайн аугментации. Для экспериментальных исследований выставлялось значение 0,2, это означает, что во время аугментации случайным образом выбирается значение из диапазона от 0 до 0,2 раз и на этот коэффициент уменьшается ширина спектрограммы, если устанавливается значение 0, то спектрограмма никак не видоизменяется. Изменение масштаба производилось в диапазоне от 0,2 до 0,2. Изменение яркости в физическом смысле означает изменение громкости сигнала и производится в диапазоне от 0 до 1, что означает увеличение яркости на выбранный коэффициент. Сдвиг по вертикали отображает изменение отображаемой полосы частот на спектрограмме и для экспериментов устанавливается значение в диапазоне от 0 до 0,2, это означает, что изображение может уменьшиться по вертикали максимум в 0,2 раза. Горизонтальное отражение отображает сигнал в обратной последовательности и его значение устанавливается в булевом виде. Угол наклона спектрограммы варьируется в диапазоне от 0 до 5°, таким образом, изображение спектрограммы может быть повернуто максимум на 5° по часовой стрелке. Результаты экспериментов с использованием аугментацией различными методами видоизменения изображений показаны в таблице. Для оценки эффективности предложенных подходов автоматического распознавания кашля использовался показатель точности (ассигасу), т.к. он позволяет произвести объективную оценку точности на сбалансированных данных.

Таблица. Результаты экспериментальных исследований

Метод аугментации изображений спектрограмм	Точность распознавания, %	Прирост точности распознавания, %
Без аугментации	61,23	–
Изменение угла поворота	70,32	9,09
Изменение масштаба	70,86	9,63
Сдвиг по ширине	70,58	9,35
Сдвиг по высоте	67,37	6,14
Изменение яркости	69,25	8,02
Горизонтальное отражение	61,74	0,51

Как видно из таблицы, большинство методов аугментации спектрограмм позволяют улучшить базовый результат распознавания кашля, который по показателю точности равному 61,23% для двух классов (кашель и не кашель). Наилучший результат 70,86% достигается при использовании онлайн аугментации методом видоизменения масштаба изображений спектрограмм. Наименьший прирост в точности всего на 0,51% показал метод аугментации при помощи горизонтального отражения, это связано с тем, что горизонтальный поворот в физическом смысле обозначает аудиосигнал в обратной последовательности, т.е. с конца до начала, для анализа неречевых акустических событий такой подход не имеет физического смысла, соответственно, его применение в машинном обучении тоже не является эффективным.

В работе рассматривается подход к распознаванию кашля с помощью анализа спектрограмм и сверточной нейронной сети. Проводятся экспериментальные исследования с различными методами аугментации спектрограмм, которые позволяют повысить результаты распознавания кашля с 61,23% до 70,86% по показателю точности распознавания антропофонов при использовании онлайн аугментации с помощью изменения масштаба изображения спектрограмм.

Литература

1. Orlandic L., Teijeiro T., Atienza D. The COUGHVID crowdsourcing dataset, a corpus for the study of large-scale cough analysis algorithms // *Scientific Data*. – 2021. – Т.8. – №1. – С. 1–10.
2. Карпов А.А., Акарун Л., Ронжин А.Л. Многомодальные ассистивные системы для интеллектуального жилого пространства // *Информатика и автоматизация*. – 2011. – №19. – С. 48–64.

3. Абрамов Г.В. и др. Анализ и использование математических методов для распознавания звуковых сигналов // Вестник Воронежского государственного университета инженерных технологий. – 2015. – №2(64). – С. 61–65.
4. Левчук П.О., Иванков А.Ю., Максимов А.В. Система анализа респираторной активности // Сборник студенческих научных работ факультета компьютерных наук ВГУ. – 2020. – С. 151–156.
5. Курашкина Н.А. Роль антропофонов как невербальных компонентов коммуникации // Homo loquens в языке, культуре, познании. – 2010. – С. 176–184.
6. Zvukogram. [Электронный ресурс]. – Режим доступа: <https://zvukogram.com> (дата обращения: 01.03.2022).
7. Zvukipro. [Электронный ресурс]. – Режим доступа: <https://zvukipro.com> (дата обращения: 01.03.2022).
8. 4uall.net. [Электронный ресурс]. – Режим доступа: 4uall.net (дата обращения: 01.03.2022).
9. Sharma N. et al. Coswara--A Database of Breathing, Cough, and Voice Sounds for COVID-19 Diagnosis // In Proceedings of the INTERSPEECH-2020. – 2020. – Pp. 4811–4815. DOI: 10.21437/Interspeech.2020-2768.
10. Brown C., Chauhan J., Grammenos A. et al. Exploring Automatic Diagnosis of COVID-19 from Crowdsourced Respiratory Sound Data // In Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining (KDD '20). – 2020. – Pp. 3474–3484. DOI: 10.1145/3394486.3412865.
11. Schuller B., Batliner A., Bergler C., et al. The INTERSPEECH 2021 Computational Paralinguistics Challenge: COVID-19 Cough, COVID-19 Speech, Escalation & Primates // In Proceedings of the INTERSPEECH-2021. – 2021. – 5 p.
12. Fran Chollet. Keras. 2015. [Электронный ресурс]. – Режим доступа: <https://keras.io> (дата обращения: 01.03.22).

УДК 711.796.542

АНАЛИЗ ПРАКТИК И ПРОЕКТОВ В ОБЛАСТИ УСТОЙЧИВОГО ТУРИЗМА И ТУРИСТСКИХ ОБЪЕКТОВ В СЕВЕРНЫХ СТРАНАХ

Дырко А.В.¹ (студент)

Научный руководитель – канд. псих. наук, доцент Шмелева И.А.¹

¹Университет ИТМО

e-mail: i_a_shmeleva@itmo.ru, alisadyrko@gmail.com

В работе рассмотрена сфера устойчивого туризма, как компонента, способствующего ускорению прогресса в достижении Целей устойчивого развития ООН. Ограничения на международные поездки, введенные в связи пандемией COVID-19, стали импульсом для развития внутреннего туризма. В России устойчивый туризм только начинает развиваться. Индустрия нуждается в разработке и применении механизмов, способствующих достижению устойчивости. В свою очередь перенимание зарубежного опыта может ускорить этот процесс.

Ключевые слова: туризм, устойчивый туризм, устойчивое развитие, экологический туризм, циркулярный туризм.

По данным ООН за 2020 год из-за пандемии туристический сектор потерял 320 миллиардов долларов, что в 3 раза больше, чем из-за экономического кризиса в 2009 году [1]. Международный туризм был практически полностью приостановлен.

С одной стороны, вызванный COVID-19 кризис, привел к ряду разрушительных экономических и социальных последствий, с другой, он поспособствовал некоторым экологическим улучшениям. Производство отходов, потребление природных ресурсов, чрезмерное давление на экосистемы и местные сообщества — все это проблемы, остающиеся в рамках туристической отрасли. Однако, в 2020 году закрытие заводов и электростанций, резкое сокращение объема авиаперевозок и ограничения на передвижение людей в автомобилях заметно сократило уровень загрязнения и выбросов парниковых газов, что положительно сказалось на экологической ситуации в регионах и городах.

Сложившаяся в 2020 году ситуация показала необходимость поиска эффективных путей адаптации к новым реалиям. Туризм должен изменяться, и в последующем формироваться таким образом, чтобы становиться лучше, уделять внимание инклюзивности и повышению устойчивости.

В то же время необходимо отметить, что результаты исследования, проведенного в 2021 году в России при поддержке Ростуризма показали катастрофическую неосведомленность руководителей туристических предприятий о концепции устойчивого развития: лишь 27% участников опроса знают Цели устойчивого развития ООН и видят их конкурентное преимущество, из них только 14% опираются на ЦУР ООН в практической деятельности, лишь у 2% имеется стратегия устойчивого развития собственной компании, но большинству слова «устойчивый туризм» неизвестны и непонятны, а это 73% участников проведенного опроса [2].

Развитие и популяризация устойчивого туризма может помочь восстановлению и поддержанию индустрии, поскольку одной из предпосылок, на которых основан устойчивый туризм, является отсутствие перенаселенности туристических направлений. Чтобы гарантировать долгосрочную устойчивость туризма, необходимо найти баланс между тремя принципами устойчивого развития: экологическим, экономическим и социокультурным аспектами.

Единое определение устойчивого туризма до настоящего времени отсутствует, трактовка ВТО имеет наибольшее распространение, согласно которой «устойчивым называется туризм, который в полной мере учитывает текущие и будущие экономические, социальные и экологические последствия туристической деятельности, удовлетворяя потребности посетителей, индустрии и местных сообществ [3].

Устойчивое развитие требует участия всех соответствующих заинтересованных сторон, в особенности сильного политического руководства. А также постоянного мониторинга последствий принятия необходимых превентивных или корректирующих мер.

В России существуют механизмы устойчивого туризма, но их применение не носит массовый характер. Таким образом, актуальность данного исследования обусловлена потребностью России перенимать успешный зарубежный опыт в области развития устойчивого туризма.

Проблемы, вызванные туристической индустрией

После доклада комиссии Г.Х. Брундтланд «Наше общее будущее» [4] в 1987 году, а особенно в последние годы, устойчивое развитие как концепция была интегрирована в различные исследования, связанные с туризмом. Тенденция роста числа научных публикаций по теме устойчивого туризма на базе Scopus за последние 5 лет показана на рисунке 1.

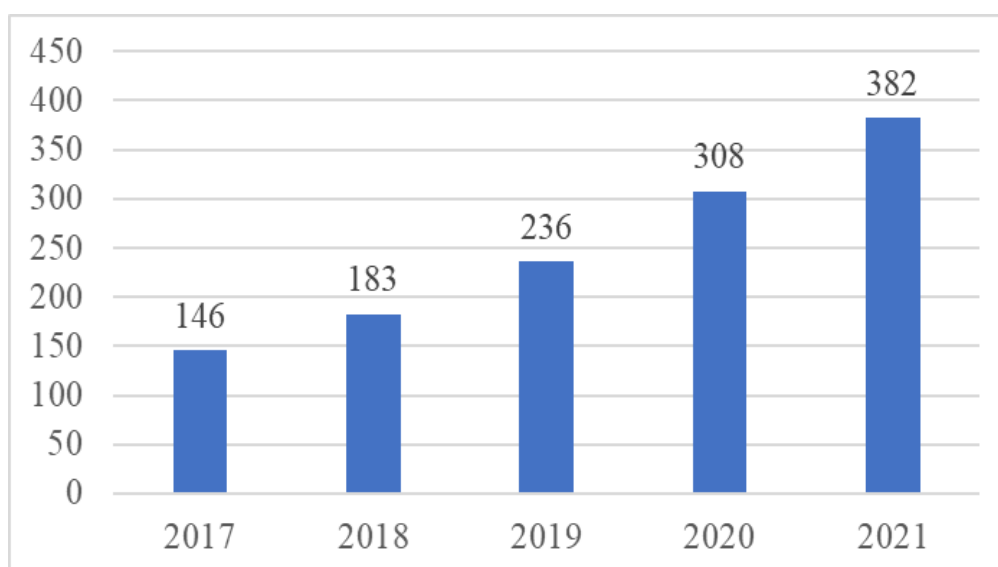


Рис. 1. Количество научных публикаций (за год) по теме устойчивого туризма на базе Scopus (декабрь 2021 года)

Как видно из рисунка 1, количество публикаций, посвященных сфере устойчивого туризма, растет с каждым годом. Анализ литературы показывает, что эта тема активно обсуждается в научном и экспертном сообществе с различных сторон.

Наибольшее распространение в академической литературе получила тема влияния туризма на изменение климата [5, 6, 7]. В частности, было отмечено, что связь между изменением климата и туризмом стала одной из наиболее важных и динамичных областей исследований в области устойчивого туризма в последние годы [5]. Приведенная исследователями оценка глобальных потоков углерода, связанных с туризмом, между 160 странами, показала, что большая часть углеродного следа приходится на страны с высоким уровнем дохода – США, Евросоюз, Россию, Китай.

Шведские ученые проанализировали теоретический потенциал сокращения выбросов за счет изменения видов используемого транспорта и направлений, а также потенциал готовности туристов изменить свое поведение в путешествиях. Помимо этого, на примере Швеции они представили методы расчета и анализа воздействия туристического транспорта на климат с учетом типов поездок туристов, расстояния и вида транспорта [6].

Еще одной важной темой является управление энергетическими и водными ресурсами, а также управление отходами. Так, например, датский остров Самсе с населением около

4000 человек обеспечен энергией за счет возобновляемых источников, поступающей от солнечных батарей и ветряных турбин. Более того, электроэнергия производится значительно больше, чем используется непосредственно жителями острова. В связи с чем администрация острова работает в направлении улучшения климатической нейтральности, за счет перехода острова к умной энергетике через сокращение экспорта энергии. В приведенном исследовании [7] также рассматриваются социальные аспекты – численность населения острова с годами уменьшается, это приводит к переизбытку энергии. Туризм играет важную роль в жизни острова, создавая около трети рабочих мест. Его развитие, по мнению авторов, может помочь привлечению новых жителей на остров, что увеличит объем потребляемой энергии непосредственно на его территории.

В академических исследованиях в области устойчивого туризма уделяется большое внимание применению сертификации по зеленым стандартам с акцентом на то, как она способствует соблюдению конкретных экологических стандартов и развитию более устойчивого туризма. Как показало исследование [8], для реализации практики устойчивого туризма в организациях по размещению туристов чрезвычайно важно отношение руководителей этих организаций к экологическим маркировкам. Еще одним важным выводом этого исследования было то, что устойчивое управление и операционная деятельность отелей значительно различались в зависимости от того, были ли они сертифицированы по международным зеленым стандартам или нет. Кроме того, изучение поведения менеджеров, предприятий и туристов имеет основополагающее значение для понимания того, какие благоприятные факторы могут способствовать внедрению практики устойчивого туризма.

Интегрирование принципов экономики замкнутого цикла в туризм – еще одна тема, рассматриваемая исследователями. Экономика замкнутого цикла, признающая фундаментальную роль окружающей среды, представлена как альтернатива линейной модели экономики. Гостиничный сектор и сектор туризма в целом подверглись критике за недостаточное решение экологических проблем и влияние на глобальное потепление. Розато П.Ф., Капуто А., Валенте Д. и Пицци С. [9] выявили основные преимущества перехода от линейной экономики к циклической в гостиничном секторе. Также исследователями была описана трехосная бизнес-модель для осуществления этого перехода.

В то же время, просвещение туристов и отношение потребителей туристических услуг к более экологичному и устойчивому виду отдыха рассматривается в качестве ключевой задачи перехода к принципам устойчивого туризма. Исследователи из Норвегии полагают, что именно понимание психологически механизмов, лежащих в основе реакции туристов на жесткие и мягкие высказывания, пропагандирующие про-экологическое поведение в туризме, могут стимулировать развитие устойчивого туризма [10].

Наконец, самые актуальные вопросы в настоящее время связаны с преодолением последствий пандемии COVID-19. Пандемия остановила мобильность во всем мире, что привело к серьезному нарушению рыночных механизмов глобального туризма. Иоаннидис В. и Пноти С. [11] рассматривают процессы выхода из предыдущих кризисов за счет государственной поддержки сектора. Помимо этого, исследователи обсуждают то, что COVID-19 предоставляет государственным, частным и академическим организациям уникальную возможность разработать и закрепить переход к более экологичному и сбалансированному туризму. Ученые, занимающиеся туризмом, например, могут сыграть ведущую роль в этом, изменив свою учебную программу, чтобы подготовить будущих лидеров отрасли к более ответственному подходу к путешествиям и туризму.

Проекты устойчивого туризма, реализованные в Северной Европе

Туризм – очень сложная отрасль, включающая в себя множество заинтересованных сторон (иногда с противоположными интересами) и требующая значительного объема ресурсов. Таким образом, туризм может иметь совершенно противоположные последствия в зависимости от способа управления данным видом деятельности. При правильном управлении туризм может сыграть позитивную роль в социальном, культурном, экономическом, экологическом и политическом развитии.

В ходе проведенного исследования были рассмотрены и проанализированы туристические организации, располагающиеся в странах Северной Европы, которые в настоящее время применяют методы снижения воздействия своего бизнеса на окружающую среду, а также способствуют социальному и экономическому развитию. В частности, были рассмотрены следующие туристические объекты: отели в Финляндии, имеющие экологический сертификат Green key, в которых, например, реализованы следующие принципы устойчивого туризма:

- а) здания отеля выполнены из местного натурального материала – дерева;
- б) в отеле сконструированы зеленые крыши, которые компенсируют растительность, потерянную за счет построенного здания;
- в) отель следит за своим углеродным следом, стараясь его минимизировать;
- г) реализуются решения по сокращению энергопотребления и экономии воды;
- д) используется возобновляемая энергия – солнечная и геотермальная;
- е) осуществляется сортировка отходов;
- ж) используются только местные продукты для приготовления пищи;
- з) отель осуществляет сотрудничество с экологическими организациями и местными сообществами.

Другой пример – гостиничная сеть в Дании, сертифицированная по зеленому стандарту DGNB и сертификату Organic Cuisine Label. Сеть реализует следующие принципы устойчивого туризма:

- а) в отелях эксплуатируются зеленые крыши;
- б) используется альтернативная энергетика, в частности установлены солнечные батареи;
- в) осуществляется сортировка и переработка отходов;
- г) используются органические местные продукты;
- д) внедряются зеленые инициативы, такие как экономия электричества, экономия расхода воды; передается просьба о повторном использовании полотенец;
- д) используется фурнитура собственного производства и органических материалов;
- ж) рекомендуется использование общественного транспорта и электромобилей;
- и) предметы интерьера и мебели передаются в дар после ремонта для дальнейшего использования.

Еще один пример гостиничной сети в Норвегии, Дании и Исландии, имеющей международный сертификат Green Globe, и разработавшей собственный план реализации принципов устойчивого развития, включающего 5 ключевых областей, а именно:

- 1) устойчивое управление;
- 2) заботу о социуме и экономике;
- 3) поддержание культурного наследия;
- 4) экологические принципы управления и заботу об окружающей среде;
- 5) оценку полученных улучшений за предыдущий год.

Таким образом, проанализированные данные проектов устойчивого туризма, реализованные в странах Северной Европы, позволили сделать следующие выводы:

1. В туристической отрасли Северной Европы большое внимание уделяется эффективному использованию ресурсов и внедрению новых технологий – в рассматриваемых объектах используется геотермальная энергия, тепловые воздушные насосы, солнечные панели.

2. Осуществляется экономия электричества и воды, за счет использования санитарных приборов с низкими расходами, а также энергоэффективных электрических приборов.

3. Туристические объекты следят за своим углеродным следом, стараясь свести его к минимуму, например, за счет использования экологически чистых материалов при постройке отелей, зданий и сооружений.

4. Во всех представленных объектах производится сортировка и переработка отходов, некоторые предприятия собирают биологические отходы в собственный компостер.

5. Помимо этого, большинство предприятий сотрудничает с поставщиками, которые разделяют их ценности в области устойчивого развития, отдавая предпочтение органическим продуктам местного производства.

6. Важным аспектом для объектов туристической отрасли, которые располагаются за пределами городов, является сохранение биологического разнообразия, мест обитания зверей и птиц, дикой природы, и сведение к минимуму ущерба, наносимого ими природе.

7. Экскурсионные туры ограничивают количество участников для уменьшения антропогенного воздействия на среду.

8. Особое внимание уделяется уважению исторического наследия. За счет продвижения аутентичной культуры, традиций и самобытности принимающих общин туристические организации стараются поддерживать и повышать качество жизни местных.

9. С точки зрения экономики, предприятия стараются максимизировать вклад туризма в экономическое процветание региона, используя продукцию местных фермеров и мастеров, а также предоставляя рабочие места местному населению.

В странах Северной Европы, таких как: Дания, Швеция, Финляндия, Норвегия и Исландия устойчивый туризм является широко признанной концепцией, что подтверждается академическими публикациями, докладами Всемирной туристской организации и Европейской комиссии, а также большим числом реализованных проектов и эксплуатируемых туристических объектов.

Анализ передового опыта в области управления ресурсами и реализации проектов устойчивого туризма занимает значительную часть исследований, поскольку распространение передовой практики является ключевым фактором повышения устойчивости в данной сфере.

Важной обсуждаемой темой является роль восприятия устойчивого туризма и поведения стейкхолдеров – менеджеров, самих туристов, организаций в продвижении практик устойчивого туризма, что рассматривается как фактор, благоприятствующий его развитию.

Перспективной бизнес-моделью развития устойчивого туризма рассматривается интеграция принципов экономики замкнутого цикла в туристический сектор.

Пандемия COVID-19 очень сильно повлияла на туристическую отрасль и ее последствия с точки зрения концептуальных основ устойчивого развития представляют значительный интерес.

Как показали результаты анализа устойчивый туризм в России только начинает развиваться. Индустрия нуждается в разработке и применении механизмов, способствующих достижению устойчивости. В свою очередь перенимание зарубежного опыта может ускорить этот процесс.

Литература

1. В ООН подсчитали потери туризма из-за пандемии коронавируса [Электронный ресурс]. – Режим доступа: <https://ria.ru/20200728/1575052058.html> (дата обращения: 10.12.21).
2. Устойчивый туризм. Академия. [Электронный ресурс]. – Режим доступа: <https://st.rostourunion.ru/kurs2022> (дата обращения: 09.03.2022).
3. UNWTO Sustainable Development Sustainable development. [Электронный ресурс]. – Режим доступа: <https://www.unwto.org/sustainable-development> (дата обращения: 09.02.22).
4. Our Common Future // Report of the World Commission on Environment and Development. – 1987. – 412 p.
5. Lenzen M., Sun Y., Faturay F., Ting Y., Geschke A., Malik A. The carbon footprint of global tourism // *Nature Climate Change*. – 2018. – V.8. – №6. – Pp. 522–528.
6. Kamb A., Lundberg E., Larson J., Nilsson J. Potentials for reducing climate impact from tourism transport behavior // *Journal of Sustainable Tourism*. – 2021. – V.29. – №8. – Pp. 1365–1382.
7. Jantzen J., Kristensen M., Christensen T. H. Sociotechnical transition to smart energy: The case of Samsø 1997–2030 // *Energy*. – 2018. – V.162. – №6. – Pp. 20–34.
8. Yilmaz Y., Unguren E., Kacmaz Y. Determination of Managers' Attitudes Towards Eco-Labeling Applied in the Context of Sustainable Tourism and Evaluation of the Effects of Eco-Labeling on Accommodation Enterprises // *Sustainability (Switzerland)*. – 2019. – V.11. – №11. – Pp. 1–27.
9. Rosato P.F., Caputo A., Valente D., Pizzi S. 2030 Agenda and sustainable business models in tourism: A bibliometric analysis // *Ecological Indicators*. – 2021. – V.121. – №106978. DOI: 10.1016/j.ecolind.2020.106978.

10. Skeiseid H., Derdowski L.A., Grahn A.H., Hansen H. Motivating sustainable change in tourism behavior: The first-and third-person effects of hard and soft messages // Sustainability (Switzerland). – 2019. – Vol.12. DOI: 10.3390/su12010235.
11. Ioannides D., Gyimóthy S. The COVID-19 crisis as an opportunity for escaping the unsustainable global tourism path // Tourism Geographies. – 2020. – V.22. – №3. – Pp. 624–632.

УДК 711;159.9.316.6;004

ИССЛЕДОВАНИЕ ПРИМЕНИМОСТИ ЦИФРОВЫХ СЕРВИСОВ В ГЕНДЕРНО-ОРИЕНТИРОВАННОМ ПРОЕКТИРОВАНИИ ГОРОДСКОЙ СРЕДЫ

Евсеева З.А.¹ (студент), Гудым А.А.¹ (студент)

Научные руководители – канд. псих. наук, доцент Шмелева И.А.¹,

канд. псих. наук Джумагулова А.Ф.¹

¹*Университет ИТМО*

e-mail: zlataevseeva007@gmail.com

В работе рассмотрена городская мобильность и транспортное поведение женщин и проведен обзор существующих интерфейсов, затрагивающих основные потребности женщин в городской среде, такие как безопасность, комфортная мобильность и ежедневное взаимодействие, возможность влияния на процессы в городе, с целью применения цифровых сервисов в гендерно-ориентированном проектировании городской среды.

Ключевые слова: городская среда, гендерно-ориентированные исследования, цифровые сервисы в городе, мобильность женщин, женщины.

В историческом контексте город был местом, строящимся мужчинами и предназначенный в первую очередь для мужчин, так как преимущественно они влияли на становление и развитие государства: например, владели землей и оружием, занимались военным делом, наследовали собственность, могли занимать государственные должности, получать образование и многое другое. Тогда как женщины почти не участвовали в городской жизни, в связи с частично или полностью ограничивающими их роль в общественной жизни социо-культурными, религиозными или законодательными исторически сложившимися нормами. Результатом этого длительного исторического процесса оказалось отсутствие учета гендерно-поведенческих особенностей женщин при проектировании городов, что создавало и продолжает создавать ряд проблем, в том числе связанных с безопасностью [1].

Актуальность темы исследования связана с назревшей необходимостью принимать во внимание потребности женщин в современном обществе, тем более в городах, где сегодня проживает более половины населения планеты [2]. Женщины ежедневно находятся в невыгодном положении, живя, перемещаясь, работая, отдыхая в пространстве города, не учитывающего их потребности и проблемы. Изменение городов в сторону создания более безопасной и инклюзивной среды, а также репрезентация образа женщины в этой среде – это длительный процесс, однако, используя современные цифровые технологии, появляется возможность изменения небезопасных и дискомфортных паттернов взаимодействия женщин с городской средой.

Мобильность женщин в городской среде

Городская мобильность является одной из ключевых проблем для женщин. Считается, что женщины, в среднем, более склонны пользоваться общественным транспортом в силу их большего участия в процессах заботы о детях и занятости домашними делами, из-за чего им приходится осуществлять несколько коротких поездок в течение дня [3]. Мужчины же больше склонны делать поездки на длинные дистанции, что считается связанным с поездками по работе. В то же время женщины предпочитают искать работу ближе к дому, что тесно связано с выполнением домашних обязанностей [4].

В одном из исследований в Испании было обнаружено, что женщины больше привержены устойчивой мобильности, к которой относятся велосипед, общественный транспорт, походы пешком, в то время как мужчины пользуются устойчивым транспортом только в 15% случаев. Таким образом, можно заметить, что качество общественного транспорта, согласно статистике, влияет больше на женское население города, чем на мужское. Авторы отмечают, что мобильность женщины кардинально меняется в зависимости от многих факторов (замужество, уровень образования, дохода, количество детей) в то время, как на мужчин это не имеет особого влияния.

Отдельно необходимо отметить ожидание транспорта на остановках, точках пересадки и в других местах [5]. Исследователи рассматривают долгое время ожидания как один из недостатков инфраструктуры, опираясь на психологические опросы, выявившие, что время в ожидании проводится тяжелее, чем нахождение в самом путешествии. Однако, женщины испытывали чувство страха больше, чем мужчины. В ходе работы было выяснено, что уменьшению ощущения страха на остановках способствует наличие карт с информацией о времени прибытия транспорта, карты района и схем остановок. Также уменьшению чувства тревоги способствуют аудиальные оповещения и табло с актуальным временем для транспортных средств.

Чувства страха и тревоги, ощущаемые женщинами субъективны, и усугубляются в связи с экономическим положением, расовым притеснением, уровнем преступности в стране и рядом других факторов. Однако, нельзя не уделять внимание данной проблеме и списать негативные эмоции женщин при передвижении по городу и ожиданию транспорта на самих женщин, не учитывая причин, почему эти страхи появляются. Так как нет четкого ответа на вопрос, почему женщины ощущают страх при использовании транспорта, а чувство беспокойства присутствует, авторы приходят к выводу, что необходимо заниматься решением данной проблемы, делая городскую среду более комфортной для женщин. Главными рекомендациями, помимо представленных выше, является улучшение работы городских властей и полиции, техническим же предложением является оснащение пространства рядом с остановками и дорогами специальными приборами для оповещения об опасности.

Транспортное поведение женщин

Репрезентативная выборка, представленная в работе китайских исследователей, описывает несколько поведенческих паттернов [6]. Например, женщины меньше мужчин путешествуют вечером и ночью. Авторы делают важное заключение о том, что традиционно публичные места больше связаны с проведением времени мужчинами и предназначенного для них же, в то время как женское пространство остается больше скрытым, «домашним». Отмечается интересная особенность – современные города не обеспечивают женщинам плавного перехода между работой и домом, поэтому стигмы, которые существуют в обществе, мешают женщинам полноценно пользоваться общественными пространствами. Важными предложениями по улучшению комфортности городов для женщин является плотная застройка, которая помогает женщинам чувствовать себя в безопасности, когда они находятся в окружении других людей. Также система транспортного обеспечения (изменение мест остановок, их техническое оснащение), дизайн аллей и улиц, все эти аспекты безопасного дизайна города могут сделать городскую среду для женщин более комфортной, что позволит женщинам чаще выходить на улицу, искать работу вдали от дома, не беспокоясь о своей безопасности во время пути на работу и обратно.

В исследованиях российских специалистов наибольшее внимание уделяется проблемам безопасности российских городов, в частности они в большей степени направлены на изучение ощущения безопасности на улице, нежели рассмотрения изменения городского дизайна с целью его большей комфортности для женщин [7, 8]. В ряде исследований методы социологических исследований выходят на передний план, что отличает их от подхода иностранных специалистов. В статье А. М. Веркеева рассматриваются различия в восприятии безопасности города разными группами людей, в зависимости от пола и возраста (женщинами, мужчинами и пожилыми людьми). Так, женщины и люди старшего возраста больше склонны испытывать беспокойство при передвижении по городу. Женщины в старшей возрастной группе также больше испытывают страх, чем мужчины в той же возрастной группе [7].

В статье Багиной Я.А рассматривается страх, испытываемый женщинами в городском пространстве, что значимо в контексте городской феминисткой географии и криминалистики [8]. Подход в данном исследовании был основан на восприятии города через пространственные истории, с помощью анализа которых автору удалось увидеть модели поведения, которые женщины используют при столкновении с опасной, в их восприятии, ситуации. Центр города, например, считается большинством женщин безопасным, а переулки с малым количеством людей и недостаточным освещением воспринимаются более тревожными местами.

Исследование «Гендерные установки в дизайне города», в котором авторы рассматривают городскую среду в контексте феминности и маскулинности показывает, что город не является объективным по своей сути, в нем крайне важным оказывается субъективный компонент восприятия [9]. Авторы приводят различные примеры элементов города, которые могут восприниматься с точки зрения гендерных установок. Таким образом, зачастую административные здания воспринимаются маскулинными в силу их дизайна, а жилая среда считается больше феминной. Исследователи говорят о возможности создания более нейтральной городской среды, в которой признается различие и равноценность пространств с гендерной точки зрения, и реализован постепенный отход от гендерных установок.

Обобщая результаты международных и российских исследований городов с учетом потребностей женщин, можно сделать вывод о том, что концепция адаптации городской среды для нужд женщин не является характерной чертой для какой-то отдельной страны или региона. Мировая практика показывает, что восприятие женщинами городской среды отличается от мужского, поэтому необходимо уделять больше внимания учету мнения женщин в вопросах о городских решениях.

Следует также обратить внимание на исследование харрасмента (домогательства, приставания) к женщинам, не только в качестве пассажирок, но и в роли работниц на транспорте [10]. Так в ходе исследования, проведенного Ассоциацией транспортных работников в Европе, было выявлено, что 69% женщин-работниц заявили о харрасменте со стороны клиентов, а 39% со стороны коллег. Таким образом, сама организация транспортной системы не является гендерно нейтральной. Для того, чтобы сделать передвижение более гендерно инклюзивным авторы статьи предлагают уменьшить число пересадок, чтобы сократить стоимость поездок и минимизировать время ожидания, которое может также вызывать тревогу у женщин, особенно в темное время суток. Специальные сервисы на смартфонах для транспорта с целью оповещения об опасности или подозрительном поведении могут быть инструментами для повышения комфортности перемещения по городу. Также автор упоминает «розовый транспорт» для женщин, что является спорным решением, так как не решает проблему неравенства, в целом, но делает путешествие по городу более комфортным. Такое такси успешно функционирует в городах по всему миру [11].

Подводя итог, можно сказать, что женщинам приходится регулировать свое передвижение по городу в большей степени опираясь не на удобство, а скорее на безопасность выбранного маршрута. Определенные сервисы могут стать опорным пунктом в решении гендерных проблем городской среды, однако, многое зависит от той политики, которые изберут власти и активные члены общества.

В современном мире у каждой горожанки существует потребность улучшения качества жизни в городе, что естественным образом включает нахождение в доступной, комфортной и безопасной городской среде, ведь «здания, архитектура, дизайн и доступ к общественному транспорту самым базовым образом влияют на имеющиеся у жителей возможности передвижения, доступа к услугам, получения работы, ощущения себя в безопасности, самовыражения, ведения частной жизни, и в целом создания качественных условия для жизни» [12].

Обзор существующих интерфейсов для женщин в городе

Преобразования городской среды в инклюзивную, дружелюбную и безопасную, способную улучшить жизнь каждого горожанина – часто есть результат длительных процессов, тогда как сталкиваться с дискомфортом и опасностью жителям и жительницам города приходится ежедневно.

Существующие современные цифровые технологии призваны упростить и улучшить ежедневную жизнь. Цифровые сервисы и возможности их использования (интерфейсы) способны улучшать взаимодействие горожан с городской средой, благодаря наделению пользователей свободой построения этого взаимодействия: например, они помогают создавать лучшие маршруты или сообщать в администрацию о волнующих проблемах в общественных пространствах.

При разработке сервисов и интерфейсов к ним в контексте взаимодействия с городской средой редко используется гендерный подход, но некоторые сервисы обладают рядом необходимых функций, способных улучшать взаимодействие с окружающей средой различных категорий женщин. В рамках исследования был проведен общий обзор существующих сервисов, для удобства классифицированных на группы, затрагивающие основные потребности женщин в городской среде: безопасность, комфортная мобильность и ежедневное взаимодействие, возможность влияния на процессы в городе (рисунок). Рассмотрим некоторые из них.

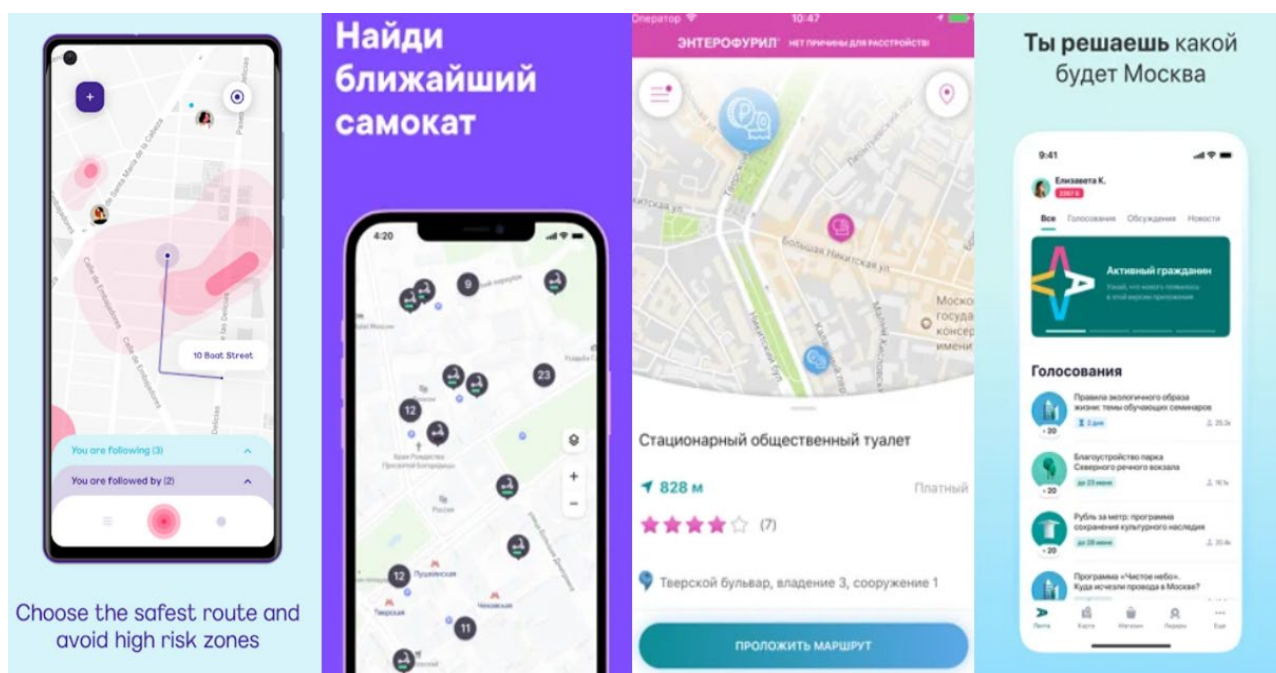


Рисунок. Примеры интерфейсов, затрагивающие основные потребности женщин в городе

Интерфейсы сервисов для безопасности при навигации в городе

В основном для проектирования интерфейсов для безопасности при навигации в городе используется один из двух принципов: спасение или контроль. К первому типу относятся, например, приложения-тревожные кнопки, отправляющие данные пользователя и его геопозицию в местную службу спасения и ряду выбранных им контактов [13]. Также среди «тревожных» интерфейсов существуют и сервисы, которые помогают защитить себя в режиме «здесь и сейчас» с помощью включения различных громких звуков или сирены [14]. Кроме интерфейсов-тревожных кнопок в мире распространены интерфейсы контроля, помогающие группе людей отслеживать местоположение друг друга на карте [15]. Самые функциональные интерфейсы в категории личной безопасности в городском пространстве располагают функциями и спасения, и контроля.

Из общих недостатков, присущих интерфейсам категории безопасности в контексте повсеместного ежедневного использования в городской среде, можно выделить их точечную специализированность и специфичность целевой аудитории. Проблемой также является их низкая распространенность и адаптированность в России: так, например, «тревожные» интерфейсы широко распространены в европейских странах и Америке, являются простыми и действенными, однако, почти не используются в России. Для того, чтобы стать полезнее и охватывать большее количество пользователей сервисам необходимо добавить мультифункциональность и увеличить возможности использования интерфейса.

Интерфейсы сервисов для комфортной мобильности

Мобильность в городской среде должна быть не только безопасной и доступной, но и удобной. В России работает достаточное количество сервисов, способных сделать перемещение граждан более комфортным. Например, к таким интерфейсам относятся навигационные мобильные приложения для контроля времени прибытия общественного транспорта [16], или sharing-интерфейсы, с помощью которых любой пользователь может арендовать велосипед или электросамокат в специальных городских точках, используя только свой телефон [17]. Для женщин такие sharing-интерфейсы имеют большое значение, позволяя ускорить перемещение в городе, а также повысить удовлетворенность от нахождения в пути. Кроме того, перемещение на электросамокате и велосипеде повышает доступность городской среды в вечернее и ночное время суток за счет увеличения скорости перемещения.

Интерфейсы сервисов комфортной мобильности повсеместно распространены в России и почти каждый из них в своей сфере влияния способен повысить доступность городской среды для широкого круга населения. Главным из основных недостатков таких интерфейсов является их сосредоточенность на краткости маршрутов и игнорирование потребностей в удобстве и безопасности при городских перемещениях, в частности для женщин-пешеходов. Создание нового комплексного навигационного интерфейса с учетом потребностей женщин и реализующего лучшие из существующих практик, способно решить эти проблемы.

Интерфейсы сервисов для комфортной жизни в городе

В категорию интерфейсов сервисов, способных сделать нахождение в городе более комфортным, входят самые разнообразные интерфейсы: от навигационных интерфейсов карт с отметками общественных туалетов, [18] до государственных интерфейсов приложений, помогающих жителям во взаимодействии с городской средой [19]. Интерфейсы сервисов для комфорта в городской среде, существующие в России, чаще всего разработаны городской администрацией для взаимодействия с рядом государственных общественных учреждений и не затрагивают взаимодействие с улицами и другими общественными пространствами. К тому же многие из них технически не адаптированы под нужды современного пользователя, тогда как частные инициативы обычно узкоспециализированы, малоизвестны и затрагивают только одну область, а также не способны оказывать сильного влияния на ежедневную жизнь большого количества граждан.

Интерфейсы сервисов для городского аудита и участия граждан в городском управлении

Аудит общественных пространств является общественно значимой функцией, помогающей улучшить взаимодействие горожан не только с городской средой, но и с органами городского управления. Для женщин существование таких сервисов является крайне важным, потому что помогает донести их проблемы и потребности в городе до тех, кто может помочь их исправить. Такие сервисы позволяют жителям оставлять жалобы по поводу сломанной детской площадки, выбоин на дороге, актов вандализма и всего того, что портит комфортную и доступную жизнь в городе [19].

Сейчас в России не существует популярного, удобного и грамотного интерфейса сервиса для аудита городской среды обычными гражданами. Работающие государственные сервисы, создающие коммуникацию между горожанами и местной администрацией, сосредоточены только в некоторых крупных городах. Не считая московского сервиса «Активный гражданин» [20], в России не существует интерфейсов сервиса для возможности оказания влияния на разрабатываемые в городах планировочные решения общественных пространств. Разработка таких сервисов является актуальной задачей создания комфортной и инклюзивной городской среды.

К сожалению, изменения во внешнем облике городов по модели городского управления “Top-Down” (инициатива, исходящая от органов власти) происходят недостаточно быстро и масштабно, а гендерно-нейтральный подход к проектированию не используется повсеместно. Эта проблема определяет в обществе новый актуальный запрос на самостоятельное повышение доступности города через изменение горожанами способов взаимодействия в условиях текущей дискомфортной среды. Как показало наше исследование, такое изменение может

происходить с помощью современных интерфейсов цифровых сервисов. Кроме того, современные технологии позволят привлечь внимание общественности к необходимости учета женского взгляда при проектировании городской среды.

Литература

1. Hirt S. A. Stuck in the suburbs? Gendered perspectives on living at the edge of the post-communist city // *Cities*. – 2008. – Vol.25. – No.6. – Pp. 340–354.
2. Handbook for Gender-Inclusive Urban Planning and Design. [Электронный ресурс]. – Режим доступа: <http://worldbank.org/en/topic/urbandevelopment/publication/handbook-for-gender-inclusive-urban-planning-and-design> (дата обращения: 14.12.2021).
3. Rethinking public transportation for women's safety and security. [Электронный ресурс]. – Режим доступа: <https://sustainablemobility.iclei.org/rethinking-public-transportation-for-womens-safety-and-security/> (дата обращения: 11.11.2021).
4. Women's Budget Group. Public transport and Gender. A pre-budget briefing from the UK Women's Budget. – 2020. – 7 p.
5. Chowdhury S., Van Wee B. Examining women's perception of safety during waiting times at public transport terminals // *Transport Policy*. – 2020. – No.94. – Pp. 102–108.
6. Zhang M., Zhao P., Tong X. Constructing women's immobility: Fear of violence and Women's constricted nocturnal travel behaviour // *Travel Behaviour and Society*. – 2022. – Vol.26. – Pp. 178–192.
7. Веркеев А.М. Неравенство в восприятии (у)личной безопасности в России// *Журнал социологии и социальной антропологии*. // *Журнал социологии и социальной антропологии*. – 2021. – №24(3). – С. 169–192.
8. Багина Я.А. Страх и тревога как часть женских пространственных историй в городе. // *Интеракция. Интервью. Интерпретация*. – 2019. – Т.11. – №17. – С. 46–60.
9. Михайлов С. М., Михайлова А.С., Белов М.И., Филимонова Ю.М. Гендерные установки в дизайне города // *Дизайн-Ревю*. – 2015. – №1–4. – С. 18–55.
10. Pirra M., Kalakou S., Carboni A., Costa M., Diana M., & Lynce A.R. A preliminary analysis on gender aspects in transport systems and mobility services: Presentation of a survey design // *Sustainability*. – 2021. – Vol.13(5). – Pp. 1–20.
11. 7 cities with taxis and subway cars only for women. [Электронный ресурс]. – Режим доступа: <https://www.businessinsider.com/transportation-options-only-for-women-2015-12#pink-taxi-in-cairo-4> (дата обращения: 11.12.2021).
12. Dish A.T., Damien C., Reardon M., Anderson T. Architectural & Urban Design // Raoul Wallenberg Institute of Human Rights and Humanitarian Law. [Электронный ресурс]. – Режим доступа: <https://rwi.lu.se/> (дата обращения: 13.12.2021).
13. Red Panic Button App. [Электронный ресурс]. – Режим доступа: <https://www.redpanicbutton.com> (дата обращения: 13.12.2021).
14. 10 лучших приложений личной безопасности для женщин. [Электронный ресурс]. – Режим доступа: <https://triu.ru/10-luchshih-prilozhenij-lichnoj-bezopasnosti-dlya-zhenshhin-android> (дата обращения: 13.12.2021).
15. Лучшие приложения для родителей. [Электронный ресурс]. – Режим доступа: <https://findmykids.org/blog/ru/7-luchshikh-prilozheniy-dlya-roditeley> (дата обращения: 13.12.2021).
16. Каким приложением-навигатором лучше пользоваться. [Электронный ресурс]. – Режим доступа: <https://club.dns-shop.ru/blog/t-328-prilojeniya/46496-kakim-prilojeniem-navigаторom-luchshe-polzovatsya> (дата обращения: 13.12.2021).
17. 7 сервисов аренды самокатов в Москве и других крупных городах. [Электронный ресурс]. – Режим доступа: <https://journal.tinkoff.ru/short/beri-i-katis/> (дата обращения: 13.12.2021).
18. Интерактивная карта общественных туалетов «Водоканала». [Электронный ресурс]. – Режим доступа: <http://www.vodokanal.spb.ru/kanalizovanie/tualety/> (дата обращения: 13.12.2021).

19. 17 онлайн-ресурсов, принимающих жалобы граждан. [Электронный ресурс]. – Режим доступа:
https://pikabu.ru/story/17_onlaynresursov_prinimayushchikh_zhalobyi_grazhdan_7997374
(дата обращения: 13.12.2021).
20. Услышать каждого: как проект «Активный гражданин» помогает жителям Москвы. [Электронный ресурс]. – Режим доступа: <https://lenta.ru/articles/2020/12/01/ag/>(дата обращения: 13.12.2021).

УДК 004.514

ИССЛЕДОВАНИЕ ПАТТЕРНОВ ВЗАИМОДЕЙСТВИЯ ПОЛЬЗОВАТЕЛЕЙ С СЕРВИСАМИ РЕКОМЕНДАЦИЙ

Ермолаева А.А.¹ (студент), Карсаков А.А.¹

Научный руководитель – канд. филос. наук Смолин А.А.¹

¹Университет ИТМО

e-mail: alinaermolaeva044@gmail.com

В работе рассмотрены процессы проектирования интерфейса новой версии приложения с рекомендательным функционалом, которое помогает студентам в поиске тем для своих НИР и НИРМА, научных руководителей, а также единомышленников в работе над научными работами, а научным руководителям - в поисках студентов для своих научных работ. А также приведены результаты тестирования новой версии и нового функционала на первых пользователях.

Ключевые слова: рекомендательный контент, лента, ведение документации, интерфейс, мобильное приложение.

Современный человек с каждым днем все чаще сталкивается с рекомендательными системами. На сегодняшний день тяжело представить приложение, которое не использовало бы в своем функционале рекомендательных алгоритмов. Это связано с тем, что система рекомендаций решает одну из основных проблем потенциальных пользователей – увеличивает скорость нахождения искомого. Рекомендуемые продукты или элементы контента особенно распространены на сайтах электронной коммерции, социальных сетях, новостных сайтах, а также в сервисах потокового видео или музыки, но их также можно найти на сайтах других жанров [1]. Также набирает популярность тенденция использовать рекомендации в приложениях, которые помогают пользователям в работе над различными проектами. Это позволяет оптимизировать поиск нужной информации, её хранение, структурирование и совместное использование.

Старая версия приложения фокусируется на событиях личных, пользовательских и университетских, а также на интересах пользователя (рис. 1). Имеет светлую и темную тему, что очень удобно для сторонников каждой версии и просмотра в разное время суток.

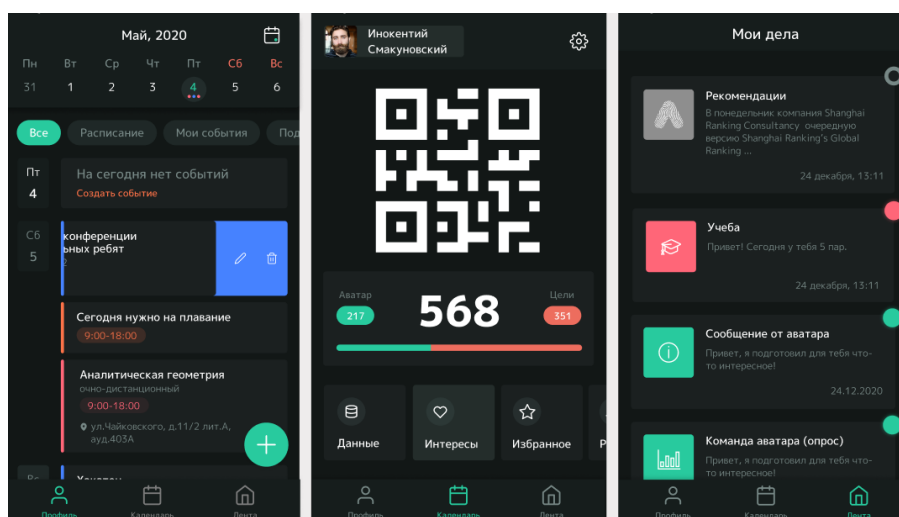


Рис. 1. Предыдущая версия приложения Аватар.ИТМО

Приложение разделено на три основных раздела: «Лента», «Календарь», «Профиль». Такое разделение обусловлено тремя основными задачами приложения: рекомендация статей и мероприятий, на основании интересов пользователя, тайм-менеджмент жизни в университете и рейтинг среди других пользователей приложения.

В разделе «Лента» пользователю рекомендовались статьи, события в университете, а также в этом разделе находились уведомления пользователя сообщения от Аватара. Данные рекомендации имеют формат stories. При открытии какого-либо уведомления или статьи есть возможность листать контент нажав вправо/влево. Такое проектирование ленты с точки зрения удобства пользователя неудобно по нескольким причинам:

1. В одном месте смешаны несколько видов контента: уведомления, опросы, игры и контент, что путает пользователя, не дает четкого понимания цели посещения данного раздела.
2. «Закрытый» формат ленты, при котором пользователю необходимо просматривать каждую единицу контента, прежде чем перейти к следующей не дает пользователю «сквозного» просмотра, как при классической ленте, что создает риск быстрого утомления от такой подачи формата или надоедания.

Раздел «Календарь» спроектирован по типичному для других приложений шаблону.

Раздел «Профиль» хранит пользовательскую информацию и основные пользовательские настройки. На главном экране мы можем увидеть QR-код для входа пользователя в университет, индивидуальный рейтинг, настройки и подразделы «Данные», «Интересы», «Избранное», «Рейтинги», «Отзыв». На виджете с индивидуальным рейтингом находится рейтинг целей и Аватар. Визуализация рейтинга интуитивно непонятна и может быть воспринята пользователями по-разному. Зайдя в раздел «Рейтинги», можно увидеть рейтинг пользователя относительно других пользователей, однако, отсутствует объяснение на основе чего строится этот рейтинг, что может ввести пользователя в заблуждение.

Таким образом, данная версия приложения имеет ряд существенных недостатков в проектировании. Наиболее важными из них являются: «закрытый» тип ленты, отсутствие описания функциональных возможностей приложения (пользователь вынужден догадываться о тех или иных функциях приложения), отсутствие ценностного предложения по сравнению с другими приложениями на рынке.

Функционал новой версии Аватар.ИТМО как раз фокусируется на помощи в ведении научных или учебных проектов в университете, используя также и рекомендательный контент.

Первым этапом определения пользовательских требований было интервью с потенциальными пользователями, на основе которого была составлена Value Proposition Canvas (далее VPC). С помощью VPC мы сможем определить основные боли и проблемы пользователя, что существенно упростит продуктивное проектирование новой версии. Также такой подход поможет сконцентрироваться на конечных предполагаемых результатах и использовании продукта конечным пользователем.

Исходя из глубинных интервью удалось определить такие задачи пользователя:

- поиск научного руководителя для работы над НИР и НИРМА;
- возможность совместной работы над НИР и НИРМА с научным руководителем и другими участниками;
- поиск статей, на тему схожую с темной НИР и НИРМА;
- совместное хранение и структурирование материалов, связанных с НИР и НИРМА, а также личных материалов;
- создание проектов и совместной работы над ними;
- просмотр возможных тем и возможных научных руководителей и партнеров в работе над НИР, НИРМА, или другими проектами на интересующие темы.

Также были определены пользовательские боли – причины, по которым они сталкиваются со сложностями в выполнении поставленных задач:

- сложность в выборе научного руководителя без четкого понимания о его компетенциях и профиле работы;
- сложность поиска единомышленника в работе над различными проектами;

- сложность в поиске статей на интересующие научные темы;
- недостаточная информация о существующих научных и не только проектах, к которым можно присоединиться;
- недостаточная информация о темах потенциальных научных руководителей.

Последнее, что было выявлено из глубинных интервью, это выгоды – то, что пользователи хотели бы увидеть в приложении, которое поможет в решении их задач:

- связь с преподавателем и участниками проекта в одном месте;
- настройка рекомендаций научных статей;
- просмотр научных достижений других студентов и преподавателей;
- поиск актуальных мероприятий и вакансий;
- возможность бронирования аудиторий для работы над проектами;
- возможность иметь ассистента, который сможет быстро ответить на интересующие вопросы;
- получать комментарии преподавателя по задачам.

На основании полученной информации было сформировано ценностное предложение – сервисы и услуги, выгоды «болеутоляющие», которые помогут наиболее полно покрыть потребности пользователя.

Сервисы и услуги:

- лента новостей, где рекомендуются актуальные темы НИР, проектов, а также научные руководители и возможные участники;
- общее пространство, в котором идет совместная работа над одним проектом в виде простых документов с возможностью их выгрузки в отдельные документы;
- текстовый редактор с возможностью добавления туда ссылок с превью, фотографий, документов и статей из ленты.

«Болеутоляющие» - функции, которые помогают пользователю наиболее удобно решать свои задачи без ранее изложенных сложностей.

1. Описание проектов, над которыми в данный момент планирует работать или уже работает преподаватель.
2. Рекомендация в ленте тем, на которые преподаватель ищет студентов.
3. Рекомендация других студентов и аспирантов, как потенциальных участников проектов.
4. Лента рекомендаций статей на выбранные темы.
5. Поиск научных статей в приложении.
6. Рекомендации существующих проектов на интересующие темы в ленте рекомендаций.
7. Рекомендации тем научных руководителей в ленте новостей, а также в профиле научных руководителей.

Выгоды:

- создание чата, где можно общаться на предмет различных тем и проектов;
- создание бота Аватар, который отвечает на вопросы пользователей, а также рекомендует те или иные действия на основании знаний о пользователе.

Новая версия приложения (рис. 2) была разделена по основным направлениям на 5 разделов: проекты, черновики, профиль, лента и коммуникатор.

В разделе «Лента» находятся рекомендации тем, на которые преподаватель ищет студентов, рекомендация других студентов и аспирантов, как потенциальных участников проектов, лента рекомендаций статей на выбранные темы, поиск научных статей в приложении, других пользователей и проектов, рекомендации существующих проектов на интересующие темы в ленте рекомендаций, рекомендации тем научных руководителей в ленте новостей, а также в профиле научных руководителей.

В разделе «Проекты» находятся проекты, над которыми пользователь в данный момент ведет работу.

В разделе «Черновики» находятся документы, которые пользователь создает для хранения своих идей и записей для будущих проектов.

«Коммуникатор» – это раздел с чатами с другими пользователями, технической поддержкой, а также чата с Аватаром – ботом, который отвечает на интересующие вопросы и дает рекомендации действий.

«Профиль» – это раздел с персональной информацией о пользователе, а также пользовательскими настройками и настройками приложения.

На следующем этапе было спроектировано взаимодействие с приложением на прототипах низкого уровня точности.

После этого была сформирована дизайн-система, по принципу «атомарного дизайна» [2].

Атомарный дизайн — нелинейный процесс. Это как мыслительная модель, которая помогает нам думать обо всем интерфейсе, но не забывать о деталях. Каждый из пяти уровней играет свою роль в иерархии систем проектирования интерфейсов [2].

Создание страниц интерфейса производилось в программе Figma.

Раздел «Лента» теперь имеет открытый вид, в котором с помощью карточек визуализированы различные виды контента: проекты, другие пользователи и статьи. Чтобы увидеть детальную информацию каждого контента, необходимо нажать на него. Также в новой версии появилась возможность поиска по рекомендациям приложения, где можно найти пользователей, проекты и статьи.

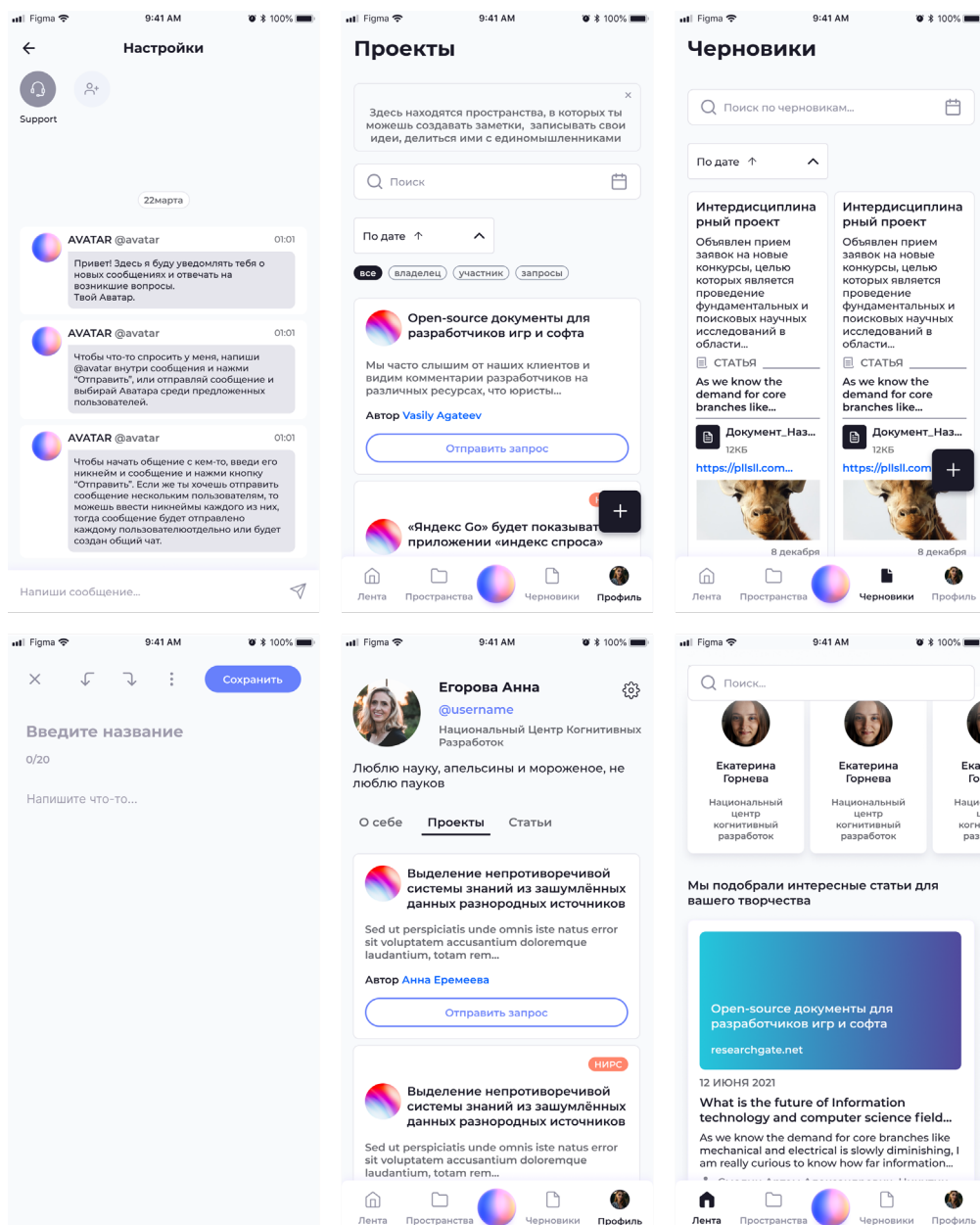


Рис. 2. Новая версия приложения

Каждую статью можно оценить, убрать из рекомендаций и добавить к себе в документ. Также каждый тип рекомендаций можно раскрыть полностью и посмотреть все рекомендуемые типы контента. Такое проектирование ленты имеет ряд существенных преимуществ перед предыдущей версией: привлечение внимания пользователя, более быстрая скорость поиска информации. Также и изменился предлагаемый контент. В новой версии рекомендуемая информация базируется на выбранных пользователем интересах и призвана помочь в работе над проектами.

Был добавлен новый раздел «Черновики» и возможность создания и редактирования документов- текстовый реактор. Раздел «Черновики» предназначен для хранения черновых вариантов документов для НИР, НИРМА и других проектов, а также в качестве использования как заметок для создания, записи и структурирования различных идей. Также в данном разделе есть возможность добавления нового черновика и поиска и фильтрации по дате и ключевым словам. Каждый отдельный черновик можно удалить, отредактировать и изменить цвет его превью для выделения на фоне других с помощью панели инструментов.

Экран создания и редактирования заметки или черновика – это текстовый редактор с базовыми функциями редактирования текста и возможностью вставки картинок, фотографий, статей из ленты, ссылок, ссылок с превью.

Раздел «Проекты» состоит из различных рабочих пространств, в которых пользователь работает над НИР, НИРМА или другими проектами. Их можно сортировать по дате или фильтровать по роли в проекте. Каждый проект состоит из документов (заметок) с различной информацией, над которым работает как один участник, так и несколько. Каждый член команды может просматривать все документы. У проекта могут быть владелец, администратор, участник и подписчик.

Коммуникатор – это раздел для общения пользователя с личным ассистентом – Аватаром и другими пользователями. На главном экране находится чат с Аватаром, в нем можно задавать вопросы, а также туда приходят важные уведомления. На эти сообщения можно отвечать и удалять их. Перейдя во все чаты, можно посмотреть всех своих собеседников или найти новых.

Далее экспериментальным путем сравнивалась эффективность взаимодействия пользователей с каждой из систем.

Гипотеза: в новой версии интерфейса пользователи потратят больше времени на выполнение задач, чем в старой; нулевая гипотеза: пользователи потратят одинаковое количество времени на решение задач в обеих версиях интерфейса.

Описание процедуры проведения эксперимента:

1. Ознакомительная анкета.
2. Каждому пользователю даются 3 задачи в каждой версии приложения. Измеряется время выполнения поставленных задач.
3. Опрос пользователей о недостатках и достоинствах приложения.

В пилотном эксперименте принимали участие 7 человек. Необходимый размер выборки был посчитан по формуле:

$$n = \frac{z^2}{4 * (p_1 - 0.5)}$$

В таблице 1 приведены результаты пилотного эксперимента.

Таблица 1. Результаты пилотного эксперимента

Желаемая статистическая значимость	Желаемая мощность	z	Размер выборки
1%	80%	2,49	24

Основной эксперимент проводился по методу теста знаков на выборке из 24 человека. Его результаты приведены в таблице 2.

Таблица 2. Результаты основного эксперимента

Количество ненулевых разностей	Размер выборки	Количество успехов	Количество неудач	p-value	Желаемая статистическая значимость	Желаемая мощность
21	24	17	4	0,36%	1%	80%

Таким образом, был получен статистически значимый результат, при котором нет оснований отвергать гипотезу о том, что пользователи тратят меньше времени на выполнение задач в новой версии. В следующих экспериментах планируется расширение гипотезы: проверка количества ошибок пользователя и успешность выполнения задач

Литература

1. Рекомендательные системы: как помочь пользователю найти то, что ему нужно? [Электронный ресурс]. – Режим доступа: <https://vc.ru/marketing/152926-rekomendatelnye-sistemy-kak-pomoch-polzovatelyu-nayti-to-cto-emu-nuzhno> (дата обращения: 25.02.2022).
2. Глава 2: Методология атомарного дизайна (часть 1). [Электронный ресурс]. – Режим доступа: <https://medium.com/atomic-design-part-2-1-eab69ea66560> (дата обращения: 25.02.2022).

УДК 004.03

РЕИНЖИНИРИНГ БИЗНЕС-ПРОЦЕССА ОПЕРАТОРА СВЯЗИ: ПРЕДПРОЕКТНОЕ ИССЛЕДОВАНИЕ

Зайцева И.С.¹ (студент)

Научный руководитель – зам. директора Трутнев Д.Р.²

¹*Университет ИТМО*

²*Центр технологий электронного правительства, CISA, PRINCE2, Agile Practitioner, PMP*

e-mail: zaitseva.itmo@yandex.ru

Материал представляет собой предпроектное исследование реинжиниринга деятельности предприятия, оказывающее телекоммуникационные услуги правительству г. Санкт-Петербурга. Описание задач, которые должен решить реинжиниринг.

Ключевые слова: телекоммуникации, бизнес-процесс, оператор связи, реинжиниринг, правительство города, предпосылки реинжиниринга, Санкт-Петербург.

СПб ГУП «АТС Смольного» (Предприятие) – оператор услуг телефонной связи и передачи данных г. Санкт-Петербурга. Предприятие является одним из старейших в Санкт-Петербурге и является подведомственным учреждением Комитета по информатизации и связи. На текущий момент предприятие обеспечивает услугами связи как правительство города, так и организации, связанные по роду своей деятельностью с органами государственной власти. Участвует в решении стратегически важных для города задач, таких как обеспечение бесперебойной трансляции на выборах, едином государственном экзамене, городские праздники, колл-центры в поликлиниках, единая региональная справочная служба 122.

Актуальной составляющей для качественного оказания услуг является эффективно выстроенные бизнес-процессы Предприятия.

Однако, на текущий день при оказании услуг связи клиентам предприятие испытывает ряд недочетов. Так для подключения телефонного номера клиент должен подать официальное письмо, которое регистрируется во внутренней системе документооборота предприятия (СЭД). Далее письмо рассматривается в отделе по работе с клиентами, создается дополнительная заявка на оказание услуг в рамках которой проверяется техническая возможность подключения, координируются работы по подключению, подписывается договор. При необходимости для подключения получить оборудование со склада требуется подать заявку сотруднику склада, который совершит необходимые операции в 1С, а если оборудование требуется купить, то это также оформляется отдельной заявкой, или проводится конкурсная процедура. Данный механизм взаимодействия приводит к увеличению сроков подключения, затруднено взаимодействие между сотрудниками, различные нюансы по подключению теряются при переносе одной системы в другую.

Для построения аналитики требуется сопоставлять данные из разных систем и приводить их в соответствие. При этом требуется большое количество времени для проведения контроля за процессом и оценки результативности процесса. Понятие оперативный контроль в данной ситуации является условным понятием.

Впервые термин реинжиниринг встретился в трудах Майкл Хаммер и Джеймс Чампи, которые описали основные принципы и возможности реинжиниринга. Авторы определяют реинжиниринг как «фундаментальное переосмысление и радикальное перепроектирование

бизнес-процесса для достижения существенных улучшений в таких ключевых для современного бизнеса показателях результативности, как затраты, качество, уровень обслуживания и оперативность» [1].

Рассмотрим реинжиниринг бизнес-процесса оператора связи, как возможность решить проблемы качественного оказания услуг связи и получения возможности дополнительного извлечения прибыли путем подключения большего количества клиентов.

Существует немало методов реинжиниринга. В.В. Репин и В.Г. Елиферов – предлагают в качестве методов классификацию методик анализа процессов [2]. Наиболее полно методы реинжиниринга бизнес-процессов рассмотрены в работах Б. Андерсена [3]. Б. Андерсен предлагает не только большое разнообразие методов реинжиниринга, но и их широкую классификацию. Например, по стадии совершенствования, по основным назначениям инструментов, по масштабам воздействия от их применения, по затраченному времени на применение метода. Организационные методы Б. Андерсен выделяет отдельной группой.

Описанные им методы горизонтального и вертикального сжатия бизнес-процесса, методы интеграции бизнес-процесса, устранения временных разрывов и дублирования, внедрение типовых бизнес-процессов, алгоритмизации процесса, уменьшение количества входов и выходов бизнес-процесса, анкетирование, интервьюирование, беседа, опрос, мозговой штурм, совещание, матрица, матрица приоритетности процессов, различные виды матриц текущих и желаемых состояний процесса, диаграмма Парето (выявление ключевых проблем), метод постановки новых целей Х. Ясухисо, метод постановки целей SMART позволяют не только выбрать подходящий из методов заинтересованным в реинжиниринге предприятиям, но и подойти к вопросу реинжиниринга комплексно, используя весь спектр разнообразия применительно к проблемным областям.

Для определения задач, которые должен решить реинжиниринг бизнес-процесса СПб ГУП «АТС Смольного», была проведена общая оценка процессной деятельности предприятия: выполнено сопоставление текущих процессов Предприятия с эталонной моделью бизнес-процессов в области телекоммуникаций eTOM, выполнена оценка системы управления бизнес-процессами процессами по методике В.В. Репина [4]. Проведен анализ бизнес-процессов на наличие разрывов.

Модель eTOM (enhanced Telecom Operations Map) — расширенная карта бизнес-процессов оператора связи – является эталоном для телекоммуникационных компаний. На рисунке представлена общая карта eTOM [5].

Выполнено сопоставление бизнес-процессов Предприятия на схему eTOM. При этом взято за аксиому, что отсутствие бизнес-процессов в какой-либо части карты означает, что деятельность в данном направлении Предприятием не осуществляется. Несколько процессов в каком-либо блоке карты eTOM предполагает, что деятельность в данном направлении ведется, но является слабо выраженной. Сопоставление с эталонной моделью eTOM в части развития и повышения эффективности бизнес-процессов Предприятия продемонстрировало, что отсутствуют методы и подходы в управлении развитии услуг и ресурсов, что не позволяет предлагать клиенту гибкие тарифы, эксклюзивные предложения, или современные услуги связи.

Также не развиты процессы, описывающие отношения с поставщиками, разработки стратегии и мероприятий, направленных на анализ эффективности предприятия и оценки рисков. Примерами отсутствующих процессов могут быть процессы управления проблемами оказания услуг, что может позволить увеличить качество оказываемых услуг. Процессы управления ресурсами, управление рисками и управление эффективностью, в том числе планирование и стратегическое развитие, управление имиджем предприятия также актуальны для повышения качества оказываемых услуг. Отсутствие стратегических целей предприятия, планов развития в области ресурсов, услуг приводит к потерям прибыли и не позволяет работать над увеличением прибыли. Таким образом, одной из задач реинжиниринга бизнес-процессов является внедрение новых процессов по развитию клиентских услуг, направленных на дополнительное увеличение прибыли и сопутствующие оказанию услуг связи процессы стратегического планирования и управления рисками.

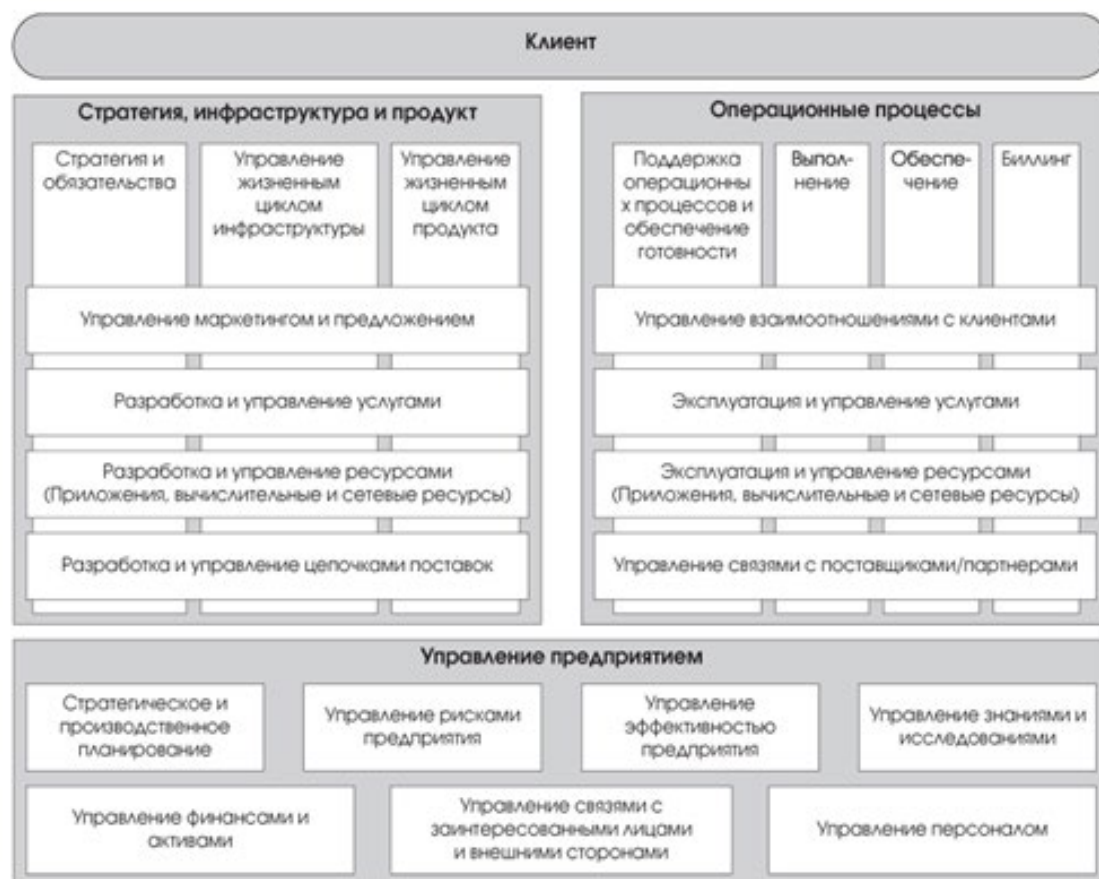


Рисунок. Схема eTOM (источник)

Для выявления следующей задачи реинжиниринга была проведена комплексная оценка системы управления бизнес-процессами Предприятия (СУБП) по методике предложенной В.В. Репиным.

Данная методика включает оценку СУБП по 10 направлениям таким как – архитектура бизнес-процессов, управление бизнес-процессами по целям и показателям, система стимулирования руководителей на улучшение бизнес-процессов по коэффициентам показателей эффективности (КПЭ), практика описания и анализа бизнес-процессов, практика оптимизации бизнес-процессов и внедрения изменений, автоматизация бизнес-процессов, стандартизация бизнес-процессов, контроль и аудит бизнес-процессов, корпоративная система обучения персонала методам процессного управления, процессный офис. Для каждого из направлений были определены критерии, их вес и значения. Такими критериями в направлении архитектуры бизнес-процессов, например, является критерий наличия реестра процессов, его границ и владельцев, а также архитектурные модели процессов. В качестве веса критериев взяты значения 0,2, 0,3 и. т. д, что в общей сумме составляет единицу для каждого из оцениваемых направлений. В качестве значений взяты значение отлично, хорошо, удовлетворительно, что позволяет рассчитать совокупное значение для каждого из направлений. Максимально возможный итоговый результат равен единице. Для увеличения качества оказания услуг и эффективности взаимодействия сотрудников при подключении услуг целесообразно проводить работу над направлениями, в результате расчёта которых получились ниже 0,5. Это такие критерии как осуществление регламентации и автоматизации процессов, оперативный контроль показателей результативности процесса, оценка рисков, введение практик рассмотрения и внедрения предложений по оптимизации.

Таким образом, еще одной из задач реинжиниринга является автоматизация бизнес-процессов, регулярная оптимизация процессов и внедрение практик внедрения предложений по оптимизации.

Как был приведен пример ранее для подключения услуг клиенту используется более трех информационных систем: СЭД, CRM (Customer Relationship Management) система, 1С, система расчета тарификации и. т. д. Помимо этого существуют и другие информационные системы для организации деятельности Предприятия. Проанализировав имеющийся реестр систем на предмет их автоматизации в той или иной системе можно привести такие цифры. Из 57 существующих бизнес-процессов двадцать семь 27 процессов содержат действия переноса данных из одной системы в другую, содержат взаимодействие в различных информационных системах, что увеличивает сроки выполнения работ и ресурсы, потерю информации при переносе.

Таким образом, устранение разрывов в процессах и создание единого информационного ресурса для координации работ по предоставлению услуг на высоком качестве является еще одной задачей реинжиниринга.

Итак, реинжиниринг бизнес-процесса оператора связи может быть направлен на решение проблемы качественного оказания услуг связи и повышение количества подключений за счет качества и разнообразия спектра услуг, позволяющих увеличить прибыль.

В рамках предпроектного обследования было выявлено, что реинжиниринг Предприятия должен включать:

- внедрение новых процессов по развитию клиентских услуг;
- развитие процессов стратегического планирования и управления рисками;
- автоматизацию бизнес-процессов;
- регулярную оптимизацию процессов и внедрение практик внедрения предложений по оптимизации;
- устранение разрывов в процессах;
- создание единого информационного пространства для обеспечения производственной деятельности Предприятия, в том числе и подключение услуг клиентам.

Литература

1. Хаммер М. Реинжиниринг корпораций: манифест революции в бизнесе / М. Хаммер, Дж. Чампи. — М: Манн, Иванов и Фербер. — 2006. — 287с.
2. Репин В.В. Процессный подход к управлению. Моделирование бизнес-процессов / В.В. Репин, В.Г. Елиферов. -3-е изд., испр. - М.: РИА «Стандарты и качество». — 2005. — 408 с.
3. Андерсен Б. Бизнес-процессы. Инструменты совершенствования / Бьерн Андерсен; пер. с англ. С.В. Ариничева; под науч. Ред. Ю.П. Адлер. - М.: РИА «Стандарты и качество». — 2005. — 272 с.
4. Репин В.В. Оценка зрелости системы управления бизнес-процессами компании. [Электронный ресурс]. — Режим доступа: https://www.businessstudio.ru/articles/article/otsenka_zrelosti_sistemy_upravleniya_biznes_prot_sessami_repin/ (дата обращения: 03.03.2022).
5. Официальный сайт JETINFO. [Электронный ресурс]. — Режим доступа: <https://www.jetinfo.ru/struktura-biznes-protsessov-dlya-otrasli-telekommunikatsionnykh-uslug/> (дата обращения: 06.03.2022).

УДК 004.891.2

РАЗРАБОТКА МЕТОДА ВЫДЕЛЕНИЯ ПРИЗНАКОВ ИЗ МЕТАГЕНОМНЫХ ОБРАЗЦОВ, ОСНОВАННОГО НА РАСКРАСКЕ ГРАФА ДЕ БРЮИНА

Звягинцева Д.А.¹ (студент), Иванов А.Б.¹ (аспирант)

Научный руководитель – канд. техн. наук Ульянов В.И.¹

¹Университет ИТМО

e-mail: daryazvyaginseva@yandex.ru

Работа выполнена в рамках темы НИР №621290 «Алгоритмы извлечения признаков для сравнительного анализа групп метагеномных образцов полногеномного секвенирования с использованием графов де Брейна и применением мультимодальных омиксных данных».

В работе предлагается новый метод выделения признаков из метагеномных образцов. Метод базируется на раскраске графа де Брейна и дальнейшем разделении его на компоненты с использованием полученной раскраски. Извлеченные признаки предполагается использовать в дальнейших исследованиях, тем самым улучшить качество диагностирования заболеваний.

Ключевые слова: метагеномика, извлечение признаков, k-меры, граф де Брейна, раскраска графа, диагностирование заболеваний.

Множество людей по всему миру страдает от воспалительных заболеваний кишечника (ВЗК). Часто неинвазивных методов бывает недостаточно для точной постановки диагноза. Так существуют болезни (болезнь Крона и язвенный колит), которые трудно отличить друг от друга. В связи с этим необходимы другие методы, которые могут облегчить диагностирование заболеваний. Ряд исследований направлен на изучение связи микробиоты и ВЗК [1, 2]. В рамках НИР также ведутся разработки методов диагностики, основанных на метагеномных данных.

Основной инструмент в проекте – программа MetaFast [3], позволяющая выделять признаки из метагеномных образцов. Исходная версия алгоритма заключается в следующем:

- выделить k-меры из метагеномных образцов;
- построить по ним граф де Брейна;
- разбить граф на компоненты, используя обход в ширину;
- посчитать покрытие образцов компонентами, тем самым получив вектора признаков.

Данный проект направлен на усовершенствование имеющегося алгоритма выделения признаков, путем добавления раскраски графа. В данной работе реализован соответствующий алгоритм и внедрен в исходный код программы MetaFast, реализованный на языке java. Для тестирования и генерации тестовых образцов использовались скрипты, реализованные на языке python и готовые библиотеки (BIO), а также Insilicoseq [4].

Методы раскраски графов

На вход программе Metafast подаются метагеномные образцы. Далее из них выделяются k-меры, по которым строится граф де Брейна. В классической реализации MetaFast дальше происходит разбиение на компоненты, путем обхода графа в ширину с ограничениями на максимальный и минимальный размер компоненты. Далее каждому образцу тестовой выборки сопоставляется вектор признаков, а именно, считается покрытие компонент графа искомым образцом. Проблема заключается в том, что размерность признаков, равная числу компонент,

получается очень большая. В связи с чем есть необходимость делить граф на компоненты более сложным способом.

Так, был предложен способ раскрасить граф. А именно, после того как построен граф, присвоим каждой из вершин (k-меров) вектор размерности C , где C – количество классов, к которым принадлежат образцы. В данном случае, таких классов будет три. Далее запишем в вектор вероятность встретить данный k-мер в образце заданного класса. Посчитаем значения по имеющимся данным. Например, если некоторый k-мер встречался у восьми здоровых пациентов, двух с язвенным колитом, и ни у кого из болеющих болезнью Крона, получим вектор $[0.8, 0.2, 0]$.

Далее попробуем определить наиболее вероятный класс. Пример представлен на рисунке 1. Однако, нельзя полагаться на большинство. Есть случаи, когда вероятности распределяются равномерно. Это, как правило, консервативные участки генома. На рисунке 1 мы обозначили такие вершины белым цветом. И данные вершины (как и вершины с достаточно близким распределением вероятностей между классами), представляют некую трудность в исследовании.

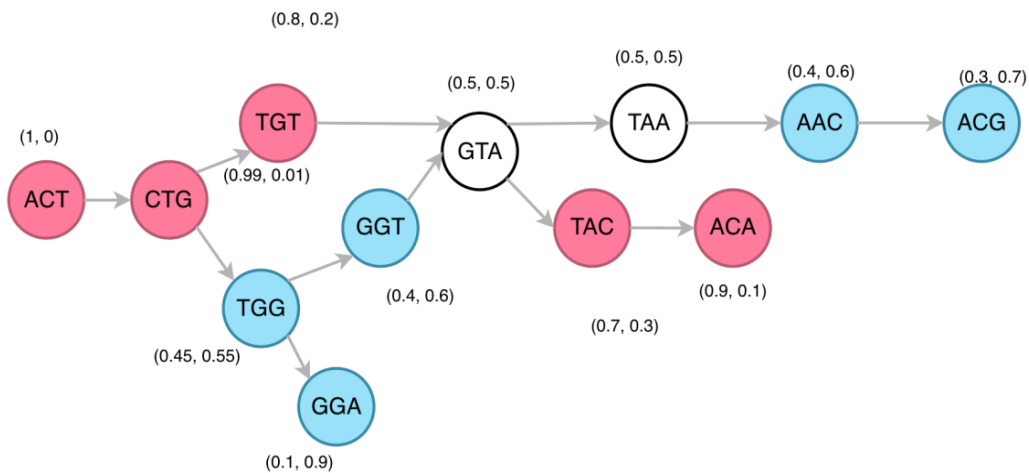


Рис. 1. Пример наивной раскраски определения класса

Далее, как и в классическом алгоритме MetaFast, мы используем обход графа в ширину, чтобы разбить его на компоненты. Однако, переход мы разрешаем только внутри одного цвета для однозначно определяемых вершин. Для неоднозначных вершин возможны варианты.

Один из способов, это считать, что если k-мер встречается равномерно во всех классах, то он принадлежит сразу ко всем классам.

Так, в примере, представленном на рисунке 2 мы получим две компоненты:

- АСТ СТГ ТГТ ГТА ТАА ТАС АСА;
- ГГА ТГГ ГГТ ГТА ТАА ААС АСГ.

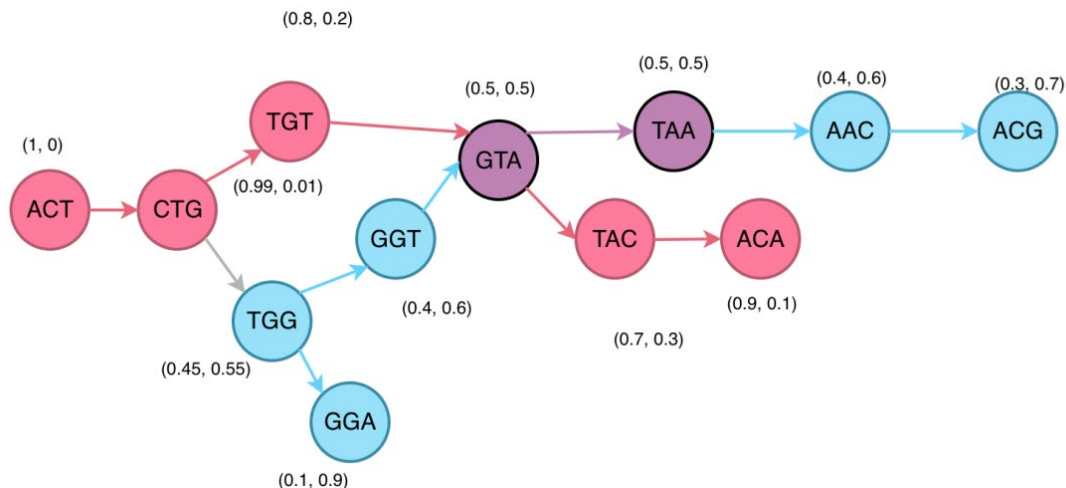


Рис. 2. Пример определения неоднозначных k-меров, как принадлежащих всем классам

Однако, в данном случае будут возникать сложности при обходе графа, т.к. мы не можем точно гарантировать число раз, которое может быть посещена вершина.

Другой способ заключается в том, чтобы определить такие вершины, как отдельный класс. В примере на рисунке 3 мы получим пять компонент:

- АСТ CTG TGT;
- TAC ACA;
- GTA TAA;
- GGA TGG GGT;
- AAC ACG.

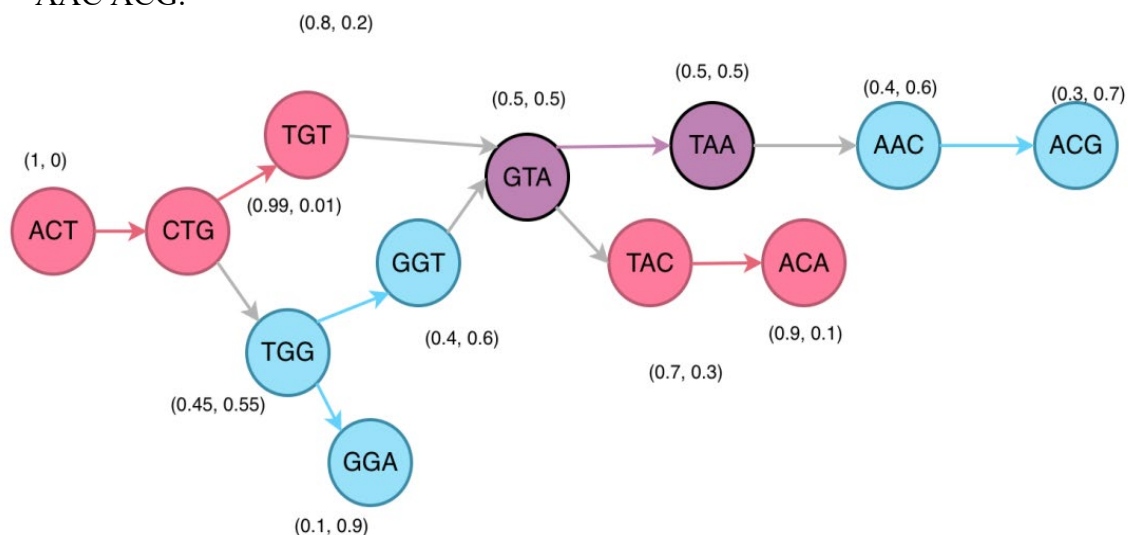


Рис. 3. Пример определения неоднозначных k-меров, как принадлежащих отдельному классу

Итоговый алгоритм

На данном этапе реализован последний алгоритм (неоднозначных k-меров, как принадлежащих отдельному классу) как наиболее простой в реализации. Текущий пайплайн представлен на рисунке 4.



Рис. 4. Итоговый пайплайн алгоритма

- 1) kmer-counter — извлечение k-меров из образцов;
- 2) kmers-color — сопоставление k-меру вектора цветов;
- 3) component-cutter-color — разделение графа на компоненты с учетом цвета;
- 4) features-calculator — получение векторов признаков по покрытию компонентами.

Код для kmers-color и component-cutter-color был разработан и реализован в рамках данного проекта и проверен на сгенерированных метагеномных образцах. На тестовых данных полученные векторы соответствовали природе образцов, однако, более точные выводы о качестве алгоритма будут сделаны после проверки на реальных данных.

В данной работе был предложен алгоритм выделения признаков, основанный на раскраске графа де Брейна. Он был реализован и внедрен в программу MetaFast.

В дальнейшем планируется:

- оптимизировать использование памяти в программе;
- провести эксперименты на реальных данных, а не моделированных;
- провести эксперименты с выделенными признаками, применяя методы машинного обучения;
- провести эксперименты по настройке алгоритма, подбору параметров;
- поменять способ раскраски неоднозначных компонент и провести еще одну серию экспериментов.

Литература

1. Lloyd-Price J. et al. Multi-omics of the gut microbial ecosystem in inflammatory bowel diseases // *Nature*. – 2019. – Т.569. – №7758. – С. 655–662.
2. Schirmer M. et al. Dynamics of metatranscription in the inflammatory bowel disease gut microbiome // *Nature microbiology*. – 2018. – Т.3. – №3. – С. 337–346.
3. Ulyantsev V.I. et al. MetaFast: fast reference-free graph-based comparison of shotgun metagenomic data // *Bioinformatics*. – 2016. – Т.32. – №18. – С. 2760–2767.
4. Gourelé H. et al. Simulating Illumina metagenomic data with InSilicoSeq // *Bioinformatics*. – 2019. – Т.35. – №3. – С. 521–522.

УДК 004.891.2

ИЗВЛЕЧЕНИЕ МЕТАГЕНОМНЫХ ПРИЗНАКОВ И МАШИННОЕ ОБУЧЕНИЕ ДЛЯ ДИАГНОСТИРОВАНИЯ ВОСПАЛИТЕЛЬНЫХ ЗАБОЛЕВАНИЙ КИШЕЧНИКА

Иванов А.Б.¹ (аспирант)

Научный руководитель – канд. техн. наук Ульянцев В.И.¹

¹Университет ИТМО

e-mail: ivart07@mail.ru

Работа выполнена в рамках темы НИР №621290 «Алгоритмы извлечения признаков для сравнительного анализа групп метагеномных образцов полногеномного секвенирования с использованием графов де Брейна и применением мультимодальных омиксных данных».

В работе предлагаются новые методы для извлечения признаков из групп метагеномных образцов. Они основаны на определении уникальных k-меров и их последующей сборке в контиги. Разработанные методы были применены для диагностирования воспалительных заболеваний кишечника на основе данных секвенирования микробиоты. На основе извлеченных признаков были обучены различные модели машинного обучения. Точность полученных классификаторов превышает значения, полученные при использовании классических подходов таксономической аннотации и анализа k-мерных спектров.

Ключевые слова: метагеномика, полногеномное секвенирование, извлечение признаков, k-меры, граф де Брейна, машинное обучение, диагностирование заболеваний.

Анализ микробных сообществ, населяющих различные ниши окружающей среды и организмы млекопитающих, является важной задачей. Например, микробиота кишечника человека играет ключевую роль в усвоении питательных веществ и формировании иммунного ответа на внешние возбудители. Кроме того, микробные сообщества могут быть исследованы на наличие генов устойчивости к антибиотикам. Развитие данной области стало возможно благодаря появлению быстрых и эффективных методов секвенирования последовательностей ДНК, а также благодаря созданию компьютерных программ для параллельной обработки огромных массивов накапливаемых данных. Методы сравнительного анализа метагеномов можно условно разделить на две категории по принципу работы.

Первая категория производит анализ с использованием референсной информации, хранящейся в базах данных (англ. reference-based methods). Как правило, в качестве такой информации выступают данные о нуклеотидных последовательностях отдельных бактерий (таксономическая информация) или данные о последовательностях отдельных генов (функциональная информация). Данная группа методов позволяет найти набор известных сущностей в образце, однако, в сложных и зашумленных метагеномных сообществах до 70 % объема информации из метагеномного образца может не поддаваться таксономической аннотации. Это может приводить к потере ключевой информации об образце и ошибкам в дальнейшем анализе.

Вторая категория методов не использует референсную информацию при анализе данных (англ. reference-free methods). Методы опираются на информацию, которую можно извлечь непосредственно из анализируемых образцов. Наиболее популярным является использование информации о k-мерах – коротких нуклеотидных последовательностях. Можно производить построение k-мерных спектров и сравнивать образцы на их основе. Также, на основе k-меров можно конструировать графы де Брейна, из которых затем можно извлекать линейные участки

(контиги) для дальнейшего анализа. Данный подход позволяет анализировать весь объем информации из образца, однако, биологическая интерпретация признаков, критически важная для биомедицинского применения, ограничена.

Данное исследование направлено на разработку методов, которые совмещали бы преимущества двух описанных подходов. Ставится цель разработать методы для извлечения признаков с использованием наиболее полной информации о метагеномных образцах, и при этом сохранить возможность для дальнейшей интерпретации получаемых признаков.

В качестве реальных данных для анализа в данной работе используются метагеномы кишечника людей с воспалительными заболеваниями кишечника (ВЗК). Заболевание поражает в основном людей в развивающихся странах и характеризуется хроническими воспалительными процессами в желудочно-кишечном тракте. Выделяют два основных типа заболеваний: язвенный колит (ЯК) и болезнь Крона (БК). Язвенный колит характеризуется непрерывным воспалением в терминальных отделах толстого кишечника и поражает в основном слизистую оболочку стенок кишечника. Для болезни Крона характерны очаговые воспаления на всем протяжении желудочно-кишечного тракта, которые могут распространяться по всей толщине стенок. Раннее диагностирование ВЗК является сложной задачей из-за часто смазанной клинической картины и недостаточной информативности биохимических тестов. Золотым стандартом диагностирования служит биопсия при колоноскопии, однако, стремление к уменьшению числа инвазивных диагностических процедур приводит к необходимости поиска новых методов диагностики. В частности, анализ данных метагеномного секвенирования микробиоты кишечника рассматривается как один из наиболее перспективных методов.

Методы извлечения метагеномных признаков

В основе предлагаемого метода лежит программное средство для анализа метагеномных данных MetaFast [1], которое позволяет работать с короткими прочтениями на уровне k-меров и графа де Брейна. На вход программа принимает метагеномные образцы полногеномного секвенирования от секвенаторов семейства Illumina, разбитые на категории. Критерием для разделения выступает вспомогательная информация об образцах, например наличие определенного заболевания у пациента или прием специфического препарата. На первом этапе производится извлечение k-меров, уникальных для определенной группы. Уникальность определяется по принадлежности k-мера к хотя бы заданному числу образцов из одной группы и полному отсутствию в образцах других групп. На втором этапе из каждого метагеномного образца отбираются все короткие прочтения, в которых встречаются уникальные k-меры. Теперь, имея специфичные прочтения, необходимо объединить их для получения признаков.

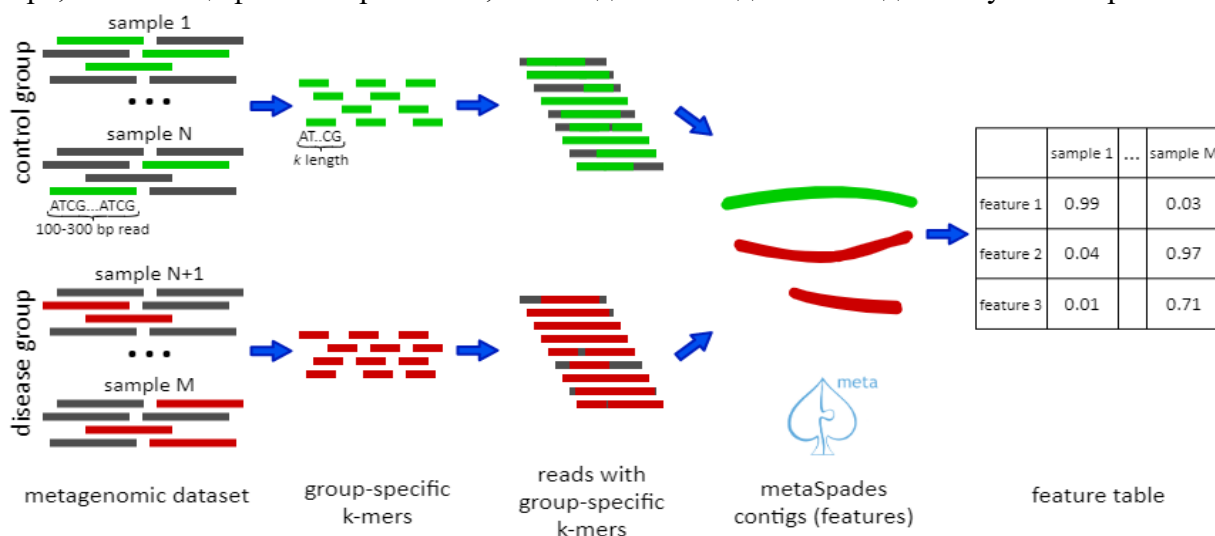


Рис. 1. Схема извлечения признаков из групп метагеномных образцов с использованием уникальных k-меров и сборщика metaSpades

Для объединения прочтений в продолжительные геномные участки (контиги) используется программа для метагеномной сборки metaSpades [2]. Ей на вход подаются наборы

отобранных прочтений, из которых затем строится граф де Брейна для различных длин k -меров, и извлекаются контиги. Каждый полученный контиг используется в дальнейшем анализе как отдельный признак. Для сопоставления каждому признаку численного значения производится вычисление глубины или ширины покрытия контига k -мерами из образца. Схема работы предлагаемого алгоритма представлена на рисунке 1.

В результате работы описанного алгоритма на выходе получается таблица численных значений признаков для каждого из образцов, которая может быть использована для дальнейшего анализа в методах машинного обучения. Помимо таблицы признаков, для анализа остаются доступными сами геномные последовательности, выступающие в роли признаков. В дальнейшем наиболее важные признаки могут быть дополнительно проанализированы на наличие в них определенных генов.

Методы машинного обучения для диагностирования ВЗК

Разработанные методы были применены для анализа данных секвенирования микробиоты кишечника здоровых людей и людей с ВЗК. Были использованы опубликованные данные из открытых источников, основная информация о которых приведена в таблице. Всего было проанализировано четыре набора данных, в которых суммарно было представлено 520 метагеномов. Отмечается некоторое преобладание образцов от пациентов с болезнью Крона, однако, в целом образцы равномерно распределены между категориями. Люди из разных исследований отличаются по особенностям питания и диеты, а также образа жизни и окружающей среды, что позволяет не только проверить предсказательную способность извлекаемых признаков, но и степень обобщения признаков из одной когорты на другие.

Таблица. Наборы метагеномных данных для вычислительных экспериментов

Набор данных	Болезнь крона	Язвенный колит	Контроль	Всего
Schirmer [3]	23	17	14	54
He [4]	63	0	53	116
Franzosa [5]	88	76	56	220
HMP2 [6]	65	38	27	130
Всего	239	131	150	520

В рамках данного исследования в качестве обучающего набора данных был использован набор образцов HMP2, полученных в рамках проекта «Микробиом человека» (англ. Integrative Human Microbiome Project). 130 образцов были разделены на три категории и подвергнуты анализу с помощью разработанного алгоритма для извлечения метагеномных признаков. В результате было получено 14537, 4169 и 4974 признаков для категорий БК, ЯК и контроль, соответственно. Полученные признаки были использованы для обучения двух моделей машинного обучения: случайный лес деревьев решений (англ. Random Forest) и нейронные сети (англ. Neural Network). Был произведен подбор оптимальных параметров для каждого из алгоритмов. Полученные признаки и предобученные модели машинного обучения были использованы для тестирования и трехклассовой классификации трех других наборов данных (Schirmer, He, Franzosa).

Для сравнения с предложенным методом, были применены другие классические методы для анализа метагеномов. Один из них – проведение таксономической аннотации для определения видов микроорганизмов, населяющих образец; другой – построение k -мерного спектра для коротких значений k (в данном случае использовались 9-меры). Также была произведена попытка эвристической классификации с использованием только уникальных k -меров и компонентов графа де Брейна из оригинального алгоритма MetaFast.

Результаты точности работы классификаторов приведены на рис. 2. Таксономическая аннотация является одним из наиболее универсальных и широко используемых методов для анализа метагеномов. Тем не менее, большинство методов, основанных на анализе k -меров, показывают более качественные результаты классификации. Для двух наборов данных Schirmer и Franzosa наилучшие результаты классификации были достигнуты при использовании предложенного в данной работе метода.

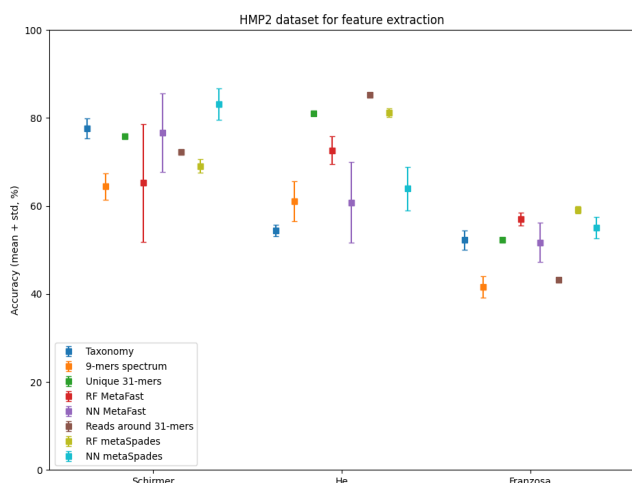


Рис. 2. График точности предсказания классов метагеномных образцов при использовании различных признаков и методов машинного обучения

В данной работе был разработан и реализован метода извлечения признаков из метагеномных данных. Предложенный метод основан на определении уникальных для групп k -меров и их последующем объединении в контиги. Полученные признаки были применены в моделях машинного обучения для классификации метагеномных образцов секвенирования микробиоты кишечника пациентов с воспалительными заболеваниями кишечника. Результаты классификации превосходят по точности результаты работы классических методов и могут стать основой для разработки диагностических и рекомендательных систем.

Литература

1. Ulyantsev V.I. et al. MetaFast: fast reference-free graph-based comparison of shotgun metagenomic data // *Bioinformatics*. – 2016. – Т.32. – №18. – С. 2760–2767.
2. Nurk S. et al. metaSPAdes: a new versatile metagenomic assembler // *Genome research*. – 2017. – Т.27. – №5. – С. 824–834.
3. Schirmer M. et al. Dynamics of metatranscription in the inflammatory bowel disease gut microbiome // *Nature microbiology*. – 2018. – Т.3. – №3. – С. 337–346.
4. He Q. et al. Two distinct metacommunities characterize the gut microbiota in Crohn's disease patients // *Gigascience*. – 2017. – Т.6. – №7. – С. gix050.
5. Franzosa E.A. et al. Gut microbiome structure and metabolic activity in inflammatory bowel disease // *Nature microbiology*. – 2019. – Т.4. – №2. – С. 293–305.
6. Lloyd-Price J. et al. Multi-omics of the gut microbial ecosystem in inflammatory bowel diseases // *Nature*. – 2019. – Т.569. – №7758. – С. 655–662.

УДК 004

ПРОЕКТИРОВАНИЕ СИСТЕМЫ ОЦЕНКИ ДОСТУПНОСТИ ОБЪЕКТОВ ГОРОДСКОЙ СРЕДЫ ПО ДАННЫМ ИЗ СОЦИАЛЬНЫХ СЕТЕЙ

Казак В.А.¹ (студент)

Научный руководитель – канд. полит. наук Чугунов А.В.¹

¹Университет ИТМО

e-mail: kazakvauniversity@gmail.com

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».

В статье рассмотрена проблема существующих методик оценки доступности объектов городского благоустройства. Исследованы существующие на данный момент методики оценки городских объектов и проведен сравнительный анализ инструментов по извлечению данных из социальных сетей. Обосновано создание оригинальной методики в качестве наиболее удобной и эффективной для оценки городской доступности для маломобильных групп населения.

Ключевые слова: парсинг, API-ключи, Извлечение информации, Социальные сети, Веб-скрейпинг, Городская среда, Доступная среда.

Сегодня социальные сети играют в жизни современного человека огромную роль. Причем люди таким образом могут не только общаться, но еще и делиться советами, профессиональными навыками, продвигать свои услуги, покупать товары и прочее.

Благодаря сети можно найти самую разную информацию, которая в данный момент интересует. И если воспользоваться всей этой информацией правильно, то это открывает широчайшие горизонты почти во всех сферах жизни общества. Одной из таких областей является сфера, связанная с развитием города. Например, для управления городскими проектами можно использовать данные из социальных сетей или из специализированных сервисов обладающими свойствами социальных сетей. Одним из таких источников информации могут послужить комментарии с рейтингами из сервиса Google Maps [1]. Для того чтобы получить эти данные для последующего анализа нужно воспользоваться специализированными инструментами.

На данный момент времени существуют различные способы оценки доступности среды, к примеру опросы, проведение наблюдений, изучение проектной документации, интернет-опросы. Однако, все эти способы трудозатратны, то есть возникает необходимость задействовать значительные кадровые ресурсы на протяжении долгого времени и в зависимости от исследуемого участка их объем может быть разным. Часто, в режиме экономии ресурсов, данные методы не используются городской властью.

Поскольку данная проблема слабо освещена не только в зарубежном, но и в отечественном зарубежном пространстве, то целесообразным будет предложить новую методику оценки доступности объектов городской среды.

Основной целью данной работы является предложение методики оценки доступности объектов городского благоустройства для маломобильных групп населения.

Объектом работы являются городские объекты определенного района города. Предметом выступают данные, находящиеся в социальных сетях (различные комментарии и отзывы по городским объектам) [2–5].

Разработка новой методики даст возможность улучшить общественное пространство и сделать его современным и доступным для маломобильных групп населения.

Предлагаемая методика подразумевает:

- определение аудитории пользователей городских объектов;
- составление словаря, включающий в себя слова, которые могут характеризовать доступность объекта;
- определение способа выгрузки данных из интернет-ресурсов;
- анализ и обработка полученных данных.

Сам метод можно разбить на несколько шагов. Шаги практической реализации метода можно изобразить схемой, которая представлена на рисунке 1.

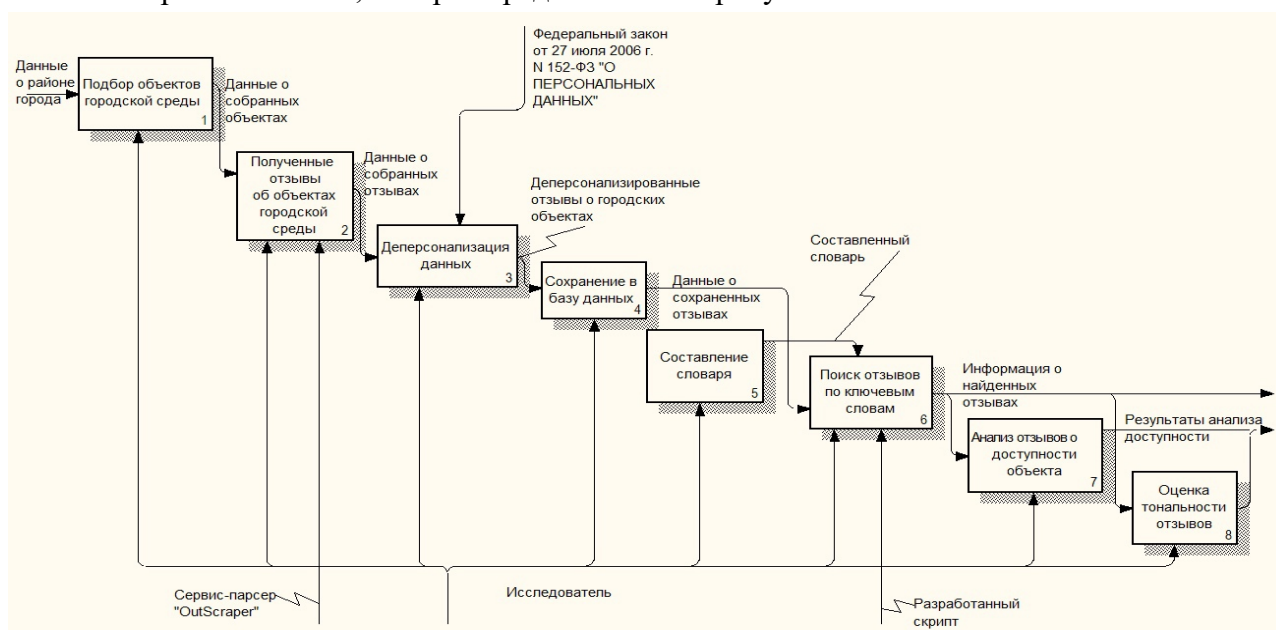


Рис. 1. Оценка доступности городской среды

Исходными данными для данного исследования являются различные текстовые сообщения жителей, содержащие информацию о том или ином городском объекте.

В качестве источника данных была рассмотрена геоинформационный сервис – Гугл карты. Гугл карты дают возможность оставлять комментарии о локации. Каждый, кто посетил объект на карте, может оставить отзыв, как положительный, так и отрицательный, при этом, кроме численной оценки, пользователь может оставить текстовый комментарий. Эта информация становится очень полезной при изучении плюсов и минусов локации. Как правило, такие отзывы оставляют на бизнес-объекты – кафе, рестораны, клубы, гостиницы, и бизнес очень успешно использует их, для улучшения качества своих услуг. Различные муниципальные учреждения – больницы, школы, полиция, тоже работают с отзывами. Что говорит о высокой информативности данных отзывов. Значит, данную информацию можно использовать и для оценки рекреационных объектов благоустройства (парки, скверы и сады).

Для проведения исследования был отобран один район в Санкт-Петербурге - Петроградский район, далее, отобраны типы объектов:

- парк;
- сад;
- сквер.

Для данных типов было отобрано 17 локаций, далее, в Гугл картах были сформированы ссылки на каждый объект, на карточку с отзывами.

Изучив методы выгрузки данных из Интернет-ресурсов, были проанализированы несколько альтернативных вариантов и было принято решение искать готовый парсер (сервис). Поэтому, был найден готовый парсер – «OutScrapet», данный инструмент дает возможность выгружать отзывы о любом объекте на Гугл картах в машиночитаемом виде.

При помощи данного сервиса были выгружены все отзывы об интересующих нас объектах, итоговая база составила 4348 отзывов, для 17 объектов. Все данные были деперсонализированы.

В данной работе мы сосредоточились на изучении доступности таких объектов для незащищенных социальных групп – люди с ограниченными возможностями, семьях с колясками и пенсионеров. Необходимо было понять, пишут ли представители данных социальных групп отзывы про парки, в которых есть информация о доступности.

На следующем этапе, в результате анализа специализированных текстов был составлен словарь, включающий 60 слов, которые могут характеризовать доступность объекта. Далее, с помощью скрипта был произведен поиск по собранной базе данных, с использованием этих слов и различных словоформ. Были отобраны все отзывы, которые содержат термины из созданного словаря. В ходе поиска было отобрано 450 отзывов, которые характеризуют доступность. Это очень ценная информация, которая может помочь улучшить эти объекты.

Далее была проведена ручная проверка, исключая отзывы, которые были малоинформативные. Итоговая база содержала 125 отзывов о доступности объектов. Дополнительно была проанализирована тональность сообщений. Около 20% таких отзывов имели негативную тональность и 57% позитивную тональность. А также был проанализирован пользовательский рейтинг в отрыве от текстового содержания комментария. Данные диаграммы говорят о том, что район в целом доступен для маломобильных групп населения, однако, есть некоторое недовольство со стороны посетителей рекреационных зон Петроградского района. Также глядя на полученные диаграммы, мы можем наблюдать то, что иногда люди могут писать негативный отзыв, но при этом ставить хорошую оценку месту посещения (4-5), что делает способ оценивать место только по рейтингу нецелесообразным

Также в рамках проделанной работы была построена концепция картографического сервиса, представленная на рисунке 2. Кроме этого был представлен прототип итогового картографического сервиса (рис. 3).

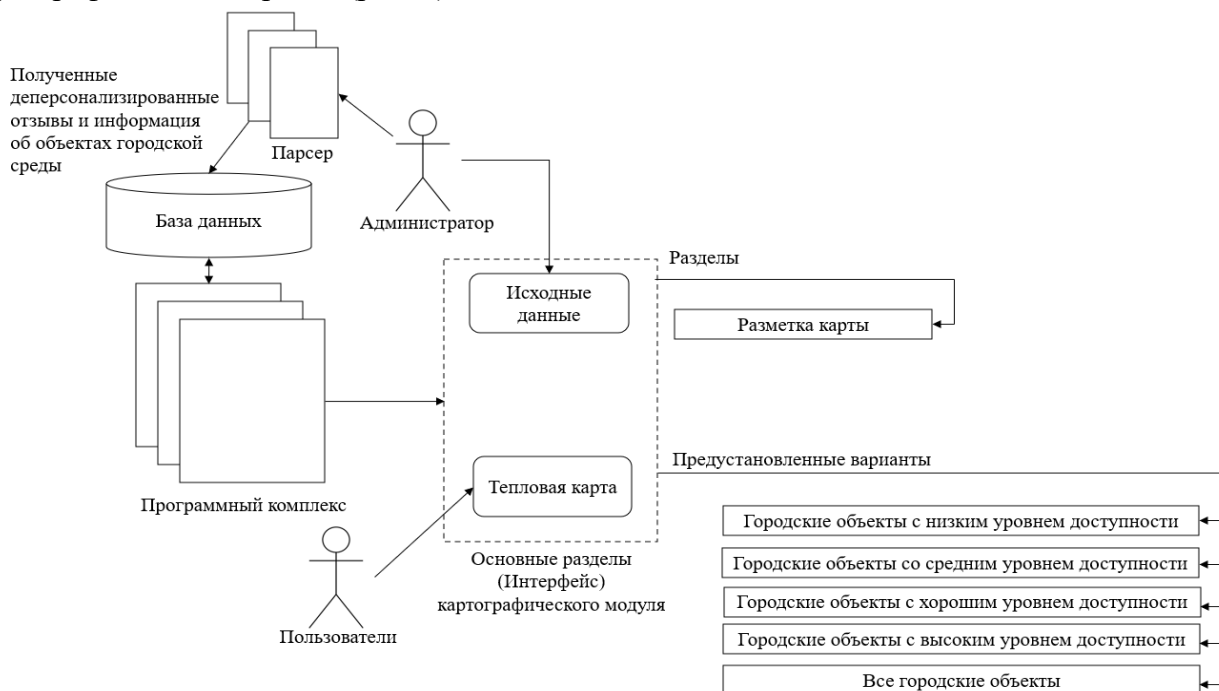


Рис. 2. Концепция картографического сервиса

В рамках выполнения работы были рассмотрены и проанализированы, научные статьи, специализированные сайты, различные труды конференций, а также были просмотрены мнения

на специализированных форумах и в каждом из этих информационных источников были затронуты проблемы, связанные с оценкой доступности городских объектов для маломобильных групп населения.

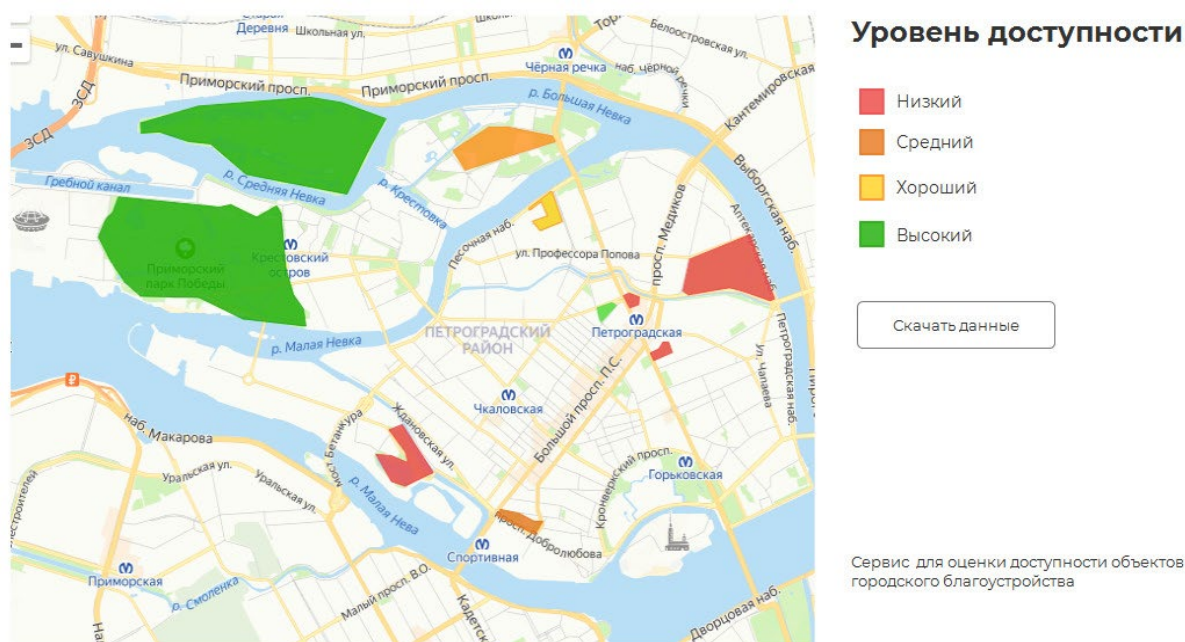


Рис. 3. Прототип картографического сервиса

В ходе проделанной работы была представлена и описана методика применения оценки доступности объектов городского благоустройства. В качестве наглядного примера были взяты объекты городской среды (парки, скверы, сады), расположенные на территории Петроградского района.

По итогам применения методики был произведен анализ полученных отзывов и, исходя из них, было выявлено, что в целом состояние объектов городского благоустройства находится в не совсем надлежащем состоянии, ведь 23% негативных отзывов показатель того, что в районе присутствуют труднодоступные территории для маломобильных групп населения, на которые администрация данного района должна обратить внимание и оперативно исправить ситуацию. Также были построены концепция и прототип картографического сервиса.

Литература

1. Ненько А.Е., Семенова А.М., Галактионова А.А. Оценка качества общественных пространств по данным отзывов в Google Maps // Научный сервис в сети Интернет: труды XXII Всероссийской научной конференции (21-25 сентября 2020 г., онлайн). — М.: ИПМ им. М.В.Келдыша. — 2020. — С. 473–485.
2. What are Location-Based Social Networks. [Электронный ресурс]. — Режим доступа: <https://www.quickanddirtytips.com/business-career/communication/what-are-location-based-social-networks> (дата обращения: 16.12.2020).
3. A vision of smarter cities How cities can lead the way into a prosperous and sustainable future, IBM Institute for Business Value. [Электронный ресурс]. — Режим доступа: <https://www.ibm.com/downloads/cas/2JYLM4ZA> (дата обращения: 27.12.2020).
4. Аналитика социальных сетей как инструмент для продвижения территорий. [Электронный ресурс]. — Режим доступа: <https://vc.ru/marketing/81241-analitika-socialnyh-setey-kak-instrument-dlya-prodvizheniya-territoriy> (дата обращения: 28.12.2020).
5. Порозова Д.Ю., Симбирцева Н.А. Особенности социальных сетей в городском пространстве // Инновации в науке. — 2011. — №4. — С. 112–116.

УДК 004.58

ИССЛЕДОВАНИЕ ПРИМЕНЕНИЯ ДИАЛОГОВЫХ АГЕНТОВ В СФЕРЕ ЗДРАВООХРАНЕНИЯ

Калинин П.С.¹ (студент), Лисицинский С.Г.² (студент), Орлов Г.М.²

¹*Санкт-Петербургский государственный университет*

²*Университет ИТМО*

e-mail: pashkalini2000@ya.ru

Статья посвящена теме применения диалоговых агентов (чат-ботов) в сфере здравоохранения для электронного взаимодействия медицинской организации и пациента. В статье рассказывается об электронном взаимодействии и цифровизации в медицинской сфере в общем, а также о применении диалоговых агентов в медицинской сфере. Приведены преимущества данного способа взаимодействия пациента и клиники. В статье приводится анализ рынка цифровых продуктов в данной сфере. Также приводятся результаты опроса пациентов ФГБУ СЗОНКЦ им. Л.Г. Соколова ФМБА России и описание наиболее востребованных сервисов среди пациентов.

Ключевые слова: медицинский чат-бот, диалоговый агент, электронное взаимодействие, цифровое здравоохранение, сервисы для пациента.

В современных реалиях цифровизации сфера здравоохранения не может оставаться в стороне. Это обуславливается следующими факторами:

- научно-технический прогресс;
- информатизация и мобильность (возможность выхода в Интернет из любой точки мира);
- пациентоцентричность (модель организации здравоохранения, построенная вокруг потребностей пациента, предполагающая комплексное управление здоровьем);
- датацентричность (массивы данных о здоровье населения, на основе которых принимаются эффективные решения для лечения отдельных пациентов, а также решения в области организации здравоохранения) [1].

Общая цель процесса цифровизации здравоохранения – упрощение взаимодействия пациента и медицинской организации, а также решение актуальных проблем, связанных с обеспечением максимального количества жителей страны медицинскими услугами, обеспечением доступности медицинской помощи в условиях транспортной разобщенности и масштабности территории. Приближение к этой цели видится в развитии и внедрении цифровых технологий для взаимодействия пациента и медицинского учреждения [2].

Цифровые технологии в здравоохранении должны решать определённые задачи и предоставлять сервисы для:

- взаимодействия простых граждан с системой здравоохранения;
- профилактики заболеваемости и пропаганды здорового образа жизни;
- дистанционного оказания медицинской помощи или «телемедицины»;
- заочного обучения специалистов здравоохранения;
- поддержки и управления в области медицинских исследований [3].

Также важно отметить, что в случае возникновения чрезвычайной ситуации, как, например, пандемии COVID-19, важным становится «использование цифровых технологий для поддержки, профилактики и контроля эпидемии» [4].

Весной 2020 года Министерством здравоохранения Российской Федерации и Федеральной службой по надзору в сфере защиты прав потребителей и благополучия человека (Роспотребнадзор) был создан интернет-ресурс «Стопкоронавирус.рф», а также одноимённые

чат-боты и каналы в различных популярных мессенджерах (Telegram, WhatsApp, Viber, ICQ), для информирования и взаимодействия с населением по вопросам коронавируса (COVID-19).

Чат-боты активно внедряются в операционные процессы различных компаний и учреждений последние несколько лет. Чат-бот – это виртуальный ассистент, программа, предназначенная для упрощения взаимодействия человека с компьютерами. В настоящее время они могут быть интегрированы на популярные платформы, такие как Facebook Messenger, Telegram, WhatsApp, из которых в настоящее время представляется целесообразным использовать только Telegram, а также веб-сайты. Практика показывает, что пользователи хорошо реагируют на коммуникацию в мессенджерах и охотнее участвуют в ней, а чат-бот помогает автоматизировать общение, быстро и точно выполняя рутинные задачи.

Рассматривая вопрос применения чат-ботов в сфере здравоохранения, важно отметить, что современному человеку, решающему вопросы в формате «здесь и сейчас», крайне необходимы сервисы, помогающие максимально упростить и ускорить посещение медицинской организации. Кроме того, экономия времени пациентов при получении медицинской помощи позволяет проводить более полные обследования и значительно повышает качество жизни граждан.

Рост популярности чат-ботов в медицинской сфере связан с многочисленными преимуществами данного способа реализации сервисов для пациента:

- получение мгновенного ответа на запрос;
- онлайн запись к подходящему врачу;
- отсутствие очередей;
- экономия времени и абсолютная конфиденциальность и многое другое.

При анализе рынка различных медицинских сервисов для пациента, которые реализованы с помощью чат-ботов, были отобраны семь вариантов. Их функционал кратко рассмотрен далее.

Чат-бот Doc+. Основная задача бота – уменьшить время, затраченное врачом и пациентом при первичном осмотре. Виртуальный помощник собирает анамнез намного быстрее, чем доктор, всего за несколько минут. Вторая задача бота – это повышение уровня обслуживания. Пациенту задается большое количество вопросов, что снижает риск того, что доктор что-то пропустит. Третья задача – сбор большого массива информации о жалобах и симптомах. При этом все данные хранятся и анализируются в структурированном виде, что повышает точность и качество обследования [5].

Telegram-бот «МСЧ №157» – @med157_bot (Медицинский центр «МСЧ № 157», Санкт-Петербург). Telegram-бот «МСЧ №157» даёт возможность в любое удобное для пациента время производить самостоятельную запись к врачам-специалистам клиники. Функционал бота включает в себя просмотр расписания всех врачей и свободное время для записи, что позволяет пациенту выбрать максимально комфортное время приема. Как только пациент произведёт самостоятельную запись к врачу, информация о ней автоматически появится в базе клиники. Также в разделе «Мои назначения» пользователь может посмотреть дату и время его записи к специалистам, а также номер кабинета и ФИО врача. В разделе «Информация» Telegram-бота «МСЧ №157» отражается справочная информация о клинике: режим работы и телефоны, по которым пациент при необходимости может связаться с администраторами медицинского центра.

Telegram-бот «MedBot» – @InnoMedBot (Медицинский центр, Иннополис). С помощью бота жители города Иннополис (Татарстан) записываются к врачу, на диспансеризацию и платные медицинские услуги. В боте доступна запись к любому врачу в городе. Пользователи могут получать напоминания при приближении записи. После приема бот предлагает оценить качество услуг по пятибалльной шкале.

Telegram-бот «JSCMedicinaBot» – @JSCMedicinaBot (Клиника АО «Медицина», Москва). С помощью бота пациент может произвести запись к врачам клиники АО «Медицина». Реализованные сервисы помогают пациентам оценивать врачей, просматривать историю болезни, результаты исследований и анализов.

Telegram-бот «avanstom» – @avanstom_bot (Стоматологическая клиника Аванстом, Москва). Данный чат-бот предоставляет клиенту клиники следующие сервисы:

- запись на прием, без звонка в клинику;
- изменение даты и времени приема;
- напоминание о запланированном посещении клиники;
- выбор лечащего врача;
- связь с лечащим врачом;
- маршрут к клинике;
- бонусы и скидки за рекомендации;
- специальные предложения;
- текущие акции;
- скидка 7% после оформления.

Telegram-бот «Migrebot» – @Migrebot («Университетская клиника головной боли», Москва). Данный бот помогает пользователю вести дневник головной боли. Дневник головной боли — верный способ получить точный диагноз и правильно лечиться. Каждый день Мигребот задаёт пациенту несколько вопросов. Время для общения пользователь выбирает самостоятельно. Если бот замечает, что пациент принимает слишком много обезболивающих препаратов или приступы стали чаще – предупреждает пользователя. Пациент может скачать дневник в любое время, а потом отправить врачу.

Telegram-бот «Вакцинация COVID-19 Ярославская область» – @vaccine76_bot. Этот инструмент поможет ответить на самые главные вопросы о вакцинации: кто и как может привиться, как проходит вакцинация и какие есть противопоказания. Здесь же пользователь найдёт информацию об адресах пунктов вакцинации и телефоны, по которым можно записаться на процедуру.

В августе 2021 года авторами был запущен опрос пациентов медицинского центра ФГБУ СЗОНКЦ им. Л.Г. Соколова ФМБА России и было получено более 700 ответов. Опрос включал в себя 6 вопросов, которые касались популярности использования мессенджера Telegram среди пациентов клиники, их знакомства с чат-ботами в целом и медицинскими ботами в частности, а также выбора наиболее нужных и важных медицинских сервисов и каналов взаимодействия с ними.

Были получены следующие результаты:

1. Более 60% пациентов пользуются Telegram (рис. 1).
2. Более 60% пациентов знают, что такое боты в Telegram (рис. 2).

3. Есть ли у Вас Telegram?

737 ответов

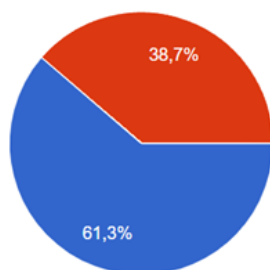


Рис. 1. Результаты опроса А

4. Знаете ли Вы, что такое боты в Telegram?

737 ответов

● Да
● Нет

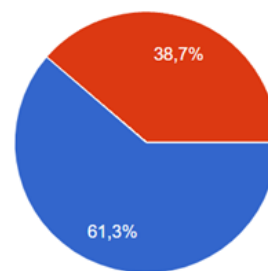


Рис. 2. Результаты опроса В

3. Некоторые пациенты (1% опрошенных) пользовались следующими медицинскими телеграм-ботами: Инвитро, Медэксперт, Меди, Мед-122, Детская поликлиника №69 Красногвардейского района СПб, боты страховых организаций.
4. Были выбраны следующие наиболее важные для пациентов сервисы (отсортировано по количеству голосов, рис. 3):
 - a. Запись на приём на конкретное время (почти 90%).
 - b. Рекомендации и заключение врача (более 65%).
 - c. Напоминание о приёме (более 55%).
 - d. Быстрая оплата услуг без кассы (более 35%).

Также были получены предложения от самих пациентов:

- а. *«Было бы неплохо, чтобы была возможность вызвать врача на дом без звонков и висения на телефоне. А также интересна возможность какой-то быстрой симптоматической консультации, может быть даже через бота, когда нет ничего серьезного, но хочется задать пару вопросов»* - последняя описанная возможность симптоматических ответов на вопросы интересна и популярна.
- б. *«Доступ к электронной карте в любое время».*

2. Выберите наиболее важные для Вас сервисы:

737 ответов

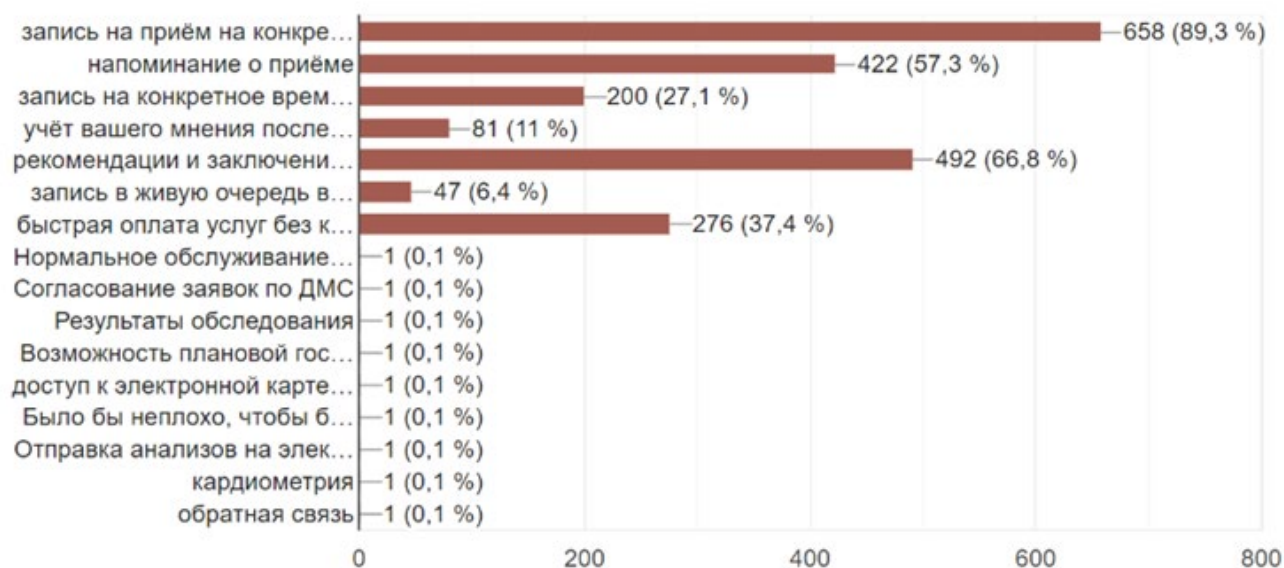


Рис. 3. Результаты опроса С

5. Были получены следующие результаты при ответе на вопрос «Какой канал связи и получения сервисов для Вас самый удобный?» (рис. 4):

- а. Возможность обратиться по телефону – 32,8%.
- б. Мобильное приложение – 22,5%.
- в. Сайт – 19,3%.
- г. Телеграм-бот – 16,7%.
- д. Личное общение с сотрудником регистратуры – 8,7%.

1. Какой канал связи и получения сервисов для Вас самый удобный?

737 ответов

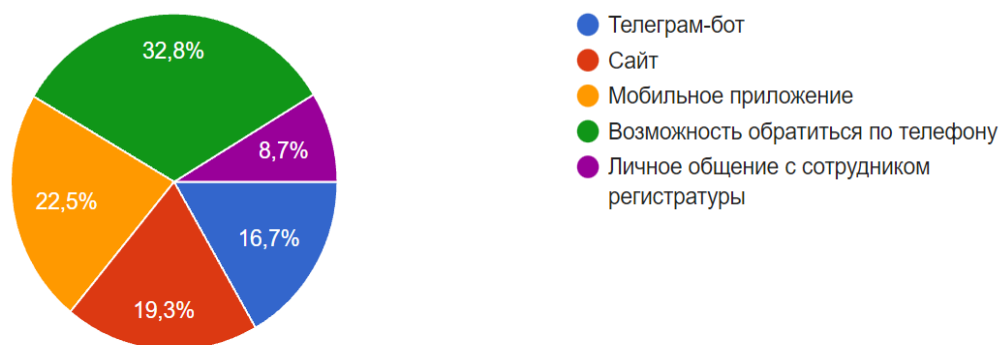


Рис. 4. Результаты опроса D

Результаты опроса

Опрос показывает, что на данный момент традиционные средства связи с медицинскими учреждениями являются приоритетными в сознании граждан, однако, прослеживается осведомленность населения о существовании других каналов связи – более 60% опрошенных пользуются Telegram и знают, что такое чат-боты. Эти цифры в сопоставлении говорят о том, что есть большой потенциал в выборе телеграм-бота как средства для реализации медицинских сервисов. Как упоминалось ранее создание собственного приложения трудоёмко, но людям психологически оно подходит. Telegram в свою очередь – это и есть приложение, причём с точки зрения безопасности более предпочтительное для пользователя, чем неизвестное приложение от медицинского центра. Также использование Telegram экономит место на мобильном устройстве, что тоже не мало важно. Учитывая результат в почти 40% (данные опроса: 22,5% за мобильное приложение и 16,7% за телеграм-бот) среди пациентов, можно с уверенностью заявить, что чат-бот на платформе Telegram будет востребован среди пациентов клиники.

По результатам проведенного исследования были выделены следующие сервисы:

1. Сервис «Запись на приём».
2. Сервис «Заключение врача».
3. Сервис «Напоминание о приёме».
4. Сервис «Оплата услуг».
5. Сервис «QR-код на вход».
6. Сервис «Консультация».

Сервис «Консультация» предполагает внедрение искусственного интеллекта для консультации пациента по симптомам того или иного заболевания. Важно: диагноз может поставить только врач, а данный сервис поможет сэкономить время приема, а также оптимизировать планирование приемных часов – принимать пациентов в зависимости от их состояния; данный сервис достаточно непрост в реализации, но абсолютно точно будет востребован среди пациентов.

Можно с уверенностью сказать, что реализация сервисов для пациента медицинской организации с помощью чат-бота актуальны, современны и будут востребованы у большого количества пациентов.

В рамках проведенного исследования были изучены возможности систем электронного взаимодействия лечебных учреждений и их пациентов, в частности были рассмотрены чат-боты медицинских организаций. Также были изучены потребности пациентов путем проведения опроса, на основе полученных результатов были выделены и описаны наиболее востребованные сервисы для пациентов.

Источники:

1. Карпов О.Э., Субботин С.А., Шишканов Д.В., Замятин М.Н. Цифровое здравоохранение. Необходимость и предпосылки // Врач и информационные технологии. – 2017. – №3. – С. 6–22.
2. Бацина Е.А., Попсуйко А.Н., Артамонова Г.В. Цифровизация здравоохранения РФ: миф или реальность? // Врач и информационные технологии. – 2020. – №3. – С. 73-80.
3. Будущее цифровых систем здравоохранения. Отчет о проведении симпозиума «Будущее цифровых систем здравоохранения в европейском регионе». [Электронный ресурс]. – Режим доступа: <https://apps.who.int/iris/bitstream/handle/10665/330370/9789289059985-rus.pdf>. (дата обращения: 01.03.2022).
4. Как устроены медицинские чат-боты — разбираемся на примере бота DOC+. [Электронный ресурс]. – Режим доступа: <https://habr.com/ru/company/docplus/blog/374493/> (дата обращения: 01.03.2022).

УДК 004

ПРИМЕНЕНИЕ КОГНИТИВНЫХ ТЕХНОЛОГИЙ В ПОСТРОЕНИИ ИНДИВИДУАЛЬНОЙ ТРАЕКТОРИИ ОБУЧЕНИЯ

Касицына М.С.¹ (студент)

Научный руководитель – канд. техн. наук, доцент Осипов Н.А.¹

¹*Университет ИТМО*

e-mail: m.kasitsina@yandex.ru

В работе рассмотрены понятия индивидуальных образовательных траекторий обучения и их роль в настоящее время. Выявлена проблематика реализации формирования индивидуальных траекторий для обучающихся, а также определены варианты их решения. Проведен обзор задач, которые решает машинное обучение, а также рассмотрен принцип работы нейронных сетей. Сделан вывод о возможности решения выявленных проблем путем применения машинного обучения.

Ключевые слова: индивидуальная образовательная траектория, машинное обучение, нейронные сети, обучение, система.

В настоящее время всё популярнее становится организация «индивидуальной образовательной траектории» (ИОТ). Кроме данного понятия существуют еще 3, которые часто используют как синонимы: «индивидуальный образовательный маршрут» (ИОМ), «индивидуальная образовательная программа» (ИОП), «индивидуальный учебный план» (ИУП) [1].

Наиболее заметно использование данного понятия в высших учебных заведениях (вузы). Однако, индивидуализацию обучения часто используют репетиторы или же преподаватели различных секций дополнительного образования. Зачастую они самостоятельно прописывают подобные учебные планы, ориентируясь на запросы своих учеников и основываясь на личном преподавательском опыте. Хороший пример такого подхода – индивидуальный тренер, который составляет план тренировки для конкретного человека, основываясь на его целях, сильных и слабых сторонах. Однако, в основной образовательной среде, ИУП скорее редкость, чем обычное явление. Это связано с тем, что обеспечение индивидуальной траекторией более миллиона человек ежегодно является достаточно трудоемкой задачей, в том числе из-за сложности определения уровня знаний и способностей каждого из них, а также определения сферы интересов.

Несмотря на вышеизложенное, вузы стараются различными способами внедрять индивидуализацию программ. К сожалению, достаточно часто обучающиеся оказываются полностью или частично не удовлетворены выбранной траекторией. В первую очередь, необходимо понимать, что в любой образовательной организации и на любом образовательном курсе существует «образовательная траектория». Иными словами, всегда имеется заранее продуманный состав знаний, необходимых для достижения требуемого или желаемого уровня квалификации в избранной области, и последовательность их получения. Индивидуальная же траектория подразумевает персонализацию образовательного пути, которая основана на принятой человеком стратегии достижения личных долгосрочных целей [2].

Учебные заведения не могут полностью положить построение траектории на плечи обучающихся, так как ограничены образовательными стандартами. Однако, на основе этих стандартов могут предложить вариативность в выборе дисциплин обучения. Таким образом,

мы получаем, что студенты сталкиваются со следующими проблемами при выборе: отсутствие выбора; выбор присутствует, но не всегда может быть осуществлен и выбор из нескольких альтернатив.

Первые две проблемы должны решаться на уровне образовательного учреждения. Третья выделенная проблема имеет несколько решений. Часть из них являются организационными: наличие у выборных дисциплин подробного описания, предполагаемых результатов обучения, полных, формализованных и всесторонне описывающих дисциплину отзывов участников. Другой же вариант – применение когнитивных или информационных технологий.

Понятие искусственный интеллект (ИИ) сейчас повсеместно используется, в связи с этим существует множество различных определений данного термина. Автором же данного понятия считается Джон Маккарти, который в 2004 в своей статье определил искусственный интеллект так: «Это наука и технология создания интеллектуальных систем, в особенности — интеллектуальных компьютерных программ. Искусственный интеллект связан с задачей использования компьютеров для понимания работы человеческого разума, но не ограничивается использованием методов, наблюдаемых в биологии» [3]. Широкое определение искусственного интеллекта не отображает картины того, какой тип ИИ используется в настоящее время. Так, существует понятие узкий искусственный интеллект (УИИ). Когда говорят об существующих технологиях и употребляют термин «искусственный интеллект», то подразумевают именно УИИ. Он позволяет запрограммировать систему на выполнение одной задачи даже в режиме реального времени. Однако, такая система работает с определенным набором данных [4].

Вопрос чем машинное обучение отличается от глубокого обучения возникает достаточно часто. В первую очередь, необходимо понять, что машинное обучение является подразделом ИИ и науки о данных, специализирующийся на использовании данных и алгоритмов для имитации процесса наработки опыта человеком с постепенным повышением точности [5, 6]. Глубокое же обучение является подвидом машинного обучения.

Основное отличие между машинным и глубоким обучением заключается в способе обучения алгоритмов. Классическое машинное обучение, а если быть точнее – его эффективность, зависит от контроля со стороны человека, тогда как большая часть изучения признаков в глубоком обучении автоматизирована, что позволяет использовать большие наборы данных [7].

Нейронные сети, также известные как искусственные нейронные сети (ИНС) или смоделированные нейронные сети (СНС), представляют собой подмножество машинного обучения и лежат в основе алгоритмов глубокого обучения. Их название и структура вдохновлены человеческим мозгом, имитируя то, как биологические нейроны подают сигналы друг другу [8].

ИНС работают по принципу черного ящика – мы, как создатели системы, не можем точно определить, что происходит внутри нейронной сети, когда она работает [9]. Чтобы нейросеть работала, ее необходимо обучить. Обучение производится с использованием трех выборок некоторого набора данных: тестовая, обучающая и контрольная. Тестовая выборка проводит проверку построенной и обученной сети, иными словами сети с найденными параметрами, на отдельном тестовом множестве. Контрольная выборка позволяет избежать переобучения сети – как только ошибка на этой выборке начинает возрастать, процесс обучения прекращается [10]. Этап обучения сети обычно самый длительный и трудоемкий. Это связано и с тем, что обучающая выборка обычно содержит около 70% наблюдений. Использование же обученной искусственной нейронной сети не требует больших вычислительных мощностей.

Искусственная нейронная сеть состоит из трех компонентов: входной слой; скрытые слои и выходной слой.

Обучение нейросетей происходит в два этапа: прямое и обратное распространение ошибки [11]. Метод обратного распространения ошибки позволяет вычислить и объяснить

ошибки, связанные с каждым нейроном, что позволяет скорректировать и адаптировать параметры модели соответствующим образом [8].

На рисунке 1 показано общее строение искусственной нейронной сети. Крайний левый слой называется входным слоем, крайний правый слой выходным слоем. Средний слой называется скрытым, потому что его значения не видны в обучающем наборе. Иными словами, скрытые слои — это вычисляемые значения и чем их больше, тем сеть глубже. А глубокой нейронной сетью называется любая ИНС с двумя или более скрытыми слоями [5].

Весы и смещения — это обучаемые параметры внутри сети. Обучаемая нейронная сеть будет подбирать случайные значения весов и смещений до того, как начнется обучение.

По мере продолжения обучения оба параметра корректируются в сторону желаемых значений и правильного вывода. Эти два параметра различаются по степени их влияния на входные данные. Проще говоря, смещение показывает насколько далеки прогнозы от их предполагаемого значения. Смещения составляют разницу между выходом функции и ее предполагаемым результатом (пороговое значение). Низкое смещение предполагает, что сеть делает больше предположений о форме вывода, тогда как высокое значение смещения делает меньше предположений о форме вывода. С другой стороны, вес можно рассматривать как силу соединения. Вес влияет на степень влияния изменения входных данных на выходные данные. Низкое значение веса не будет иметь никаких изменений на входе, и наоборот, большее значение веса более значительно изменит выход [12].

Для понимания как именно работает нейросеть необходимо знать о линейной регрессии. Линейный регрессионный анализ (линейная регрессия) используется для прогнозирования значения переменной на основе значения другой переменной. Независимой переменной называется переменная, которая используется для предсказания значения другой переменной. А зависимой переменной – переменная, значение которой предсказывается. Эта форма анализа заключается в подборе таких коэффициентов линейного уравнения с одной или несколькими независимыми переменными, чтобы это уравнение наилучшим образом предсказывало значение зависимой переменной [13].

Каждый узел нейронной сети можно представить как модель линейной регрессии. Такая модель состоит из входных данных, весовых коэффициентов, порогового значения или смещения и выходных данных. Эту модель можно представить следующими формулами (1, 2) [11]:

$$\sum_{i=1}^m w_i x_i + bias, \quad (1)$$

где w_i – вес,
 x_i - входные данные,
 $bias$ - смещение.

$$output = f(x) = \begin{cases} 1 & \text{if } \sum_{i=1}^m w_i x_i + bias \geq 0 \\ 0 & \text{if } \sum_{i=1}^m w_i x_i + bias < 0 \end{cases} \quad (2)$$

Первый шаг – определение слоя входных данных, иными словами, необходимо задать значения каждому x_i . Затем назначаются весовые коэффициенты. Затем производится расчет выходных данных: произведения входных данных и соответствующих им весовых коэффициентов суммируются. После чего, полученные значения передаются через функцию активации, которая вычисляет результат (2). Если полученный результат превышает установленное пороговое значение, узел срабатывает (активируется), передавая данные на следующий слой сети. Выходные данные одного узла становятся входными данными для следующего узла. Такой последовательный процесс передачи данных между слоями характерен для нейронных сетей прямого распространения [8].

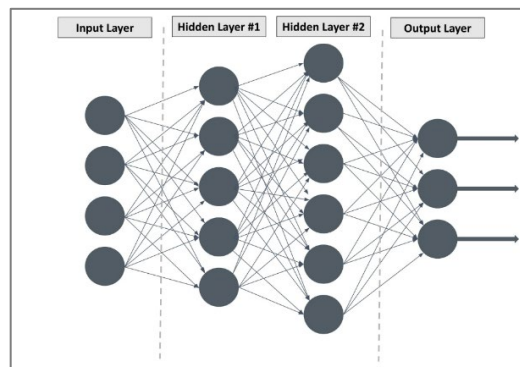


Рис. 1. Простая искусственная нейронная сеть

В ходе обучения модели нам потребуется оценить точность с помощью функции стоимости (или потерь). Это так называемая среднеквадратическая ошибка (MSE). В уравнении ниже (3) используются следующие обозначения:

- i – индекс выборки,
- \hat{y} – прогнозируемый результат,
- y – фактическое значение,
- m – число выборок.

$$MSE = \frac{1}{2m} \sum_{i=1}^m (\hat{y} - y)^2. \quad (3)$$

Согласно методологии CrispDM [14], внедрение машинного обучения включает в себя следующие этапы:

1. Постановка задачи.
2. Изучение данных и их подготовка.
3. Моделирование (построение прогнозирующей модели и подбор ее параметров).
4. Оценка полученной модели.
5. Внедрение.
6. Поддержка.

Под постановкой задачи имеется в виду ее уточнение, то есть необходимо сопоставить проблему с задачей, которую будет решать искусственный интеллект, а также установить целесообразность применения методов машинного обучения [15]. Задачи, решаемые машинным обучением, можно сгруппировать в 5 категорий [16]:

- задача регрессии. Задача состоит в получении (предсказании) вещественного ответа на основании различных признаков. Например, предсказание прибыли магазина в следующем месяце;
- задача классификации. Данная задача состоит в получении категориального ответа на основании различных признаков. Другими словами, ответов в такой задаче конечное количество. Например, является ли письмо спамом или находится на фотографии еж;
- задача кластеризации. Здесь происходит разбиение данных на похожие категории. Например, разбиение космических объектов на похожие (галактики, планеты, звезды и так далее);
- задача уменьшения размерности. Здесь требуется научиться описывать наши данные не N признаками, а меньшим числом. В качестве примера можно привести задачи визуализации и сжатие данных;
- задача выявления аномалий. Задача состоит в том, чтобы научиться различать отличать аномалии от «не-аномалий» на основании признаков. Задача похожа на задачу классификации, однако, здесь есть особенность: примеров аномалий для тренировки модели либо очень мало, либо нет совсем, поэтому такую задачу нельзя решать, как задачу классификации. Например, определение мошеннических транзакций по банковской карте.

В результате изучения предметной области была выявлена проблематика вопроса о построении индивидуальной траектории обучения, а также актуальность создания системы, способствующей грамотному и полноценному составлению таких учебных планов.

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.

Машинное обучение в образовательном процессе применяется для решения задач классификации и кластеризации (рекомендательные системы). При реализации применяются искусственные нейронные сети методов прямого и обратного распространения ошибки, в том числе глубокие нейронные сети.

Основная трудность использования нейронных сетей – 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.

Литература

1. Индивидуальная образовательная траектория обучающегося. [Электронный ресурс]. – Режим доступа: <https://ug.ru/individualnaya-obrazovatel'naya-traektoriya-obuchayushhegosya/> (дата обращения: 08.01.2022).
2. Индивидуальная образовательная траектория. [Электронный ресурс]. – Режим доступа: <https://strategium.space/lesson/personal-education-trajectory/> (дата обращения: 08.01.2022.)
3. John McCarthy "What is artificial intelligence?" [Электронный ресурс]. – Режим доступа: https://borghese.di.unimi.it/Teaching/AdvancedIntelligentSystems/Old/IntelligentSystems_2008_2009/Old/IntelligentSystems_2005_2006/Documents/Symbolic/04_McCarthy_whatissai.pdf (дата обращения: 08.01.2022).
4. Distinguishing between Narrow AI, General AI and Super AI. [Электронный ресурс]. – Режим доступа: <https://medium.com/mapping-out-2050/distinguishing-between-narrow-ai-general-ai-and-super-ai-a4bc44172e22> (дата обращения: 08.01.2022).
5. Deep learning vs. machine learning – What's the difference? [Электронный ресурс]. – Режим доступа: <https://levity.ai/blog/difference-machine-learning-deep-learning> (дата обращения: 08.01.2022).
6. Machine Learning. [Электронный ресурс]. – Режим доступа: <https://www.ibm.com/cloud/learn/machine-learning>. (дата обращения: 08.01.2022).
7. Difference Between Machine Learning and Deep Learning. [Электронный ресурс]. – Режим доступа: <https://www.geeksforgeeks.org/difference-between-machine-learning-and-deep-learning/> (дата обращения: 08.01.2022).
8. Neural Networks. [Электронный ресурс]. – Режим доступа: <https://www.ibm.com/cloud/learn/neural-networks> (дата обращения: 08.01.2022).
9. Как работают нейронные сети? [Электронный ресурс]. – Режим доступа: <https://progkids.com/blog/how-neuron-network-works> (дата обращения: 08.01.2022).
10. Data mining: выбор приоритетов и нейронные сети. [Электронный ресурс]. – Режим доступа: <https://softline.ru/about/blog/data-mining-vybor-prioritetov-i-neyronnyie-seti> (дата обращения: 08.01.2022).
11. Создаём нейронную сеть на Python с нуля. [Электронный ресурс]. – Режим доступа: <https://otus.ru/nest/post/736/> (дата обращения: 08.01.2022).
12. Weight (Artificial Neural Network). [Электронный ресурс]. – Режим доступа: <https://deeprai.org/machine-learning-glossary-and-terms/weight-artificial-neural-network> (дата обращения: 08.01.2022).
13. Linear regression. [Электронный ресурс]. – Режим доступа: <https://www.ibm.com/topics/linear-regression> (дата обращения: 08.01.2022).
14. CRISP-DM: Towards a Standard Process Model for Data Mining. [Электронный ресурс]. – Режим доступа: <http://www.cs.unibo.it/~montesi/CBD/Beatriz/10.1.1.198.5133.pdf> (дата обращения: 08.01.2022).
15. 6 шагов по созданию проектов машинного обучения. [Электронный ресурс]. – Режим доступа: <https://tproger.ru/translations/6-step-for-building-machine-learning-projects/> (дата обращения: 08.01.2022).
16. Машинное обучение для чайников. [Электронный ресурс]. – Режим доступа: <https://newtonew.com/tech/machine-learning-novice> (дата обращения: 08.01.2022).

УДК 004.925, 4.942

ИССЛЕДОВАНИЕ НЕЙРОСЕТЕВЫХ ТЕХНОЛОГИЙ ДЛЯ СОЗДАНИЯ УСЛОВИЙ, СООТВЕТСТВУЮЩИХ ЕСТЕСТВЕННОЙ АККОМОДАЦИИ

Кинёв И.Е.¹ (студент), Гебель Г.В.¹ (студент)

Научный руководитель - канд. физ.-мат. наук, доцент Жданов Д.Д.¹

¹Университет ИТМО

e-mail: igorkinevitmo@gmail.com, gebel.german@yandex.ru, ddzhdanov@mail.ru

Работа выполнена в рамках темы НИР №620164 «Методы искусственного интеллекта для киберфизических систем».

Проведено исследование методов по снижению дискомфорта зрительного восприятия, возникающего при длительном использовании устройств виртуальной реальности. Основное решение заключается в попытке сформировать изображение виртуальных объектов, в соответствии с естественной аккомодацией, что помогает устранить конфликт вергенции-аккомодации (VAC). Рассмотрены методы решения с помощью голографических, томографических, ретинальных микродисплеев и нейронных сетей.

Ключевые слова: конфликт вергенции-аккомодации, виртуальная/дополненная реальность, нейронные сети, система трекинга глаз, микродисплеи.

При длительном использовании устройств виртуальной и дополненной реальности у человека может появиться дискомфорт зрительного восприятия, проявляющийся в виде рези в глазах, спазма мышц, двоения изображения, тошноты, дезориентации и т.д. Данные симптомы могут проявиться из-за так называемого конфликта вергенции-аккомодации (vergence-accommodation conflict), когда вид изображения не соответствующий условиям вергенции-аккомодации человеческого зрения.

При наблюдении двумя глазами мозг объединяет два изображения одного и того же объекта в одно пространственное изображение. Однако, при этом должны быть выполнены определенные условия, что обеспечивается согласованием аккомодации точки конвергенции глаз человека.

В работе [1], автор описал проблемы, которые возникают при использовании систем виртуальной реальности. Первая заключается в том, что при формировании виртуального изображения частично (не полностью) учитывается движения зрачков глаз при смене их положения. На рисунке 1 взятым из этой работы можно увидеть, что на фрагменте (а) положение изображений двух различных объектов А и В на сетчатке глаза зависит от ориентации центров зрачков С и С'. Данные точки могут изображаться либо на одну точку сетчатки глаза, либо на две. Однако, на фрагменте (б) показано, что если задана статическая плоскость носителя, то точки, разнесенные в пространстве, проецируются не на две точки сетчатки, а на одну. Этот эффект приводит к дискомфорту при просмотре изображения.

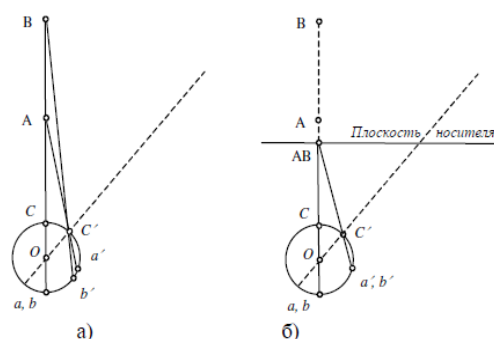


Рис. 1. Влияние оптического центра глаза на видимость предметных точек: а) в пространстве; б) на проекции [1]

Второй проблемой является вышеописанный конфликт вергенции-аккомодации. При фиксации глаза на объекте сцены возникает рефлекторная зависимость между аккомодацией глаз и ориентацией их зрачков. То есть, наблюдаемый объект находится в фокусе, а ближний и дальний планы при этом размываются.

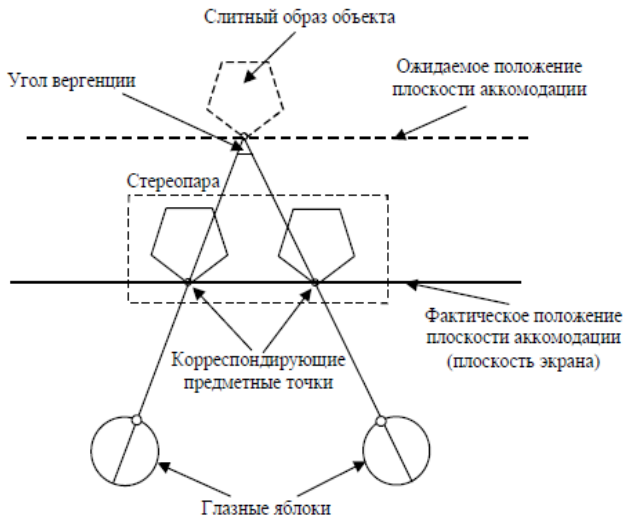


Рис. 2. Несоответствие угла вергенции и положения плоскости аккомодации [1]

При наблюдении изображения на плоском экране данная зависимость нарушается, что приводит к появлению дискомфорта. На рисунке 2 приведен пример, показывающий несоответствие между вергенцией и аккомодацией человеческого зрения.

Проблема VAC в таких дисплеях связана с тем фактом, что источником изображения в основном является плоская двумерная поверхность, расположенная на фиксированном расстоянии от глаза. Таким образом, комфортом для глаз будет, если при просмотре на дальние предметы, размываются ближние объекты и наоборот, при наблюдении дальних объектов размываются ближние. Данный эффект показан на рисунке 3.

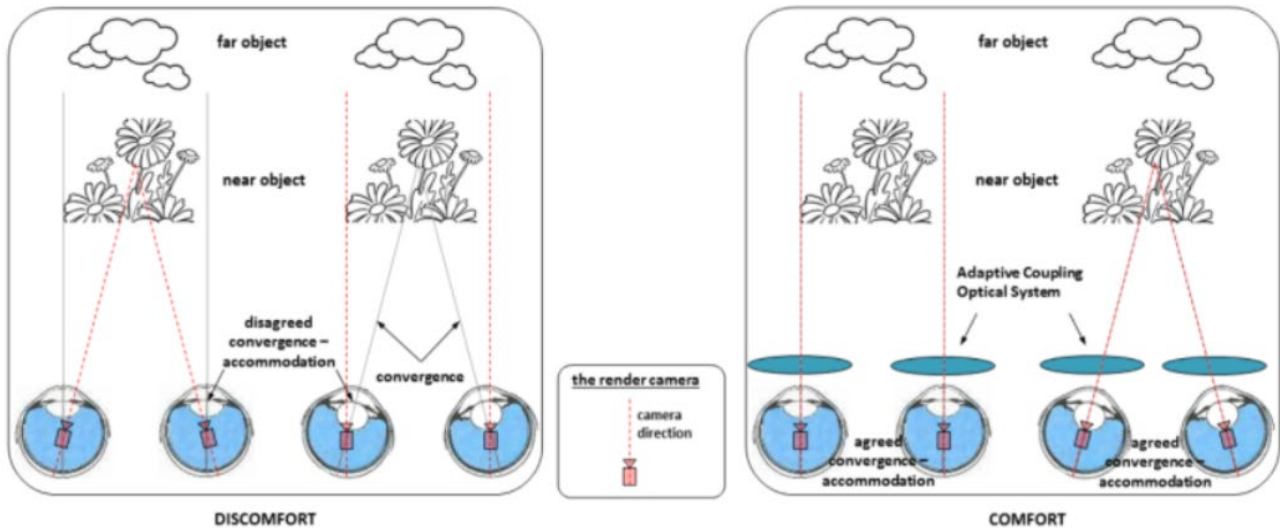


Рис. 3. Комфорт и дискомфорт от просмотра изображения в системе виртуальной реальности



Рис. 4. Сцена

Реальные условия аккомодации создаются следующим образом: при фокусировании на объекте, наш мозг воспринимает размытие объектов, не находящихся в области конвергенции, как естественный эффект, не вызывающий конфликта и способствующий возникновению стереоскопического восприятия. Рисунки 4–5 схематично иллюстрируют естественно связанные действия аккомодации и конвергенции глаз наряду с эффектом размытия сетчатки при просмотре естественной сцены.

Глубина вергенции глаз совпадает с глубиной их аккомодации, и размытость изображения объектов на сетчатке увеличивается с расстояниями от точки фиксации глаза до других точек на разной глубине.

Сцена реализована в программе Lumincept. Рассчитаны оптические показатели путём моделирования методом обратной трассировки лучей при различных фокусных расстояниях:

- 28 м конус;
- 21 м сфера;
- 13 м куб.

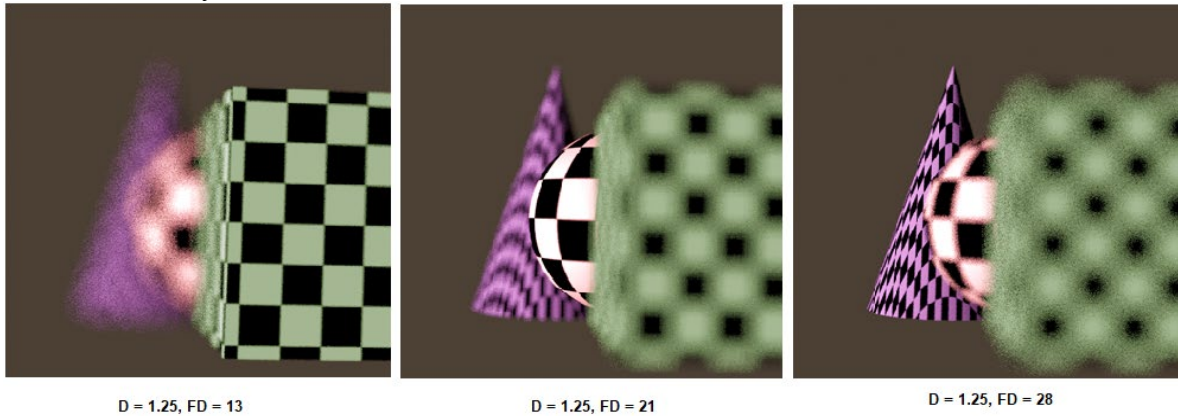


Рис. 5. Естественно связанные действия аккомодации и конвергенции глаз в реальном мире: глубина вергенции глаза, где пересекаются зрительные оси, совпадает с глубиной его аккомодации, а размытие изображения объектов на сетчатке изменяется в зависимости от глубины аккомодации глаза (куб, сфера и конус)

В большинстве случаев для дополненной и виртуальной реальности используется близкие к глазу дисплеи (near-eye display), которые формируют два изображения, для левого и правого глаз. С помощью соответствующих окуляров человек воспринимает изображения как одно и видит реалистичную и объемную сцену. Для внедрения устройств дополненной и виртуальной реальности в повседневную жизнь необходимо устранить эффекты, вызывающие у людей чувство дискомфорта. На данный момент прототипы соответствуют требованиям к аппаратной части устройства, таким как задержка отклика и пространственное разрешение. Однако, их также необходимо создание системы автоматической фокусировки окуляров в соответствии с состоянием вергенции и аккомодации глаз человека [2]. В статьях [3][4][5] авторы предлагают использовать ретинальные дисплеи для решения проблемы VAC. Преимущество данных дисплеев с голографическими оптическими элементами заключается в том, что они имеют малый объем и вес, что позволяет человеку носить устройства виртуальной и дополненной реальности дольше. Также в статье [6] авторы предложили компактный голографический дисплей близкий к глазу для виртуальной и дополненной реальности, основанный на фазовой голографической проекции. Кроме того, нововведением являются томографические дисплеи [7].

Существуют методы решения конфликта VAC с помощью нейросетевых технологий. В работе [8] авторы предлагают архитектуру нейронной сети DeepFocus. Данная нейронная сеть принимает на вход изображение с картой глубин, с помощью которой определяется расстояние до объекта. Далее по заданной точке на изображении и её глубине, формируется эффект размытия, который создает условия близкие к естественной аккомодации человеческого глаза, пример представлен на рисунке 6.

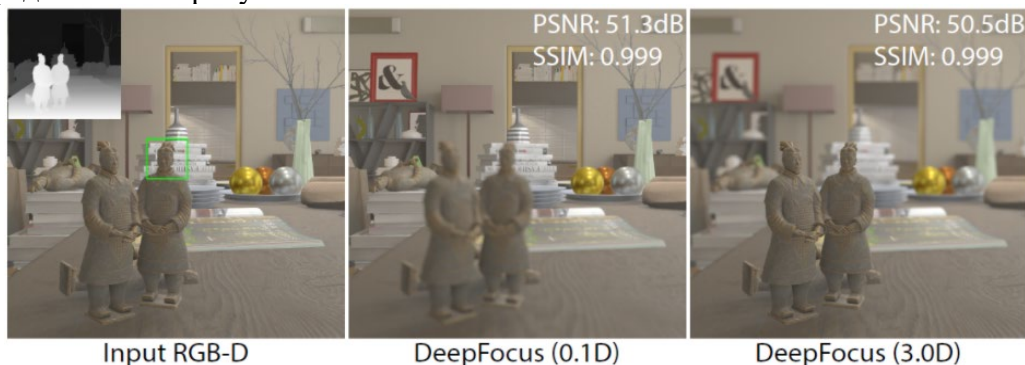


Рис. 6. Создание условий естественной аккомодации с помощью нейронной сети Deep Focus

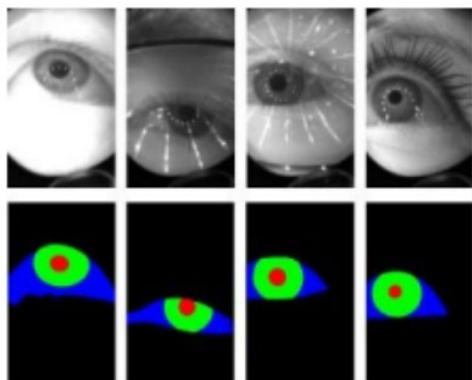


Рис. 7. OpenEDS Semantic Eye Segmentation Challenge датасет

Чтобы такая система работала в устройствах виртуальной реальности, необходимо отслеживать направление взгляда, т.е., на какую точку изображения направлен взгляд. Также необходимо анализировать положение зрачка и радужки, для поиска точки фокусировки глаз. Для решения такой задачи компанией Facebook было создано соревнование “OpenEDS Semantic Eye Segmentation Challenge”, где предоставляется достаточно большой набор данных, чтобы сегментировать нужные области глаз (рис. 7). Также компания предложила свое решение данной задачи с помощью нейронной сети EyeNet, которая справляется с задачей с высокой точностью [9].

Таким образом, если использовать данную систему сегментации глаз и нейронную сеть для создания условий естественной аккомодации, получается система, которая сможет убрать дискомфорт при использовании устройств виртуальной реальности.

Проведенное исследование дает представление о том, что вызывает конфликт вергенции-аккомодации, и какие методы существуют для решения данной проблемы. Также был исследован метод трекинга (сегментации) глаз. Данный метод будет полезен, для определения точки фокусирования. В качестве продолжения данной работы планируется создание алгоритма адаптивного фильтра дефокусировки изображения и сравнение производительности и качества его работы с методами дефокусировки, основанными на нейронных сетях. Также планируется провести исследование влияния ошибки определения карты глубин сцены на корректность работы алгоритма дефокусировки изображения.

Литература

1. Афанасьев В.О. Развитие модели формирования бинокулярного изображения виртуальной 3D-среды. // Программные продукты и системы. – 2004. – №4. – С. 27–32.
2. Lee B., Jo Y., Yoo D., Lee J. Recent progresses of near eye display for AR and VR // SPIE Optical Metrology. – 2021. – V.11785. – Pp. 1–6.
3. Jeong J., Lee J., Yoo C., Moon S., Lee B., Lee B. Holographically customized optical combiner for eye-box extended near-eye display. // Opt. Express. – 2019. – V.27. – Pp. 38006–38018.
4. Jo Y., Yoo C., Bang K., Lee B., Lee B. Eye-box extended retinal projection type near-eye display with multiple independent viewpoints. // Appl. Opt. – 2021. V.60. – Pp. 268–276.
5. Jang C., Bang K., Li G., Lee B. Holographic near-eye display with expanded eye-box. // ACM Trans. Graph. – 2018. – V.37. – Pp. 1–14.
6. Maimone A., Georgiou A., Kollin J.S. Holographic near-eye displays for virtual and augmented reality. // ACM Trans. Graph. – 2017. – V.36. – Pp. 1–16.
7. Lee S., Jo Y., Yoo D., Cho J., Lee D., Lee B. Tomographic near-eye displays. // Nature Comm. – 2019. – V.10. – Pp. 2497–2507.
8. Lei Xiao, Anton Kaplanyan, Alexander Fix, Matthew Chapman, Douglas Lanman. DeepFocus: learned image synthesis for computational displays. // ACM Transactions on Graphics. – 2018. – V.36. – Pp. 1–13.
9. Kansal P. Devanathan S. EyeNet: and Attention based convolutional encoder-decoder network for eye region segmentation // IEEE/CVF International Conference on Computer Vision Workshop (ICCVW). – 2019. – Pp. 3688–3693.

УДК 004.055

МОНЕТИЗАЦИЯ НОВЫХ МОБИЛЬНЫХ ИГР

Клокова Е.И.¹ (студент)

Научный руководитель – канд. педаг. наук, доцент Авксентьева Е.Ю.¹

¹Университет ИТМО

e-mail: lizaklokova@mail.ru

В работе было проведено исследование существующих методов монетизации мобильных игр, в рамках которого все изученные методы были разделены на несколько групп. На основе анализа полученных групп был определён набор базовых методов монетизации мобильных игр, обладающих свойством универсальности. Также были изучены различные ключевые показатели эффективности приложений и выбраны те из них, которые лучшим образом подходят для оценки и последующего улучшения выбранной стратегии монетизации.

Ключевые слова: монетизация, мобильные игры, ключевые показатели эффективности, прибыль, реклама.

В современном мире мобильные телефоны стали неотъемлемой частью жизни практически каждого человека. Одной из самых обширных областей использования мобильных телефонов стали мобильные игры. В связи с этим в определенный момент разработчики игр решили монетизировать эту повсеместно распространённую сферу развлечений. Каждый день на рынке появляются новые мобильные игры со своим планом получения прибыли от разработанного продукта. Однако, не все из них действительно получают заработок со своего приложения. Надоедливая реклама, постоянное вымогательство денег, платная подписка – всё это и многое другое лишь отпугивает потенциальных покупателей, которые просто удаляют приложения.

Именно поэтому целью данной работы стало нахождение набора базовых методов монетизации, с которых стоит начинать стратегию получения прибыли от разработанной мобильной игры. Для максимизации получаемой прибыли необходимо будет использовать сразу несколько подходящих методов монетизации, но при этом важно не нарушить баланс, когда пользователи будут уходить из игры из-за большого количества платного контента или надоедливой рекламы. Выбранные методы монетизации должны обладать свойством универсальности, то есть подходить к абсолютно любой игре, даже новой от малоизвестного на рынке мобильных игр разработчика.

Для начала нужно было чётко определиться с понятием монетизации игры. Монетизация – это совокупность процессов получения прибыли от разработанного продукта. Соответственно, монетизация игры предполагает получение прибыли от разработанной игры.

Для поиска необходимого начального набора методов монетизации для новой разрабатываемой игры нужно было провести исследование существующих методов монетизации. На сегодняшний день отрасль мобильных игр разрослась настолько широко, что появилось огромное количество различных методов монетизации мобильных игр. После проведенного исследования, на основе изученной информации из источников был составлен список существующих методов монетизации мобильных игр [1–3]. В связи с тем, что некоторые из приводимых методов монетизации являются слишком схожими и имеют какую-либо общую основу, было принято решение объединить такие методы в группы, представленные в таблице.

Таблица. Группы методов монетизации мобильных игр

№	Название группы	Методы монетизации
1	Предметы	«Внутриигровая валюта», «Дополнительные предметы», «Акционные предметы», «Случайные предметы», «Дополнительные локации, уровни», «Лимитированные предметы», «Всплывающее окно подсказки», «Подарок при покупке», «Предложения пользователей», «Дополнительное время»
2	Реклама	«Баннеры», «Полноэкранная реклама», «Видео-реклама»
3	Персонажи	«Дополнительные персонажи», «Улучшение персонажа»
4	Партнёры	«Партнёрские игры», Спонсорские предметы, «Партнёрские купоны», «Инвестиции»
5	Подписка	«Платная игра», «Платная игра без рекламы», «Бонусы в платном аккаунте»
6	Продажа ресурсов	«Продажа исходного кода», «Продажа пользовательских данных», «Сдача в аренду пользовательских данных»
7	Взаимодействие пользователей	«Турниры», «Ставки»
8	Продажа символики	Футболки, кружки, игрушки, аксессуары для компьютеров и телефонов и др.
9	Донаты	Любая сумма денег

Получившаяся группа «Предметы» предполагает получение игроками различных внутриигровых предметов, бонусов или преимуществ за внесение определённой платы. Группа «Реклама» предполагает получение прибыли не от игроков, а от рекламодателей, которые будут платить деньги за количество пользователей, которые будут смотреть рекламу, кликать по ней и приобретать предлагаемый товар. Некоторые методы получения прибыли были объединены в группу «Персонажи», так как они направлены на покупку игроками различных персонажей и их улучшений. Группа «Партнёры» предполагает привлечение партнёров, спонсоров или инвесторов, которые будут готовы заплатить деньги создателю мобильной игры в ожидании приумножения их прибыли напрямую или с помощью привлечения новых пользователей к их продукции. Несколько методов монетизации мобильных игр были объединены в группу «Подписка», которая предполагает получение игроком большого количества дополнительных бонусов на длительный период времени за более крупную сумму денег. Группа «Продажа ресурсов» предполагает наличие определённых ресурсов у создателя игры, которые он может продавать в сторонние организации. Это не самый честный способ получения прибыли от мобильной игры, но он существует в современном мире. Следующая группа «Взаимодействие пользователей» включает в себя различные способы взаимодействия между разными пользователями для получения прибыли от мобильной игры. В качестве ещё одной группы методов монетизации была выделена группа «Продажа символики», которая предполагает запуск собственного производства брендовых вещей с символикой игры после её успешного запуска и получения достаточной аудитории. В качестве брендовых вещей могут выступать футболки, кружки, игрушки, аксессуары для компьютеров и телефонов и многое другое. Последняя рассмотренная группа «Донаты» предполагает добровольные безвозвратные пожертвования пользователями определённой суммы денег на развитие и поддержку игры. Пользователь сам выбирает, какую сумму безвозмездно отдать игре.

После изучения существующих методов монетизации необходимо было найти те из них, которые составят набор базовых методов монетизации, обладающих свойством универсальности. Т.е. найти такие методы монетизации, которые могут быть применимы в качестве способа получения прибыли из абсолютно любой игры (независимо от её жанра, количества уровней или пользователей), в том числе из абсолютно новой мобильной игры от неизвестных разработчиков.

После проведённого исследования был составлен следующий набор базовых методов монетизации: это методы монетизации из групп «Предметы», «Реклама», «Донаты». Однако, в случае успеха выпущенной мобильной игры, увеличения числа пользователей, привлечения

сторонних инвесторов или наступления других факторов, в конкретную разрабатываемую игру могут быть добавлены дополнительные методы монетизации из групп «Персонажи», «Партнёры», «Подписка», «Продажа ресурсов», «Взаимодействие пользователей» и «Продажа символики».

Таким образом, после проведённого анализа существующих методов получения прибыли от мобильных игр был определён набор базовых методов монетизации, которые можно внедрять в новое разработанное приложение. Однако, нельзя просто взять и добавить выбранный метод монетизации. Например, сложно определить стоимость подсказки в логической игре, когда пользователь зашёл в тупик. Ведь при неправильном использовании метода монетизации пользователи начнут уходить из игры, а, следовательно, прибыль, получаемая от разработанного продукта, будет уменьшаться.

Чтобы лучше понимать, устраивает ли пользователей разработанная мобильная игра с её стратегией монетизации, были придуманы специальные ключевые показатели эффективности мобильных игр. На сегодняшний день существует большое число таких показателей, однако, для решения поставленной задачи балансировки базовых методов монетизации в качестве наиболее информативных показателей по результатам агрегации информации из нескольких изученных источников были выбраны следующие ключевые показатели эффективности: Daily Active Users (DAU – Ежедневные активные пользователи), Retention Rate (RR – Коэффициент удержания), Churn Rate (CR – Коэффициент оттока), Average Revenue Per User (ARPU – Средний доход на пользователя), Lifetime Value (LTV – Пожизненная ценность), Session Length (SL – Продолжительность сеанса) [4–5].

При внедрении в приложение определённого метода монетизации необходимо снять выбранные метрики за некоторый константный период времени. После любого изменения параметров метода монетизации через выбранный промежуток времени необходимо заново рассчитать все указанные выше показатели эффективности и сравнить с исходными. Если значения выбранных метрик значительно ухудшились, стоит вернуться к предыдущей стратегии методов монетизации. Если они не сильно изменились, то можно продолжать изменять параметры выбранных методов получения прибыли с целью её максимизации.

Таким образом, в рамках проведения данной работы были проанализированы различные методы получения прибыли от мобильных игр, на основе которых был составлен набор базовых методов монетизации, подходящих к абсолютно любой мобильной игре, в том числе к новой разрабатываемой игре. Для балансировки использования выбранных методов монетизации и максимизации прибыли от мобильной игры были изучены различные ключевые показатели эффективности мобильных приложений и выбраны наиболее подходящие, значения которых необходимо отслеживать при изменении стратегии монетизации. Дальнейшая работа по данной теме предполагает тестирование выбранных методов монетизации на мобильной игре с целью нахождения оптимальной стратегии внедрения методов монетизации, при которой прибыль будет максимальна, а выбранные ключевые показатели эффективности приложения будут лучшими на всей исследуемой выборке.

Литература

1. Appel G., Libai B., Muller E., Shachar R. On the monetization of mobile apps // *International Journal of Research in Marketing*. – 2020. – №37(1). – С. 93–107.
2. 59 Ways To Monetize Your Indie Game. [Электронный ресурс]. – Режим доступа: <http://gdu.io/blog/monetization/> (дата обращения: 09.12.2022).
3. Mobile Game Monetization Trends: Best Strategies to Monetize Your Game in 2022. [Электронный ресурс]. – Режим доступа: <https://medium.com/udonis/mobile-game-monetization-trends-best-strategies-to-monetize-your-game-in-2020-fe2de42d4463> (дата обращения: 12.12.2021).
4. 13 Key Mobile Game Metrics That Developers MUST Track. [Электронный ресурс]. – Режим доступа: <https://www.blog.udonis.co/mobile-marketing/mobile-games/key-mobile-game-metrics> (дата обращения: 24.12.2021).
5. 50+ KPIs to Measure Your Mobile Game or App. [Электронный ресурс]. – Режим доступа: <https://gameanalytics.com/blog/50-kpi-measure-mobile-game-app/> (дата обращения: 27.12.2021).

УДК 004.8

ОБЗОР ПРАКТИКИ ПРИМЕНЕНИЯ ТЕХНОЛОГИЙ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА ДЛЯ АНАЛИЗА ПОВЕДЕНИЯ ПОЛЬЗОВАТЕЛЕЙ ТУРИСТИЧЕСКИХ СЕРВИСОВ

Козырева А.И.¹ (студент)

Научный руководитель – канд. экон. наук. Кононова О.В.¹

¹*Университет ИТМО*

e-mail: alyonakozyr21@gmail.com

Работа выполнена в рамках темы НИР №621300 «Разработка сервиса оценки доступности объектов городского благоустройства для маломобильных групп населения на основе данных из социальных сетей».

Цель работы заключается в изучении понятия поведения пользователей по отношению к туристическим услугам, а также существующих и используемых в настоящее время методов его анализа. Для достижения поставленной цели предлагается определить значение понятия поведения пользователей интернет-ресурсов, выявить возможности измерения данных о поведении пользователей, а также изучить, какие технологии искусственного интеллекта можно использовать для анализа поведения пользователей.

Результат исследования позволит сформировать представления о возможностях анализа пользовательского поведения и задать вектор дальнейших исследований. В ходе работы предстоит определить и описать методы анализа поведения пользователей в настоящее время, а также изучить возможности проведения анализа поведения с использованием искусственного интеллекта.

Ключевые слова: умный туризм, электронные сервисы, поведение пользователей, технологии искусственного интеллекта, интернет-коммуникация, цифровая культура.

В настоящее время невозможно представить культуру и туризм без использования электронных сервисов. Спрос на туристические услуги растет еще и потому, что большинство туристов пользуются ими при планировании своего путешествия. В фокусе развития туризма в государственной программе развития до 2030 года [1] весьма актуальным является развитие культурно-познавательного туризма. Это приводит к широкому спектру исследований по анализу, мониторингу и оценке поведения пользователей.

Сфера туризма в целом все больше подвергается цифровизации, особенно на городском уровне. Туристических сервисов становится больше, а число их пользователей растет. Это подтверждает посещаемость туристических веб-ресурсов и количество HTTP-запросов к веб-ресурсам по бронированию, которые на начало августа 2020 года увеличились более чем на 14% по сравнению со значениями на начало февраля 2020 [2]. Эту статистику подтверждает и бюллетень о текущих тенденциях в российской экономике. Так, согласно исследованию Яндекса по туризму в 2020 году, за последние три года доля запросов с мобильных устройств удвоилась до 54%, в том числе запросы с мобильных устройств продолжали расти в первом квартале [3]. В исследовании Сбербанка за октябрь 2021 г. рассчитывался туристический индекс на основе таких параметров, как количество уникальных туристов, сумма их трат, доля лояльных туристов, средняя продолжительность поездки, выручка компаний туристической отрасли [4]. Этот показатель также показывает постоянный рост с начала 2018 года по настоящее время.

Технологии искусственного интеллекта в настоящее время являются одним из наиболее распространенных и удобных инструментов анализа больших данных, в том числе в туристической отрасли. Искусственный интеллект и статистические методы позволяют систематизировать, просчитывать и прогнозировать поведение пользователя и в результате получить срез, график или статистику. Методы интеллектуального анализа данных, такие как классификация, кластеризация, частые наборы элементов, обнаружение циклов и аномалий, уже широко используются для извлечения и анализа траектории поведения пользователей в индустрии туризма [5]. Полученные поведенческие данные можно в дальнейшем использовать для предсказания и прогнозирования, что также находит применение во многих областях.

Целью работы является изучение понятия пользовательского поведения по отношению к туристическим услугам, а также существующих и применяемых в настоящее время методов его анализа.

Первичным результатом анализа выбранных источников стало определение понятия поведения пользователя. Востребованность подобных исследований в настоящее время приводит к появлению широкого круга исследований по анализу, мониторингу и оценке поведения пользователей. Поведение пользователя может означать активность на сайте или в приложении, эмоции при использовании сервиса. Также оно может рассматриваться в контексте личных предпочтений, потребительских запросов и поведенческих характеристик пользователей туристических услуг. Анализ поведения проводится для того, чтобы понять, в чем чаще всего проявляются проблемы и решить их, чтобы сделать использование таких сервисов максимально комфортным для большинства пользователей.

В данной работе под поведением пользователей в сети будет пониматься:

- совершение сделок;
- переходы и возвраты к ресурсу;
- время, проведенное на сервисе;
- запрашиваемые и приобретаемые услуги и продукты, а также суммы трат на них;
- оценочные проявления: отзывы, лайки, результаты опроса и т.д.

Таким образом, анализ поведения пользователей должен начинаться с отслеживания интересов и предпочтений, а также частоты использования сервисов. Для анализа поведения пользователей могут использоваться как текстовые, так и нетекстовые данные. В качестве текстовых данных могут быть упомянуты отзывы и оценочные проявления. Нетекстовые данные включают количественные данные в формате времени, денег и чисел. Оба типа этих данных требуют различных инструментов анализа.

Был изучен опыт использования технологий искусственного интеллекта в электронном туризме для анализа поведения пользователей как в России, так и за рубежом. Одним из малоизученных, но в то же время достаточно перспективных направлений анализа явился анализ символов и знаков, используемых посетителями на сервисах. Их можно назвать современным индикатором поведения пользователей онлайн-сервисов.

Чтобы понять общую картину, была построена сводная таблица с некоторыми исследованиями поведения пользователей с использованием кластеризации (табл. 1). Из таблицы видно, что количество типов в рассмотренных исследованиях сильно различается, и в большинстве исследований анализируются эмоции, а в качестве методов используется искусственный интеллект, что подтверждает актуальность использования технологий искусственного интеллекта в сфере анализа поведения пользователей туристических услуг.

На основе анализа предметной области были также сформулированы возможные методы анализа поведения пользователей. Выделены конкретные моменты поведения, что именно под этим подразумевается; были выдвинуты предложения о том, как это можно измерить, а также о возможном результате и практической значимости, которые может дать исследование. Была составлена сводная таблица с предложенными методами (табл. 2). Из таблицы видно, что можно анализировать совершенно разные типы данных, которые можно разделить на текстовые и нетекстовые. Результаты анализа данных могут применяться в различных сферах деятельности и полезны как для пользователей, так и для государства.

Таблица 1. Сводная таблица исследований, использующих кластеризацию

Тема исследования	Количество групп/кластеров	Названия этих групп/кластеров	Используемые технологии	Факторы и признаки
Сентимент-анализ	8	Интерес, Радость, Удивление, Тоска, Гнев, Страх, Отвращение, Стыд	Разработанный алгоритм бинарной классификации	Лексемы с заглавными буквами и фиксирующие цифры, буквенно-цифровые обозначения и эмодзи
Поведенческие профили потребителей онлайн-турагентств	2	Удовлетворен/ Не удовлетворен	Кластерный анализ, алгоритм k-средних	Чувство безопасности, функциональность, простота использования, полезность, лояльность, качество информации, репутация и востребованность, удовлетворенность
Инструмент аналитики Google и его практическое применение в туристическом бизнесе	—	—	Инструменты веб-анализа, карты активности посетителей, Google Analytics	Движения мыши, щелчки, нажатия клавиш, прокрутка
Туристическая типология онлайн и личных социальных контактов	6	Путешественник с полным погружением, путешественник с цифровой детоксикацией, путешественник с отвлекающим маневром, путешественник с двумя зонами, контролер повседневной жизни и зависимый от социальных сетей	Проведение интервью	Мотивация путешествия, онлайн-социальный контакт, личный социальный контакт, роль туристического направления и режим туристического опыта

Таблица 2. Сводная таблица методов

Поведение	Измерения	Результаты
Анализ запросов, перехода (интереса) к товарам/услугам или категориям товаров, анализ покупок/заказов, анализ активности, транзакциям и суммам, региону	Статистика: агрегированные данные по пользователям	Маркетинговая стратегия, закупки, логистика, создание условий для формирования и продвижения качественного туристского продукта, конкурентоспособного на внутреннем и мировом рынках, стимулирование спроса на туристский продукт
	Данные по конкретному пользователю	Рекомендации/предложения/ рассылка
Анализ работы пользователя с сайтом/ сервисом по категориям пользователей (возраст, социальный/профессиональный статус, другие характеристики)	Время задержки на странице (работы с сервисом), маршруты, повторные обращения, количество обращений, закончившихся получением услуги; анализ жалоб и предложений по работе ресурса	Модернизация интерфейса, улучшение функциональности платформы/портала/сайта; снижение транзакционных издержек; повышение информированности участников цифровых сервисов и платформ совершенствование электронных форм государственных услуг и электронных форм отчетности, связанных с туристской деятельностью; интеграция ГИС, связанных с обеспечением туристской деятельности; совершенствование цифровых решений взаимодействия с предпринимательским и экспертным сообществом
Анализ настроений (портал/сайт/социальные сети/карты)	Лайки/текстовые сообщения/фото	Выявление потребностей/ожиданий, отношения к ресурсу или услуге: усиление социальной роли туризма, увеличение доступности услуг туризма, отдыха и оздоровления для жителей РФ
Анализ обращений, степени вовлеченности в реализацию городских инициатив, активности граждан, индивидуальных инициатив	Запросы/отзывы/инициативы Обсуждения/статистика/опросы	Просвещение, обучение, развлечение (культурно-познавательный туризм), оздоровление/ социальная помощь: некоммерческие сервисы, городские сервисы (лечебно-оздоровительный туризм), сбор научной информации, проведение научных исследований, посещение/освещение научных мероприятий (научный туризм, промышленный туризм / деловой туризм); развитие электронной системы оценки качества туристских услуг; создание рейтинга туристских услуг и объектов
Прогнозирование поведения	Психоанализ инфо из аккаунтов социальных сетей, комментариев в отзывах, приобретаемых пакетов услуг	Противодействие экстремизму, нарушению правопорядка, загрязнение территорий (экологический туризм)

Было изучено использование смарт-технологий в туризме как в мире, так и в России. Потенциал умного туризма в России высок, а использование интеллектуальных систем и технологий со временем будет расти, что говорит об актуальности их использования, в том числе для аналитики туристических сервисов и поведения пользователей на них.

Рассмотрены существующие исследования типологии пользователей, типов их поведения при использовании туристических услуг и влияния эмоциональной составляющей их поведения, в том числе установок, на характеристики символов, наиболее часто используемых в онлайн-общении. Также были выделены методы анализа тональности текстов, а также методы веб-аналитики.

На основании рассмотренных исследований и их сравнительного обзора были сделаны предложения по дальнейшему использованию результатов данного исследования. В дальнейшем планируется начать изучение поведения пользователей туристических услуг по отношению к русскоязычному сегменту, а именно к конкретным сервисам Санкт-Петербурга и Ленинградской области.

Литература

1. Распоряжение Правительства РФ от 20.09.2019 № 2129-р. «Об утверждении Стратегии развития туризма в Российской Федерации на период до 2035 года». [Электронный ресурс]. – Режим доступа: <http://static.government.ru/media/files/FjJ74rYOaVA4yzPAshEulYxmWSpB4lrM.pdf>. (дата обращения: 20.02.2022).
2. Трафик веб-ресурсов туристического сегмента возвращается к доковидным значениям. [Электронный ресурс]. – Режим доступа: <https://ngenix.net/news/trafik-veb-resursov-turisticheskogo-segmenta-vozvrashhaetsya-k-dokovidnym-znacheniyam> (дата обращения: 20.02.2022).
3. Исследование Яндекса: туризм в 2020 году. [Электронный ресурс]. – Режим доступа: <https://hoteliernews.ru/turizm-v-2020-ot-kaliningrada-do-kamchatki> (дата обращения: 20.02.2022).
4. СберИндекс: Индекс внутреннего туризма. октябрь 2021. [Электронный ресурс]. – Режим доступа: <https://sberindex.ru/ru/dashboards/indeks-vnutrennego-turizma-pomesyachnaya-dinamika> (дата обращения: 20.02.2022).
5. Zhang J., Wu T., Fan Z. Research on Precision Marketing Model of Tourism Industry Based on User's Mobile Behavior Trajectory // Mobile Information Systems. – 2019. – 6560848. DOI: 10.1155/2019/6560848.

УДК 004.04

РАЗРАБОТКА МЕТОДА ВЫБОРА УЧАСТКОВ УЛИЧНО-ДОРОЖНОЙ СЕТИ ДЛЯ СОЗДАНИЯ ПЕШЕХОДНЫХ ОБЩЕСТВЕННЫХ ПРОСТРАНСТВ НА ПРИМЕРЕ САНКТ-ПЕТЕРБУРГА

Красноперова Ю.А.¹ (студент), Соловьёва В.В.¹ (аспирант)

Научный руководитель – канд. техн. наук Митягин С.А.¹

¹Университет ИТМО

e-mail: jkras1912@gmail.com

Работа посвящена описанию разработанного метода по выбору участков улично-дорожной сети для создания пешеходных общественных пространств. Сформирована проблематика и актуальность исследования. Определены критерии оценки потенциала и необходимости ограничения движения на улице и создания общественного пространства. Приведена апробация разработанного метода на примере Центрального района Санкт-Петербурга.

Ключевые слова: ограничение автомобильного движения, пешеходизация, критерии оценки потенциала и необходимости ограничения движения на улице, городская среда, урбанистика.

Исследования показывают, что возможность пешей ходьбы является одним из составляющих пригодности для жизни городов, способствует созданию устойчивой городской среды [1]. Ограничение движения на одной или нескольких улицах в городе/районе и продуманное создание общественного пространства может стать живым экономическим капиталом, который важен для города [2]. Таким образом, становится актуальным поиск инструментов и разработка метода выбора улиц для ограничения транзитного автомобильного движения с применением комплексного анализа городской ситуации.



Рис. 1. Критерии (показатели) оценки потенциала и необходимости ограничения движения на улице и создания общественного пространства

В настоящее время в России не существует единого комплексного метода по определению наиболее подходящих и востребованных улиц для пешеходизации и возможности ее осуществления. Решения об ограничении движения принимаются экспертами в области транспортного планирования, архитектуры и урбанистики, а сами улицы предлагаются горожанами или городскими сообществами.

Для формирования полноценного общественного пространства на пешеходной улице необходимо определить наиболее важные критерии, анализ показателей которых выявит сильные и слабые стороны ограничения движения автомобилей на улице.

Основные критерии предложено разделить на три группы (рис. 1). В составе критерия «экология» предлагается учитывать такие экологические показатели, как уровень шумового загрязнения и уровень загрязнения выбросами автотранспорта.

Так, ограничение движения автомобилей на улице может способствовать улучшению экологической ситуации в городе, районе путем образования «холодных» коридоров, которые способствуют снижению температуры в границах городского теплового острова [3].

Критерий «инфраструктура» содержит такие оценочные показатели, как: обеспеченность района/города различными общественными пространствами на «открытом» воздухе, а также количество инфраструктурных учреждений, находящихся на предполагаемой будущей пешеходной улице. Первый показатель обосновывает необходимость увеличения количества общественных пространств за счет пешеходизации улицы, второй – отвечает за спрос и посещаемость данной улицы. Отсутствие автомобильного движения само по себе не сделает улицу привлекательной для пешеходов, людям нужна цель, чтобы прийти или пройти через данную улицу и задержаться на ней.

В составе критерия «транзит» предлагается учитывать показатели по транспортной загруженности, а также по популярности улиц у пешеходов, велосипедистов и людей, использующих средства индивидуальной мобильности. У улицы, как общественного пространства, есть такая важная функция, как обеспечение функционирования города за счет образования транспортно-коммуникационного каркаса вместе с другими элементами транспортной инфраструктуры. Поэтому важно, чтобы ограничение автомобильного движения на улице не ухудшило транспортную ситуацию на соседних улицах и не вызвало проблем с доступностью. Помимо исследования транспортных потоков необходимо также изучить пешеходные потоки. Будет преимуществом то, что будущая пешеходная улица является частью существующего транзитного пешеходного или велосипедного маршрута. Для этого также необходимо, чтобы у предполагаемой улицы были начальная и конечная точка на пути к объектам притяжения. Это могут быть станции метрополитена, площади, парки, детские площадки и другие общественные пространства.

В качестве метода выбора участков улично-дорожной сети для создания пешеходных общественных пространств предлагается разработка системы поддержки принятия решений об ограничении автомобильного движения и определения улиц для пешеходизации. Основными участниками системы являются представители администрации города/района с одной стороны, и сотрудники компаний, которые будут работать с территорией, с другой. Алгоритм работы данной системы можно увидеть на UML-диаграмме, изображенной на рисунке 2.

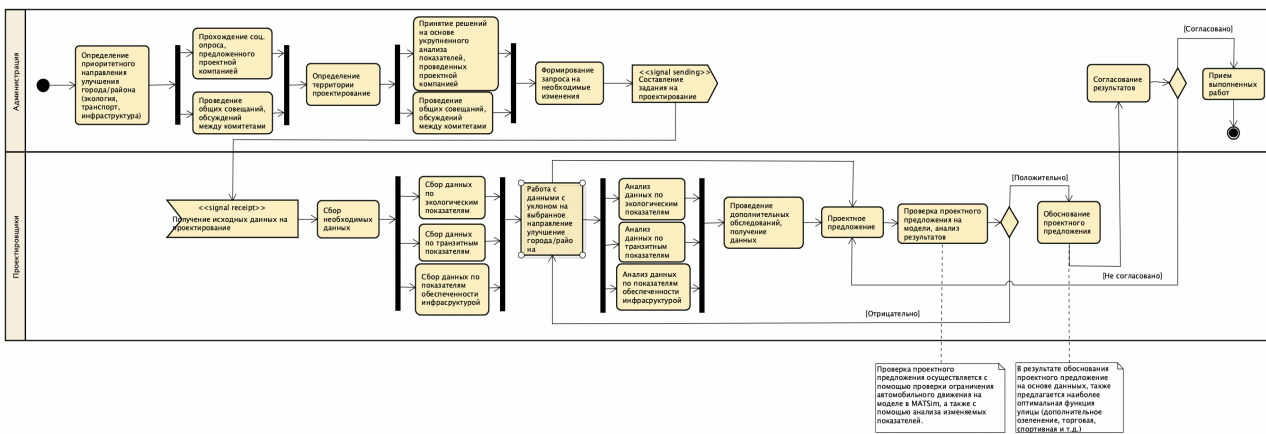


Рис. 2. Диаграмма активностей разработанной системы

На первых шагах производится работа с администрацией, выявляются основные запросы по направлениям улучшения города/района с применением соц. опроса или анализа имеющихся укрупненных показателей и обозначаются варианты улиц с уже имеющимся запросом на ограничение движения (если есть) или определяется территория, требующая развития.

В процессе работы с администрацией формируется первоначальный запрос, который будет уточняться в процессе выполнения следующих этапов.

Далее происходит первичный сбор данных по предложенным критериям. Он осуществляется преимущественно четырьмя способами: с помощью различных существующих сервисов и служб, с использованием транспортной модели города, расчет по опубликованным методикам или натурным исследованиям.

Для того, чтобы сделать вывод о возможности и целесообразности ограничения движения автомобилей на улице необходимо оценить все собранные показатели по определенным критериям.

Для оценки каждого показателя предлагается составить шкалу значений, затем ввести интегральные оценочные характеристики в виде баллов по шкале от 0 до 10 для сравнения и сопоставления показателей на основе действующих СП, СанПиН, ГОСТ.

Все интегральные оценки показателей условно можно разделить на две группы (рис. 3) (важность, необходимость и возможность ограничения транзитного автомобильного движения и потенциал работы улицы как пешеходной), баллы показателей складываются. Показатель по интенсивности автомобильного движения не включается в расчёт, а показывает необходимость работы с транспортной моделью.

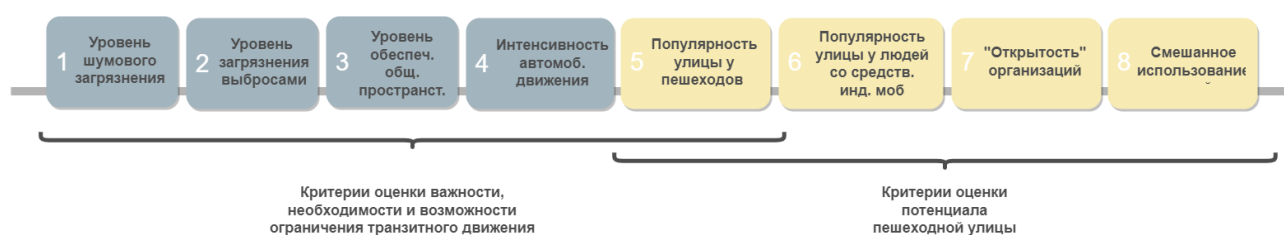


Рис. 3. Распределение показателей по обоснованию ограничения движения и потенциала общественного пространства

Для универсализации разработанного метода и учета характеристик и особенностей исследуемой территории, предлагается ввести весовой коэффициент. В зависимости от выявленных приоритетов развития собранные показатели предлагается умножить на интегральный коэффициент веса в зависимости от установленного приоритета.

После того, как будет сделан вывод в сторону ограничения транзитного автомобильного движения на улице, необходимо перейти к дополнительным обследованиям - проверке ограничения движения на транспортной модели, с целью не допустить ухудшения существующей транспортной ситуации (то есть необходимо проверить не образуются ли заторы, не нарушается ли связность, обеспечиваются подъезды, а также среднее время в пути не изменяется или изменяется незначительно).

На последних этапах формируется проектное предложение о наиболее оптимальных сценариях использования пешеходной улицы.

Разработанный метод возможно применять как при уже известных потенциальных пешеходных улицах, для выбора оптимальной, так и для исследования улиц на ограниченной территории и оценки их возможности и необходимости пешеходизации, а также потенциала улицы как общественного пространства.

Метод был апробирован на Центральном районе Санкт-Петербурга. После сбора и обработки данных по выделенным критериям и их интегральной оценки получили распределение баллов по линкам (рис. 4). Наибольшее количество баллов набрали улицы с наиболее интенсивным автомобильным движением, однако, также высокие баллы наблюдаются и улицы с низким и средним транспортным значением. Например, Большая Конюшенная ул., наб. реки Мойки, наб. Канала Грибоедова, Пушкинская ул., ул. Марата и др.

Далее проверили несколько различных по масштабу сценариев с ограничением автомобильного движения, а именно несколько улиц в туристическом центре города, одну улицу и часть улицы (участок от перекрестка до перекрестка).

Наибольший интерес представляет сценарий с ограничением транзитного автомобильного движения сразу на нескольких улиц. Среднее время в пути относительно базового сценария изменилось на 14 секунд, что может считаться в пределах погрешности.

Потоки перераспределились на ЗСД, Благовещенский и Ушаковский мост и Октябрьскую набережную (рис. 5), а коэффициент загрузки изменился более чем на 0,2 как в сторону уменьшения, так и в сторону увеличения (рис. 6).

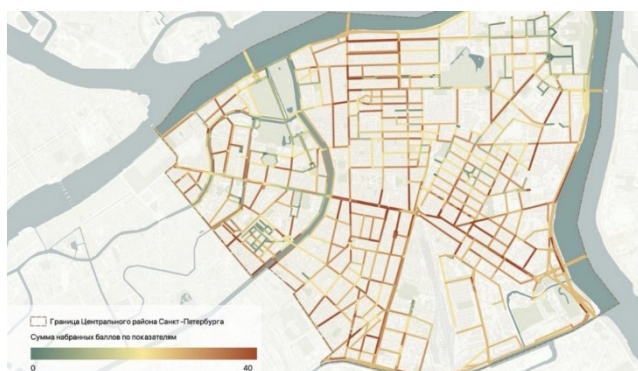


Рис. 4. Распределение баллов по линиям для Центрального района Санкт-Петербурга



Рис. 5. Перераспределение потоков в вечерний час пик

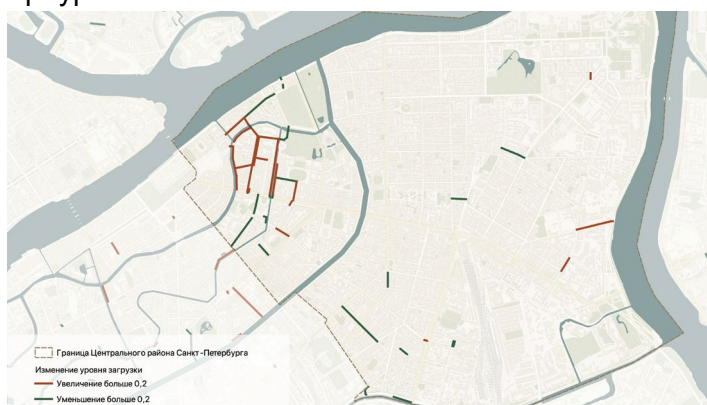


Рис. 6. Изменение уровня загрузки в Центральном районе Санкт-Петербурга

В таком случае стоит взвешенно принимать решение руководствуясь сравнением пользы нового общественного пространства и улучшением экологической составляющей в одном месте и ухудшением экологической составляющей и увеличением заторов в другом.

Выработанный и апробированный на Центральном районе Санкт-Петербурга метод может стать инструментом, помогающим городским планировщикам, специалистам по транспорту и администрации городов и районов принимать решения по пешеходизации улиц как на временной, так и на сезонной или постоянной основе. Разработанный метод может быть использован не только для анализа наиболее востребованных и возможных улиц для пешеходизации, но также для принятий решений о сокращении ширины проезжей части или парковочного пространства и расширения тротуаров.

Пешеходизация улиц в городах – это один из способов достижения устойчивого развития городов, рационального распределения ограниченных ресурсов. Пешеходизация системы улиц поможет в решении задач по нехватке открытых общественных пространств и решению локальных экологических проблем. Пешеходные улицы города выступают драйверами экономического развития, стимулируя рост доходов бизнеса, повышая конкуренцию и стоимость жилья. Грамотно продуманные пешеходные улицы обеспечивают присутствие людей на улице в течение всего дня, что оказывает положительное влияние на уровень безопасности в городе.

Литература

1. Shuhana Shamsuddin N.R.A.H. ASEAN Conference on Environment-Behaviour Studies // Walkable Environment in Increasing the Liveability of a City. Bangkok. – 2012. – №50. – С. 167–178.
2. Создание комфортной среды — это выгода или убыток для городского бюджета. [Электронный ресурс]. – Режим доступа: <https://urtmag.ru> (дата обращения: 01.06.2021).
3. Kieron J. Doick A.P.T.R.H. The role of one large greenspace in mitigating London's nocturnal urban heat island // Science of The Total Environment. – No.493. – 2014. – Pp. 662-671.

УДК 004.584

ИССЛЕДОВАНИЕ И РАЗРАБОТКА ПОДХОДОВ К СОЗДАНИЮ ЧАТ-БОТА НА ПРИМЕРЕ ИСУ ИТМО

Курбатов М.С.¹ (студент)

Научный руководитель – канд. техн. наук, доцент Зудилова Т.В.¹

¹Университет ИТМО

e-mail: maksimkurb@gmail.com, zudilova@ifmo.spb.ru

В данной работе исследуются возможности использования внутренней информационной системы управления (ИСУ ИТМО) и ставятся задачи разработки чат-бота, который будет помогать с навигацией по системе. Также в работе исследуются типы чат-ботов, способы их создания и оценки качества ответов на вопросы при их использовании.

Ключевые слова: чат-бот, средства естественного общения, распознавание естественного языка, разработка, искусственный интеллект.

Бизнес использует чат-боты для автоматизации внутренних процессов и для разгрузки первой линии поддержки [1]. Такие чат-боты отвечают на самые частые вопросы пользователя и маршрутизируют нестандартные обращения к конкретному сотруднику поддержки, имеющему необходимые компетенции для решения таких вопросов.

По мере развития технологий распознавания и генерации письменной речи, общение людей с чат-ботами будет всё меньше отличаться от общения с живым человеком. Всё больше сценариев взаимодействия клиента с сервисом можно будет сделать доступным с помощью анализа текстовых запросов. Вместо того, чтобы долго искать кнопку на сайте, пользователь может выразить своё желание написав боту и получить полную информацию об услуге и возможность тут же воспользоваться ею.

В данной работе рассматриваются подходы к реализации существующих чат-ботов, а также как чат-бот для ИСУ ИТМО может упростить взаимодействие пользователей с этой системой.

Подходы к технической реализации чат-ботов

С точки зрения технической реализации, чат-боты можно разделить на три типа [2]:

- чат-боты, основанные на бизнес-правилах: такие чат-боты работают на основе заранее заложенного дерева сценариев. Диалог с пользователем идёт по строгому пути, который заложил разработчик в момент реализации бизнес-правил. Такие боты обычно не содержат вопросов, предполагающих ответ в свободной форме, также эти боты активно используют заранее заготовленные кнопки для ответа (т.н. чат-бот клавиатуры). Для определения ветви диалога используются алгоритмы, которые рассматривают каждый вопрос в виде «мешка слов» и находят схожую ветвь диалога среди заложенных по количеству совпадающих слов в вопросе пользователя и в триггере ветви диалога;
- чат-боты, основанные на искусственном интеллекте: такие боты не имеют заранее заложенного пути диалога и опираются на модель искусственного интеллекта, которая заранее обучена на тренировочных данных. В идеале, такие чат-боты должны имитировать переписку с живым человеком, однако, для создания качественного бота, полностью основанного на искусственном интеллекте, требуются большие наборы тренировочных данных. При этом, чат-боты,

использующие искусственный интеллект, способны выдавать более персонализированные ответы пользователям;

- гибридные чат-боты: такие чат-боты включают в себя всё лучшее из двух предыдущих подходов. Такие чат-боты имеют заложенные бизнес-правила для ведения диалога по определённому сценарию, но применяют искусственный интеллект для обработки фраз пользователя, вычленения из них намерений и полезной нагрузки (номера договоров, адреса и тому подобное). Данный тип чат-ботов является наиболее распространённым в коммерческих приложениях.

Гибридный тип чат-ботов будет наиболее предпочтительным для новых решений, поскольку позволяет добиться наилучшего понимания намерения пользователя, а также решения его вопроса с помощью заранее заложенных бизнес-правил.

Способы создания чат-бота для ИСУ ИТМО

Для создания чат-ботов, которые не требуют реализации сложной логики или интеграции с внешними системами (кроме поддерживаемых), существуют специальные конструкторы чат-ботов. Примеры таких систем: SendPulse, Recast.ai, 13chats, HubSpot [3]. Функциональность чат-ботов, разработанных с использованием конструкторов, ограничивается возможностями платформы. Такие конструкторы могут включать в себя функции обработки естественного языка, интеграцию с внешними системами управления взаимоотношениями с клиентами (CRM), аналитику, отправку рассылок пользователям чат-бота и административную панель для управления чат-ботом.

Плюсы такого подхода: для создания чат-бота в конструкторе не требуется навыков программирования, скорость разработки и стоимость поддержки чат-бота становится ниже.

Среди минусов стоит выделить то, что платформа может не поддерживать интеграцию с интересующей системой, а разработчики конструктора могут долго внедрять новые необходимые функции.

Альтернативой такому подходу является написание чат-бота вручную с использованием средств разработки приложений. Созданные таким образом чат-боты могут обрабатывать абсолютно любые запросы и быть интегрированы с любыми внешними системами. Однако, у таких чат-ботов есть и минусы: они могут быть дорогими в разработке, а сама разработка может занимать много времени. Помимо этого, необходимо поддерживать хостинг для размещения чат-бота.

Поскольку ИСУ ИТМО является закрытой системой, создание чат-бота с использованием конструктора становится менее предпочтительным, так как даже если у конструктора будет возможность интеграции со своими системами, доработка ИСУ ИТМО всё равно потребует навыков программирования. В таком случае лучше потратить время на написание чат-бота вручную, чтобы впоследствии иметь возможность неограниченного расширения его функциональности.

Поддержка и корректирование чат-бота

В процессе разработки чат-бота трудно предугадать все возможные сценарии диалогов бота с пользователями. Для того, чтобы чат-бот не тратил время пользователя, а чётко отвечал на поставленные вопросы и решал проблемы человека, необходимо постоянно оценивать результаты диалогов с клиентами (в нашем случае – со студентами и преподавателями университета).

Основным критерием оценки эффективности чат-бота является процент общего количества диалогов, которые бот обработал сам, без помощи оператора [4]. В случае с чат-ботом для ИСУ ИТМО, в роли оператора (при условии того, что работает с чат-ботом студент, а не преподаватель) может выступать студенческий офис, который может ответить на различные вопросы студентов, а также выдать необходимые справки.

Количество вернувшихся пользователей (то есть тех, кто использовал чат-бот повторно) позволяет оценить, насколько эти пользователи были удовлетворены ответом чат-бота ранее и готовы выбрать новый канал коммуникации [5].

Индекс потребительской активности (NPS) — это оценка лояльности пользователей по отношению к продукту и компании. Данную метрику очень удобно рассчитывать при

использовании формата чат-бота: в конце диалога бот может задавать вопрос «Остались ли вы довольны ответом?» [5]. В качестве негативных вариантов ответа бот должен предлагать следующие: «Я спрашивал про другое», «Ответ не решил проблему», «Я не понял ответ». Этот ответ поможет понять причину низкой оценки при анализе истории диалогов.

Определение наиболее часто используемых и трудных разделов ИСУ ИТМО

ИСУ ИТМО имеет два основных раздела: Интернет-портал и Личный кабинет. Интернет-портал содержит разделы, касающиеся ресурсов и инструментов всего университета (например, структура университета, научные статьи, документооборот и тому подобное), тогда как Личный кабинет содержит информацию о конкретном человеке внутри университета: его расписание, электронная зачётная книжка (для студентов), личные публикации, заказ документов, финансы, доступ к VPN и другое.

Часть функциональности вынесена в отдельный веб-портал «ИСУ Центр приложений» (см. рисунок). Портал позволяет проводить анкетирование, бронировать помещения. Также существует отдельный портал «my.itmo», в котором дублируются самые важные разделы ИСУ: расписание студента, размер стипендии, электронная зачётная книжка, выбор дисциплин для магистрантов и др.

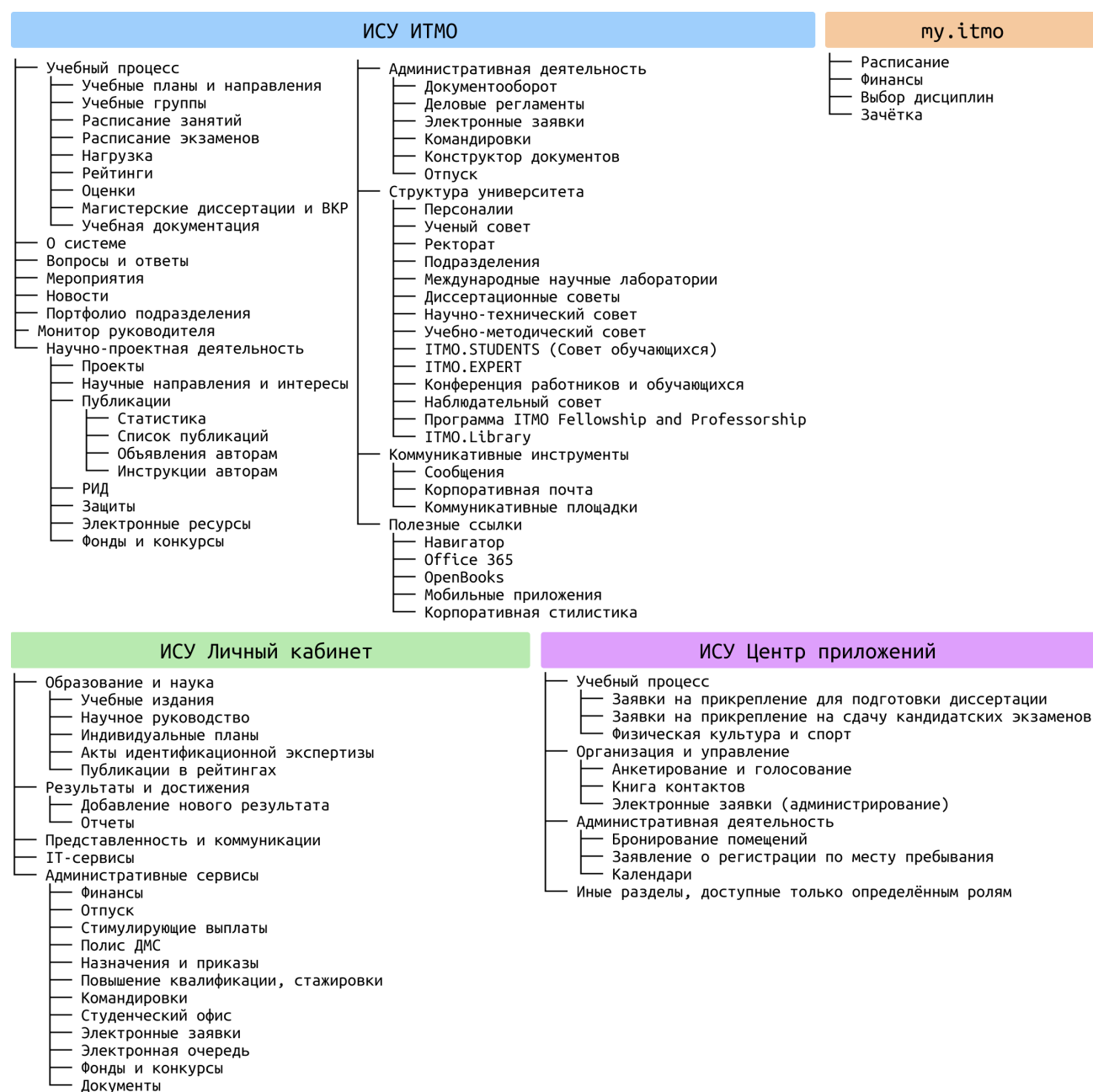


Рисунок. Схема разделов ИСУ ИТМО

Самые частые сценарии использования ИСУ:

- просмотр своего расписания;
- просмотр электронной зачётной книжки;
- получение пароля wi-fi;
- персоналии (поиск студентов и сотрудников);
- расписание по группам, преподавателям, дисциплинам, кабинетам;
- подача заявок в студенческий офис;
- изучение информации о выборных дисциплинах.

ИСУ содержит множество разделов и подразделов, о части которых некоторые студенты или преподаватели могут не догадываться годами. Так, например, многие студенты не знают о существовании приложения-навигатора по корпусам ИТМО, а также о том, что они могут связаться с любым человеком прямо через мессенджер ИСУ или получить бесплатную лицензию Microsoft Office на время обучения. Некоторые возможности ИСУ спрятаны в неочевидном месте. Так, например, студенческие объединения (студклубы) находятся в разделе «ИТМО.STUDENTS» и не могут быть найдены с помощью поиска в ИСУ: пользователь должен как-то узнать, в каком разделе ему следует искать необходимую информацию.

Данные проблемы можно решить с помощью создания чат-бота, который позволит реализовать часто используемые сценарии с помощью мессенджеров и получить помощь по разделам, с которыми у пользователя могут возникнуть проблемы.

Составление функциональных требований к чат-боту и плана по его разработке и внедрению

Чат-бот для информационной системы ИТМО должен предоставлять возможность пользоваться самыми популярными разделами ИСУ, а именно узнавать расписание, находить контакты студентов или сотрудников, изучать информацию по той или иной дисциплине, или помочь найти нужный кабинет.

Навигатор по корпусам университета ИТМО может быть встроен непосредственно в чат-бот, для этого необходимо создать карты корпусов. В качестве результата бот может выдавать готовое изображение с нарисованным путём.

Также, ввиду сложности навигации в интерфейсе ИСУ, чат-бот должен уметь отвечать на вопросы пользователя как найти ту или иную информацию, давая прямую ссылку на необходимый раздел в системе. Для этого необходимо составить карту навигации по ИСУ

Для информирования студентов и сотрудников о существовании различных разделов ИСУ, чат-бот может рассылать информативные сообщения вида «А вы знали, что ...» после ответа на запрос пользователя.

Для получения информации о том, какими функциями чат-бота пользуются чаще всего и какого функционала не хватает, в чат-бот может быть встроена система сбора статистики и учёта неопознанных запросов пользователей. На первых порах такая система позволит понять разработчику, что пользователи ожидают от бота и какие функции стоит дорабатывать в первую очередь.

Для интеграции чат-бота с данными из ИСУ необходимо получить доступы к базе данных ИСУ, проанализировать структуру базы данных и подготовить SQL-запросы. Размещать бот можно на мощностях университета ИТМО.

Чат-бот должен быть реализован в виде модульного приложения для возможности расширения функциональности.

Просуммировав всё вышесказанное, получится следующий список задач, решаемых в процессе разработки и внедрения чат-бота для ИСУ ИТМО:

- подготовка технического задания на разработку чат-бота;
- сравнительный анализ и выбор подходящей платформы для создания гибридного чат-бота;
- изучение структуры базы данных ИСУ ИТМО;
- создание справочников для навигации по ИСУ;
- реализация модуля распознавания намерения пользователя (обработка текстовых запросов);
- создание модуля аутентификации;

- создание модуля сбора статистики по использованию бота;
- создание модулей, обрабатывающих запросы пользователя;
- написание модульных и интеграционных тестов;
- запуск чат-бота на серверах университета ИТМО;
- поддержка чат-бота.

Чат-бот является перспективным средством для коммуникации клиентов с компанией. Этот инструмент более привычен пользователю, потому что воспринимается как общение с живым человеком. Также чат-боты помогают компаниям автоматизировать процессы и экономить человеческие ресурсы.

Чат-бот для ИСУ ИТМО позволит студентам и преподавателям взаимодействовать с внутренними ресурсами университета, не выходя за рамки привычных средств коммуникации, чем упростит взаимодействие пользователей с этой системой.

Гибридный вариант чат-бота, использующего искусственный интеллект и бизнес-правила для решения задач, предпочтителен для ИСУ ИТМО, потому что такой чат-бот сможет решать конкретные проблемы пользователя, правильно понимая его запросы.

Литература

1. Веретено А.А., Кушков Е.А. Мессенджеры и чат-боты-незаменимые помощники в формировании бренда // Евразийский Союз Ученых (ЕСУ). – 2019. – №5(62). – С. 39.
2. Ураев Д.А. Классификация и методы создания чат-бот приложений // International scientific review. – 2019. №LXIV. – С. 30-33.
3. Самописный чат-бот vs конструктор: как выбрать решение под свой проект? [Электронный ресурс]. – Режим доступа: <https://vc.ru/services/201315-samopisnyy-chat-bot-vs-konstruktor-kak-vybrat-reshenie-pod-svoyu-proekt> (дата обращения: 29.12.2021).
4. Почему чат-бота надо проверять и корректировать. [Электронный ресурс]. – Режим доступа: <https://www.itweek.ru/idea/article/detail.php?ID=221732> (дата обращения: 29.12.2021).
5. Ураев Д.А. Метрики для оценки качества чат-бот приложений // Наука, техника и образование. – 2019. – №9(62). – С. 36–40.

УДК 004.4

ПРИМЕНЕНИЕ БЛОКЧЕЙНА НА ГОСУДАРСТВЕННОМ УРОВНЕ И АНАЛИЗ ЕГО УСТОЙЧИВОСТИ

Матвеев Г.А.¹ (студент)

Научный руководитель – канд. техн. наук Войтюк Т.Е.¹

¹*Университет ИТМО*

e-mail: 310026@niuitmo.ru, voitiukt@itmo.ru

В статье рассматриваются варианты применения блокчейна в государственном секторе. Приводится обзор технологии блокчейн, рассматриваются цели и сферы ее применения в различных странах, определяются выгоды от использования блокчейна для государственного сектора. В заключении приводится анализ устойчивости технологии блокчейн.

Ключевые слова: блокчейн, государственный сектор, устойчивость технологии, безопасность данных, защищённость транзакций.

Блокчейн – это одноранговая сеть, которая формирует распределенную базу данных, поддерживающую постоянно растущий список записей. Эти упорядоченные записи называются блоками, которые содержат транзакции. Транзакция может содержать различные типы данных. Как правило, все блокчейны работают над тем, чтобы децентрализованные узлы достигли соглашения об общем порядке транзакций. Соглашение о порядке транзакции осуществляется с помощью криптографии и базового механизма консенсуса [1]. Каждый блок в цепочке содержит временную метку и хэш-указатель, который ссылается на предыдущий блок [2]. В работе Клавина и др. [1] описываются эти механизмы, использующие односторонние хэш-функции, генерирующие уникальные буквенно-цифровые текстовые результаты при вводе списка транзакций. Если бы кто-то изменил что-то в списке транзакций, результирующий хэш был бы значительно другим, и цепочка не достигла бы консенсуса.

Именно одноранговая сеть существенно отличает базы данных типа блокчейна от традиционных централизованных баз данных, где есть только одно центральное местоположение, известное как узел (компьютер или кластер компьютеров) для хранения данных и операций. Централизованные базы данных обычно реплицируются на несколько узлов для обеспечения механизмов резервного копирования и отказоустойчивости, но они функционируют синхронно. В технологии блокчейн каждый узел в одноранговой сети всегда обладает полной копией всего (распределенный реестр), и все изменения одновременно отправляются на каждый узел. Каждый узел запускает алгоритм консенсуса, который обеспечивает способ достижения соглашения с каждым другим узлом рассматриваемой транзакции.

Зачастую государственный сектор отстает в технологиях по сравнению с частным. Технология блокчейн появилась относительно недавно, и, казалось бы, она пока не должна широко использоваться в государственном секторе. Однако, в огромном количестве стран блокчейн используется чрезвычайно активно уже несколько лет.

В данной работе рассмотрены лишь некоторые страны, интегрирующие технологию блокчейн в свои государственные системы, а также описаны общие области использования этой технологии по всему миру.

В конце 2014 года пять стран объединились для обмена опытом в области информационных технологий и назвали себя D5 – Великобритания, Южная Корея, Эстония, Новая Зеландия и Израиль. За пять лет, к концу 2019 года, в этой организации уже было девять стран и множество

созданных совместных проектов, но прежде всего интересен опыт этих пяти стран и их успехи в развитии блокчейна на момент появления его в государственном секторе (к концу 2016 года).

Технология блокчейн привлекательна тем, что она хранит данные о транзакциях на нескольких устройствах, поэтому обладает очень высокой степенью надежности данных и невозможностью их фальсификации. Поэтому, в первую очередь, страны хотели внедрить блокчейн в системы, где необходима безопасность – для искоренения коррупции и фальсификации данных.

Наиболее интересные решения стран D5 в государственной сфере следующие [3]:

- Эстония в свое время совершила прорыв в области электронной регистрации граждан, реализовав возможность использования государственных услуг резидентами страны, даже при нахождении за ее пределами. Разумеется, речь идет об услугах, которые были реализованы в интернете. Но это не главное новшество этой системы, самым интересным фактором было то, что электронным резидентом страны мог стать практически любой желающий – даже бизнесмены из разных стран могли подать заявку на получение электронного вида на жительство и получить права на ведение бизнеса в Европе по правилам местного гражданина. Даже не задумываясь о деталях системы, понятно, насколько важно защитить данные от подделки, и именно для этой цели был использован блокчейн, позволяющий хранить огромное количество информации на разных ресурсах. Система получила название *Bitnation*. Она очень быстро распространилась, ее начали интегрировать во многие проекты, в частности для идентификации пользователей.
- Южная Корея использовала технологию блокчейн для создания систем голосования и в финансовом секторе.
- В Новой Зеландии государственный сектор занимался "зелеными" технологиями – были созданы системы для установления четкого контроля над производством продуктов питания. Все данные обрабатывались блокчейном: где продукт был произведен, когда и кем. Это очень помогает в контроле оборота, что позволяет свести к минимуму количество контрабандных, контрафактных и некачественных товаров.
- Израиль, в свою очередь, начал внедрять систему, способную контролировать действия с криптовалютой во всех финансовых сферах.
- Великобритания, как и Израиль, в первую очередь решила использовать технологию блокчейн в сфере финансов, в областях, финансируемых государством: гранты, пособия, пенсии, кредиты.

Кроме представленных первых примеров использования блокчейна в государственном секторе стран D5, в данной отрасли развивались самостоятельно – США, Китай, многие европейские страны и другие.

Может создаться впечатление, что блокчейн используется в основном для контроля в банковской и финансовой сферах. Действительно, эта технология появилась в государственном секторе именно в этих областях. Однако, в настоящее время она получила чрезвычайно широкое распространение и используется везде, где требуется тщательный контроль.

Одним из основных человеческих ресурсов является здоровье, поэтому блокчейн также популярен в медицинской сфере. Например, в Соединенных Штатах создали систему, с помощью которой контролируют распространение гепатита. В современных реалиях система помогает в борьбе с COVID-19.

Важно понимать, что блокчейн – это не набор большого количества технологий, способных решить все задачи и проблемы. Это надежный и хорошо документированный базовый аспект многих систем. Блокчейн активно используется совместно с машинным обучением, анализом данных и подобными технологиями, которые могут улучшить результаты работы. Аналогичная система в США, основанная на искусственном интеллекте и блокчейне, смогла модернизировать систему хранения информации департамента здравоохранения, что, по прогнозам, позволит сэкономить сотни миллионов долларов.

Помимо активного развития медицинской системы по всему миру, блокчейн активно используется в производственных сферах, где он вместе с другими технологиями

(чаще всего с интернетом вещей) позволяет снизить процент системных сбоев, а также повышает отказоустойчивость систем.

Как для правительств, так и для частных предприятий преимущества использования блокчейна очевидны:

- открытость информации (там, где это необходимо);
- безопасность системы;
- подробное протоколирование всех действий.

Данные преимущества технологии позволяют в государственном секторе:

- сделать все услуги более согласованными. Например, разрешить гражданам регистрироваться в разных системах под одной учетной записью;
- уменьшить уровень коррупции;
- обеспечить детальное хранение данных о жителях страны, финансовых операциях, контрактах.

Наличие множества узлов, работающих по всему миру, повышает безопасность системы, однако, приводит к высокому потреблению энергии.

Основные источники потребления энергии в технологии блокчейн можно разделить на:

1. Майнинг – добавление блоков в цепочку.
2. Избыточность – наличие всех данных, хранящихся на многих компьютерах.

В то время как избыточность считается небольшой проблемой, майнинг является реальным источником, подлежащим анализу, поскольку является основой транзакций [4].

Потребление энергии, вызванное блокчейном, имеет значительный углеродный след и может участвовать в глобальном потеплении. В статье [4] де Врис рассматривает биткойны, однако, энергопотребление может быть приблизительно рассчитано для других приложений блокчейна с тем же рабочим механизмом. В этой статье количество энергии, потраченной впустую Биткоином, сравнивается с количеством небольших отходов ИТ и телекоммуникационного оборудования, производимых такой страной, как Нидерланды. Подчеркивается тот факт, что биткойны все еще малочисленны и будут расти все больше, вызывая большее потребление энергии.

В статье Седлмейра описан прямой анализ энергопотребления, вызванного блокчейном [5]. В ней определяются нижняя и верхняя границы для источника, требующего энергию, связанную с майнингом, и благодаря некоторому предположению (об аппаратном обеспечении всех узлов блокчейна и экономическом состоянии), рассчитывается реалистичная оценка. Основной проанализированной реализацией блокчейна является разработка, основанная на технологии доказательства работы, которая является самой энергозатратной архитектурой. Она спроектирована таким образом, что высокое энергопотребление является страховкой в целях безопасности, поскольку для того, чтобы взломать её, требуется контроль 25 – 50% узлов всей цепочки, что приводит к 25 – 50% потребляемой мощности. Таким образом, чем выше потребность в энергии, тем большее количество энергии должен потреблять злоумышленник, чтобы взломать её. Однако, несмотря на высокое энергопотребление не представляется большая угроза для климата, так как потребление энергии блокчейнами существенно не увеличивается, когда они обрабатывают больше транзакций.

Блокчейн распространяется, и, вероятно, станет основой многих будущих реализаций. Страны по всему миру действительно могут заинтересоваться внедрением какого-либо приложения на основе этой технологии, следуя примеру Эстонии, где произведена попытка воспользоваться всем, что может предложить блокчейн. Его высокая безопасность может быть очень интересна, когда речь заходит об избирательной системе или обработке высокочувствительных данных. Кроме того, люди могли бы быть более уверенными в использовании технологий, основанных на блокчейне, поскольку они могли бы напрямую взаимодействовать без посредников, обрабатывающих их данные, в результате чего приложение или система, которые действительно могли бы использоваться гражданами, сохраняли бы безопасность.

Литература

1. Clavin J., Duan S., Zhang H., Janeja V.P., Joshi K.P., Yesha Y., Erickson L.C., Li J.D. Blockchains for Government: Use Cases and Challenges // Digit. Gov.: Res. Pract. – 2020. – Vol.1. – No.3, Article. – №22. – Pp. 1–21.
2. Ølnes S., Jansen A. Blockchain Technology as Infrastructure in Public Sector – an Analytical Framework // Proceedings of the 19th Annual International Conference on Digital Government Research: Governance in the Data Age. Article. – No. 77.– 2018. – Pp. 1–10.
3. Ojo A., Adebayo S. Blockchain as a Next Generation Government Information Infrastructure: A Review of Initiatives in D5 Countries // Government 3.0 – Next Generation Government Technology Infrastructure and Services. – 2017. – Pp. 283–298.
4. A. de Vries, C. Stoll. Bitcoin’s Growing e-waste Problem // Resources, conservation and recycling 175. – 2021. – Pp. 64–73.
5. Sedlmeir J., Buhl H., Fridgen G., Keller R. The Energy Consumption of Blockchain Technology: Beyond Myth // Business information systems engineering 62.6. – 2020. – Pp. 599–608.

УДК 004.891.2

РАЗРАБОТКА КЛАССИФИКАТОРА ОБЛАЧНЫХ УСЛУГ ДЛЯ СИСТЕМЫ УПРАВЛЕНИЯ ЗАТРАТАМИ И ПЛАНИРОВАНИЯ ОБЛАЧНОЙ ИТ-ИНФРАСТРУКТУРЫ ПРЕДПРИЯТИЙ

Меркушев А.Е.¹ (аспирант)

Научный руководитель – д-р экон. наук, профессор Максимова Т.Г.¹

¹Университет ИТМО

e-mail: alex.e.merkushev@gmail.com

Одновременно с наращиванием объемов использования облачных ресурсов, растет объем и сложность отчетности об использовании. Ведется работа по созданию модели рекомендаций для управления облачными ресурсами и затратами на них. В рамках данного доклада освещен текущий прогресс в работе и применение имеющейся модели классификации облачных сервисов. Был разработан классификатор облачных сервисов, основанный на имеющейся теоретической базе в целях возможности однозначной классификации каждого подтипа облачной услуги и дальнейшего формирования кост-модели, представления данных в разных разрезах в зависимости от того, что необходимо конкретному пользователю.

Ключевые слова: облачные инфраструктуры, классификатор, кост-модель, облачные сервисы, облачные ресурсы.

Рынок облачных технологий показывает небывалый рост в течении последних 10 лет, что подтверждается исследованиями ведущих мировых консалтинговых и аналитических компаний, таких как ежегодный отчет Gartner [1]. Трендов на приостановку роста не наблюдается, а только наоборот, прогнозируется кратное увеличение спроса на Облачные услуги в ближайшие 3 года. Современные реалии жизни во время и после пандемии COVID-19 бросили новые вызовы пользователям корпоративных ресурсов – использование частного облака и нагрузка на него выросли кратно, что сподвигло бизнес-пользователей перейти в сторону гибридных облачных инфраструктур и использования частного облака наряду с публичными облачными сервисами. Как результат, менеджменту сложнее понимать комплексные отчеты, еще сложнее выделить из них ценность, чтобы работа облака была максимально эффективной.

Для этого необходимо начать с однозначной классификации ресурсов, которая бы работала для всех провайдеров и давала бы срез в разных интересующих бизнес плоскостях.

Каждый из провайдеров и поставщиков сервисов использует свой формат выгрузки, растет разнообразие сервисов – у мирового лидера отрасли Amazon Web Services на данный момент более 190 сервисов, разобраться в которых становится все сложнее. Также растут объемы обрабатываемой информации, что делает эту задачу еще более трудновыполнимой. Традиционные средства в виде таблиц требуют сложной предобработки для представления всех источников в единообразном виде, а имеющиеся и предлагаемые поставщиками продукты привязаны к одному провайдеру, и аналитика всего спектра предложений не обладает исчерпывающей функциональностью. Как результат, все еще большую роль в задачах аналитики и оптимизации выполняет человек, цена ошибки которого может исчерпываться тысячами долларов при ошибке или невнимательности в управлении облачными затратами в крупной организации. Полезным видится разработка инструмента, основывающегося на своей кросспровайдерной методологии, который помог бы менеджерскому составу без углубленной

аналитики отчетности предотвращать финансовые потери и увеличивать эффективность облачных инфраструктур на основании заложенных сценариев.

В связи с быстрым темпом развития области, данная ее часть была исследована в основном корпоративными институтами, такими как Arptio. В некоммерческой сфере лидирует некоммерческая организация Совет по управлению технологическими процессами (TBM Council) [2]. Крупнейшей наработкой является их труд по созданию сборника таксономии в сфере управления технологическими бизнес-процессами – TBM Taxonomy, а совместно с Arptio была создана унифицированная модель ATUM (Arptio TBM Unified Model) [3], которая позволяет классифицировать любой ресурс в IT-инфраструктуре компании (рис. 1).

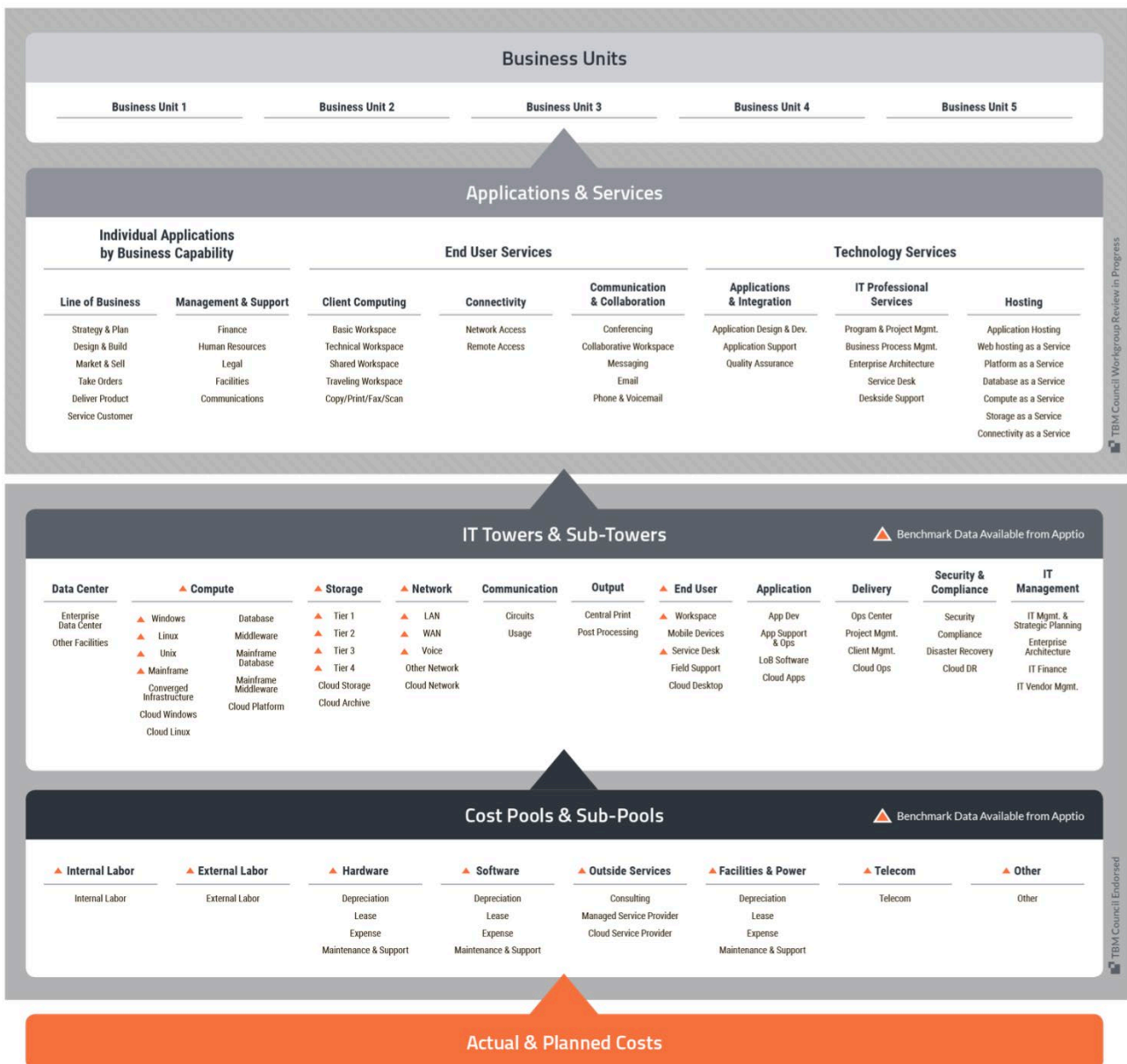


Рис. 1. Классификатор ATUM

Академическая разработанность темы управления IT активами подтверждается в основном книгой «Измерение ITSM» [4]. Благодаря этому, можно оперировать основной таксономией области и теоретической классификацией ресурсов, сервисов и активов внутри инфраструктуры. Стоит отметить, что имеющиеся материалы не являются исчерпывающими, а лишь разбирают общие теоретические наработки либо частные случаи оптимизации [5].

В ходе работы была взята модель ATUM как самая универсальная и всеобъемлющая, но требующая значительной симплификации для упрощения пользования и приведения сервисов разных провайдеров к одному знаменателю (рис. 2).

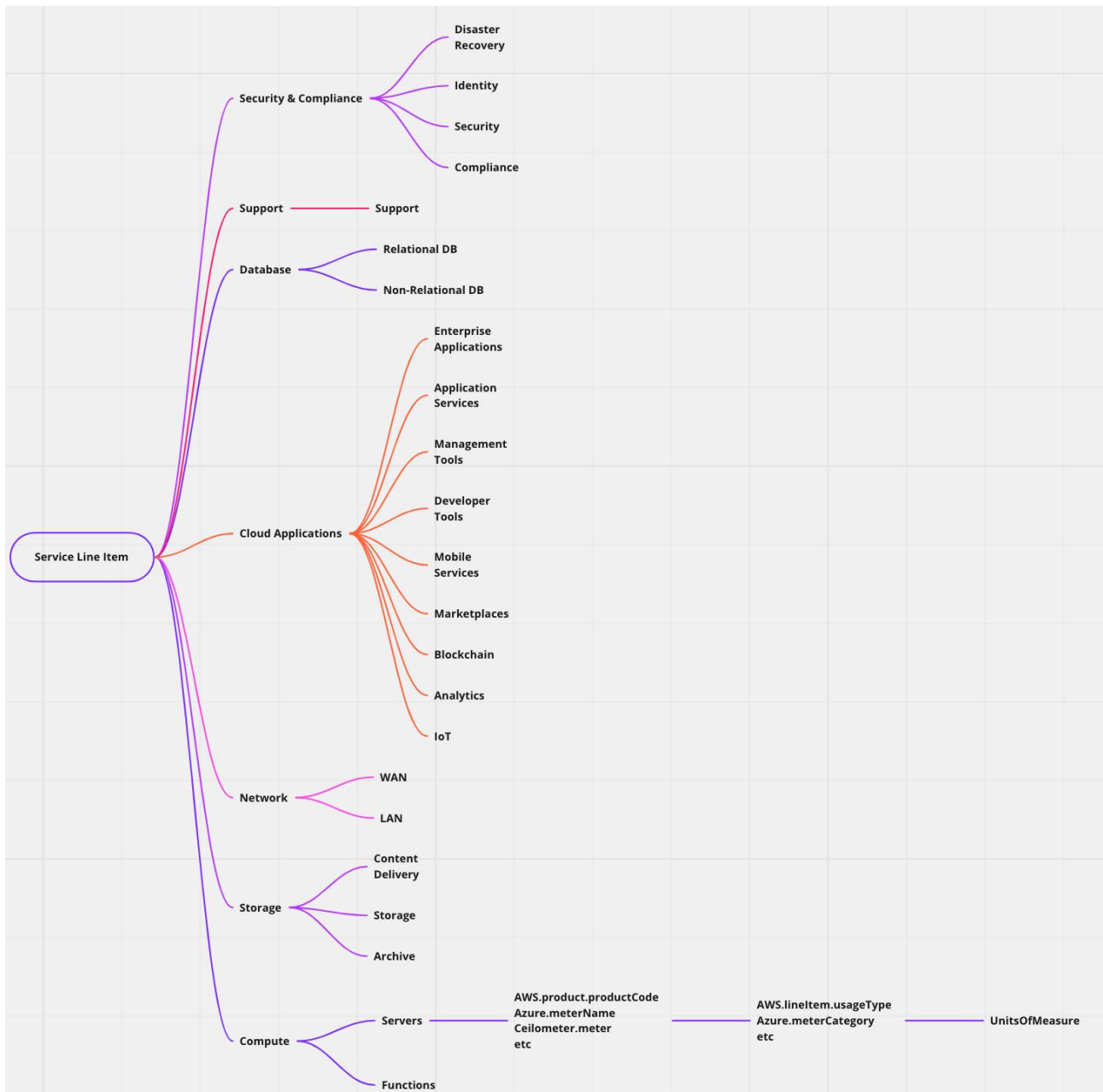


Рис. 2. Разработанный классификатор облачных услуг

Была сохранена иерархическая структура с выделением Столпов, Семейств, Типов Сервисов, Подтипов и Типов Использования Сервисов (IT Tower – Service Family – Service Type – Service Sub-Type – Service Usage Type). Столпы были выделены в рамках применимых значений для управления ИТ инфраструктуров, а далее разбиты на семейства сервисов в различии от их назначения. Крупный пул сервисов занимают SaaS сервисы, выделенные в Cloud Apps и разбивающиеся на семейства по применению. Типом сервиса предложено называть его имя от провайдера (напр. Amazon S3, Azure Monitor, Yandex DataLens, OpenStack Neutron, etc). Каждый из них распадается на подсервисы (у хранилища S3 есть несколько способов хранения данных в зависимости от времени хранения), а далее поставлен тип использования, идентифицирующий за что конкретно была начислена плата (запрос, хранение данных, время использования, обработанные данные и т.д.)

Данный классификатор может однозначно задать положение сервиса в кросспровайдерной системе управления затратах на облачные ресурсы. Поставленные цели достигнуты, задачи выполнены.

Далее можно проводить работу по разработке методологии для оптимизации затрат, так как для каждого из сервисов будет разный подход и разное свойственное поведение. Так,

анализируя поведение каждого сервиса, можно переходить к выведению паттернов поведения или выявлению финансово рискованных конфигураций. На данный момент отсутствует методология и модель классификации процессов оптимизации и предотвращения финансовых потерь в облачных инфраструктурах. Поэтому предлагается создать унифицированную систему поддержки принятия решения для системы управления затратами и планирования облачной IT-инфраструктуры предприятий.

При успешном выполнении работы по разработке модели поддержки принятия решения для системы управления затратами и планирования облачной IT-инфраструктуры предприятий видятся следующие основные значимые выгоды:

- научное сообщество получит методологические материалы для поддержки принятия решений в сфере управления IT-активами компаний;
- бизнес-сообщество получит набор рекомендаций и практик в сфере управления IT-активами компаний и управления технологическим бизнесом;
- практическая значимость работы также заключается в расширении имеющихся систем биллинга и отчетности, чтобы данные системы могли быть оснащены модулями рекомендаций с возможными способами сокращения затрат, грамотного планирования IT-инфраструктуры и предотвращения финансово опасных случаев использования облачных сервисов;
- технический менеджмент получит способ быстрее принимать решения по управлению IT-активами;
- финансовый менеджмент получит рекомендации по сокращению затрат на облачную IT-инфраструктуру;
- разработчики и технические команды получают инструмент по контролю и предотвращению финансово рискованных конфигураций.

Литература

1. Costello K., Hippold S. Gartner Forecasts Worldwide Public Cloud Revenue to Grow 17.3 Percent in 2019. [Электронный ресурс]. – Режим доступа: <https://www.gartner.com/en/newsroom/press-releases/2019-04-02-gartner-forecasts-worldwide-public-cloud-revenue-to-g> (дата обращения: 20.02.2022).
2. TBM-Council T. B. M. Framework. – 2018.
3. Apptio TBM Unified Model 2021. [Электронный ресурс]. – Режим доступа: <https://www.apptio.com/platform/atum/> (дата обращения: 20.02.2022).
4. Steinberg R.A. Measuring ITSM. – Trafford Publishing. – 2013. – 193 p.
5. Mishra S. et al. Pareto-optimal cost optimization for large scale cloud systems using joint allocation of resources // Journal of Ambient Intelligence and Humanized Computing. – 2019. – С. 1–19.

УДК 004.934

СИСТЕМА АВТОМАТИЧЕСКОГО ПОДБОРА ПАРАМЕТРОВ АЛГОРИТМОВ ДИАРИЗАЦИИ

Мещеряков И.Д.¹ (студент)

Научный руководитель – Попов Д.В.²

¹*Университет ИТМО*

²*ООО «ЦРТ-Инновации»*

e-mail: ilya_mesheryakov_1999@mail.ru

В данной работе исследуется диаризация и осуществляется разработка оптимизатора гиперпараметров диаризации. Результатом работы является программный код оптимизатора гиперпараметров. В качестве гиперпараметров диаризации были приняты параметры первичной кластеризации. В качестве результатов внедрения оптимизатора предоставляются табличные результаты его работы, доказывающие улучшения работы диаризации на 13% в результате работы оптимизатора.

Ключевые слова: диаризация, байесовская оптимизация, гауссовские процессы, параллельные вычисления, кеширование.

Гиперпараметрами в машинном обучении принято называть параметры, которые не обучаются (изменяются) в процессе обучения модели. Поскольку на данный момент не существует определенного алгоритма по выбору наилучшей комбинации этих параметров еще до начала обучения модели, эта проблема решается подбором таких коэффициентов, основываясь на результатах работы модели. Также, данная проблема относится и к области AutoML («автоматизация поиска гиперпараметров алгоритмов машинного обучения»).

Рассмотрим несколько алгоритмов оптимизации гиперпараметров. Одни из наиболее знаменитых алгоритмов – это «поиск по сетке» и «случайный поиск». «Поиск по сетке» является наиболее простым методом. Для него задается конечный набор значений для каждого гиперпараметра, а алгоритм оценивает значение ошибки модели при каждом векторе гиперпараметров, входящим в результат декартова произведения наборов значений гиперпараметров. «Поиск по сетке» страдает от «проклятия размерности», поскольку требуемое количество оценок функций растет экспоненциально с увеличением размерности областей значений. Дополнительная проблема «поиска по сетке» заключается в том, что увеличение дискретизации гиперпараметров (т.е. уменьшение шага в конечном наборе значений без изменения граничных точек) существенно увеличивает требуемое количество оценок функций [1].

Простой альтернативой поиску по сетке является «случайный поиск» [2]. Как следует из названия, выборки в данном случае выбираются случайным образом до тех пор, пока не будет исчерпан определенный лимит вычислений для поиска. Это работает лучше, чем «поиск по сетке», когда некоторые гиперпараметры намного важнее других. «Случайный поиск» также является полезным методом для инициализации процесса поиска, поскольку он исследует все пространство конфигурации и, таким образом, часто находит настройки с разумной производительностью. Тем не менее, это не является универсальным решением и алгоритм часто занимает гораздо больше времени, чем методы управляемого поиска, для определения одной из наиболее эффективных конфигураций гиперпараметров.

Байесовская оптимизация – это алгоритм оптимизации дорогостоящих функций «черного ящика», которая недавно получила распространение в оптимизации гиперпараметров

благодаря получению новых результатов в настройке глубоких нейронных сетей, а также демонстрации широкой применимости к различным задачам [3]. Байесовская оптимизация является итеративным алгоритмом с двумя ключевыми компонентами: вероятностной суррогатной моделью и функцией сбора данных, позволяющей решить, какую точку оценивать следующей. На каждой итерации суррогатная модель «подгоняется» под все наблюдения целевой функции, сделанные до сих пор. Затем функция сбора данных, которая использует прогнозирующее распределение вероятностной модели, определяет полезность различных точек-кандидатов, находя баланс между исследованием новых точек и использованием уже имеющихся точек. По сравнению с оценкой дорогостоящей функции «черного ящика», функция сбора данных требует гораздо меньших вычислительных ресурсов, и поэтому может быть тщательно оптимизирована [4].

Рассмотрим несколько суррогатных функций байесовской оптимизации. Традиционно, байесовская оптимизация в качестве суррогатной функции использует гауссовские процессы (GP) [5] для моделирования целевой функции из-за их выразительности, плавных и хорошо откалиброванных оценок неопределенности и вычислимости прогнозного распределения в замкнутой форме. При гауссовских процессах целевая функция аппроксимируется в виде апостериорного распределения значения при выбранных параметрах и известной истории. В качестве априорного распределения используется нормальное распределение с константным матожиданием и изменяемой функцией ковариации.

Суррогатная функция Tree Parzen Estimator (TPE), вместо моделирования вероятности значений при выбранных параметрах, моделирует вероятностное распределение значений и вероятностное распределение гиперпараметров при определенных значениях. Для этого априорные вероятности распределения каждого из гиперпараметров заменяются на их параметрические аналоги. TPE продемонстрировал хорошую производительность в таких структурированных задачах оптимизации гиперпараметров [5], концептуально прост и естественным образом распараллеливается.

Теперь рассмотрим оптимизируемую целевую функцию. Диаризация – это процесс разделения звукового потока на несколько в соответствии с принадлежностью сегментов тому или иному говорящему. Здесь и далее под задачей диаризации будет подразумеваться задача разделения звукового потока на звуковой поток «оператора» и звуковой поток, содержащий голоса клиентов, шумы и прочее (словом – все, кроме голоса оператора).

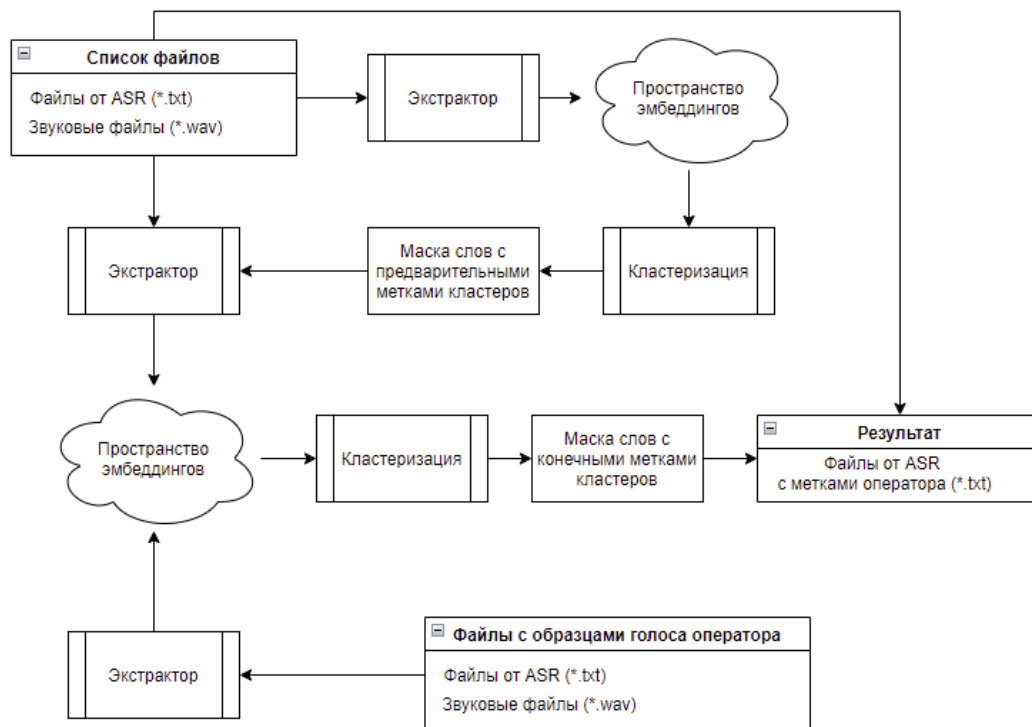


Рисунок. диаграмма работы биометрической диаризации

Пайплайн диаризации представлен на рисунке. Первый этап – получение данных от системы ASR по звуковым файлам, которые мы хотим обработать. Система ASR, в свою очередь, выдает файлы формата *txt, содержащие пары «слово – временная метка, в которую это слово было произнесено». Далее ASR-файлы вместе со звуковыми файлами передаются в «экстрактор», который извлекает биометрические признаки слов (например, мел-частотные кепстральные коэффициенты) и затем строит пространство эмбедингов для каждого слова. Это пространство кластеризуется на несколько кластеров (от 2), а метки кластеров для каждого слова («предварительные метки кластеризации»), вместе с ASR-файлами, снова передаются в экстрактор. На данном этапе экстрактор строит пространство эмбедингов для каждого кластера (то есть количество точек равно количеству кластеров, а не количеству слов, как было в первом экстракторе). На данном этапе к этому пространству добавляется эмбединг оператора, извлеченный экстрактором по ASR-файлам, полученным по звуковым дорожкам оператора. Здесь происходит финальная кластеризация (т.е. на 2 кластера). Затем конечные метки кластеров (для каждого слова) добавляются в ASR-файлы, что считается финальным этапом пайплайна кластеризации.

Для того, чтобы использовать пайплайн диаризации в байесовской оптимизации, нужно его привести к виду черного ящика – то есть, на вход «черного ящика» будет подаваться вектор гиперпараметров, а на выходе должно быть одно число, которое байесовская оптимизация и будет минимизировать.

Таким образом, согласно рисунку, в пайплайн подаются неизменяемые (не относящиеся к оптимизируемому пространству области значений гиперпараметров) протоколы обработки данных (инструкции по извлечению всех необходимых для диаризации файлов) и оптимизируемые гиперпараметры, а на выходе пайплайна диаризованные файлы будут сравниваться с эталонной разметкой, и на основе этих данных будет считаться некоторая ошибка, которая будет являться выходом «черного ящика».

В качестве метрики диаризации будем использовать WER (Word Error Rate) of operator (`wer_oper`) – частота ошибок в словах оператора.

Гиперпараметрами, в случае диаризации, будут являться параметры первичной кластеризации, поскольку она в наибольшей степени влияет на исследуемую метрику. Кластеризация также в наибольшей степени подвержена влиянию разных условий записи данных. Таким образом, для каждого набора данных может существовать свой набор параметров алгоритма кластеризации, показывающий наилучшие результаты.

Таким образом, в качестве алгоритма оптимизации была выбрана Байесовская оптимизация из-за большой вычислительной сложности исследуемой целевой функции

Пайплайн диаризации, работая как черный ящик, будет вычисляться множество раз в рамках оптимизации гиперпараметров. Чтобы сократить время вычисления байесовской оптимизации, было предложено кешировать часть данных, которая остается неизменной для каждого запуска пайплайна диаризации. Согласно рисунку, этой частью являются эмбединги слов, подаваемые в первичную кластеризацию. Результат работы диаризации с кешированием (на одной видеокарте GPU NVidia 1080ti с 11 Gb памяти) представлен в таблице 1. Таким образом, для системы оптимизации гиперпараметров удалось достичь сокращения времени выполнения диаризации на 40%.

Таблица 1. Результат работы кеширования в диаризации

	Время выполнения диаризации на всех файлах, мин.
С кешированием	10,5
Без кеширования	17,5

В рамках работы над оптимизатором также было реализовано распараллеливание вычислений. Диаризация принимает на вход файлы, а на выходе, при наличии файлов разметки, выдает метрики диаризации в JSON-формате. Далее оттуда извлекается метрика `wer_oper` для каждого из файлов, и считается среднее по ней. Идея параллельных вычислений состоит в том, чтобы на 3 видеокартах NVidia GeForce 2080ti с 11 Gb памяти запустить диаризации в 3 потока, предварительно разделив набор данных на 3 части, а затем собрать `wer_oper` со всех файлов

каждого из потоков. Для этого необходимо незначительно изменить формат выхода диаризации – теперь вместо среднего `wer_opr` по всем файлам диаризация должна возвращать массив `wer_opr` по всем файлам. Таким образом, после параллельных вычислений выходом является 3 массива, которые затем объединяются и по объединенному массиву считается среднее. Результаты работы диаризации с распараллеливанием вычислений представлены в таблице 2.

Таблица 2. Результат работы распараллеливания вычислений

	Время выполнения диаризации на всех файлах, мин.
С распараллеливанием (3 видеокарты NVidia GeForce 2080ti с 11 Gb видеопамяти)	8,5
Без распараллеливания (1 видеокарта NVidia GeForce 1080ti с 11 Gb видеопамяти)	10,5

Из таблицы видно, что количество времени не прямо пропорционально количеству видеокарт – время вычислений уменьшилось лишь на 19%.

Результат работы оптимизации гиперпараметров с различными суррогатными функциями, работающими при одинаковых условиях, представлены в таблице 3. Данные предоставлены для пилотного проекта крупной сетью автозаправочных станций. Данные содержат 153 аудиофайла разговоров оператора с клиентами.

Таблица 3. Результат работы оптимизации гиперпараметров на выборке данных

Библиотека	Суррогатная функция	Кол-во итераций	Время выполнения, мин	WER на выборке, %	WER на всех данных, %
Scikit-optimize	GP	100	47	51,52	57,64
		360	102	51,52	57,64
	Случайный поиск	100	25	54,64	58,11
		360	90	52,48	58,97
Hyperopt	TPE	100	31	52,60	59,14
		360	98	51,52	57,64

Как видно из таблицы, наилучшую сходимость показывает использование стандартных гауссовских процессов в качестве суррогатной функции. Также можно заметить, что TPE при 360 итерациях сошелся в то же значение WER, что и GP при 100 и 360 итерациях.

Обратим внимание на время работы оптимизатора с разными суррогатными функциями. При одинаковом количестве итераций быстрее всего работает случайный поиск (из-за того, что у него нет суррогатной модели, а соответственно он не тратит время и вычислительные ресурсы на ее вычисления).

Столбец «WER на выборке, %» означает результат работы оптимизатора на 2 файлах из полного набора. Найденный минимум является значением ячейки, а найденный вектор гиперпараметров применяется для вычисления значения столбца «WER на всех данных, %» на полном наборе данных.

Таким образом, GP является одновременно самой эффективной суррогатной функцией из исследуемых, и самой длительной с точки зрения времени работы оптимизатора. Точность работы диаризации по метрике «WER of operator» до применения оптимизатора составляла 66,43%. В результате работы оптимизатора ошибка снизилась с 66,43% до 57,64%.

В рамках данной работы были рассмотрены классические алгоритмы оптимизации гиперпараметров, особенности работы байесовской оптимизации и ее различные суррогатные функции. Также, был рассмотрен алгоритм работы диаризации и методы корректировки пайплайна диаризации для ее применения в алгоритме байесовской оптимизации как «черного ящика». В качестве гиперпараметров диаризации были приняты параметры первичной кластеризации. В процессе разработки оптимизатора были применены кеширование и распараллеливание вычислений с целью сократить время вычислений. В результате работы оптимизатора ошибка снизилась с 66,43 до 57,64, что составляет 13% от изначальной ошибки.

Литература

1. Lavanya Gupta Comparison of Hyperparameter Tuning algorithms: Grid search, Random search, Bayesian optimization. [Электронный ресурс]. – Режим доступа: <https://medium.com/analytics-vidhya/comparison-of-hyperparameter-tuning-algorithms-grid-search-random-search-bayesian-optimization-5326aaef1bd1> (дата обращения: 14.01.2022).
2. Bergstra J., Bengio Y. Random search for hyper-parameter optimization // Journal of Machine Learning Research. – 2013. – Pp. 281–305.
3. Will Koehrsen A Conceptual Explanation of Bayesian Hyperparameter Optimization for Machine Learning. [Электронный ресурс]. – Режим доступа: <https://towardsdatascience.com/a-conceptual-explanation-of-bayesian-model-based-hyperparameter-optimization-for-machine-learning-b8172278050f> (дата обращения: 14.01.2022).
4. Jones D., Schonlau M., Welch W. Efficient global optimization of expensive black box functions // Journal of Global Optimization. – 2013. – Pp. 455–492.
5. Bergstra J., Bardenet R., Bengio Y. Algorithms for hyper-parameter optimization // International Conference on Advances in Neural Information Processing Systems. – 2011. – №25. – Pp. 2546–2554.

УДК 004

РАЗРАБОТКА МЕТОДА ОЦЕНКИ ДОСТУПНОСТИ ГОРОДСКОЙ СРЕДЫ С ИСПОЛЬЗОВАНИЕМ ДАННЫХ СОЦИАЛЬНЫХ СЕТЕЙ

Низомутдинов Б.А.¹ (аспирант), Углова А.Б.²

¹*Университет ИТМО*

²*Российский государственный педагогический университет им. А.И. Герцена*

e-mail: boris@itmo.ru, anna.uglova@list.ru

Работа выполнена в рамках темы НИР №621300 «Разработка сервиса оценки доступности объектов городского благоустройства для маломобильных групп населения на основе данных из социальных сетей».

Исследование было направлено на разработку методики оценки доступности объектов городского благоустройства для маломобильных групп населения на основе анализа визуальных, текстовых данных из социальных сетей и социально-психологического благополучия жителей города. В статье описывается модель сбора данных из геоинформационных сервисов, которая в продолжении нашего исследования позволит в актуальном времени определять проблемы и ресурсы различных экологических зон города, доступность городского пространства для маломобильных групп населения и обозначить фокус внимания жителей города к тем или иным локациям. Для верификации данных, полученных через автоматизированный анализ, было проведено пилотное исследование мнения жителей города о социально-психологическом благополучии и качестве жизни в городе. В исследовании приняли участие 123 человека (80% женщин, 20% мужчин), из которых 62% составили представители маломобильных групп населения (родители с малолетними детьми, многодетные семьи, инвалиды, пенсионеры).

Ключевые слова: доступность, управление городом, парсинг, социальные сети, городская среда, анализ отзывов.

Доступность городской среды сегодня является важной задачей, в России реализуется большое количество программ по развитию городов, например, федеральная программа "Комфортная городская среда" [1] призвана развивать общественные пространства. Кроме того, реализуется большое количество адресных программ по благоустройству территорий в отдельных микрорайонах города и новостройках. Комфортная городская среда, как комплексный феномен, включает в себя благоустройство общественных пространств, дворов и парков, обеспечение безопасности, экологичности, а главное доступности всех частей города. Понятие доступности можно рассматривать с разных сторон, в данном исследовании изучается доступность рекреационных зон в городе для маломобильных граждан, например, для пенсионеров, инвалидов, родителей с маленькими детьми и других групп населения. В качестве рекреационных зон изучаются парки, сады и скверы, на примере Санкт-Петербурга.

Рассмотрим существующие методики оценки доступности среды для маломобильных групп, представленные в России, на федеральном и на региональном уровне. Примеры методик приведены в таблице.

Все эти методики подразумевают полевые исследования или анализ проектной документации, т.е. требуются ресурсы для оценки, кроме того, данные могут устаревать. В ситуации, когда речь идет о крупном городе, где может быть несколько сотен парков, получать данные в оперативном формате может быть затруднительно.

Таблица. Список методик оценки

Название документа	Год	Пояснение
Рекреационные зоны: парки и сады, бульвары и скверы	2008	Таблица анкеты состоит из пяти граф. Входная зона, пешеходные пути, площадки для отдыха, общественные туалеты, выводы
Оценка парков Москвы	2012	Была принята система критериев, использованных для Green Flag Award ³ , и адаптирована для условий Москвы. Были отобраны следующие 50 критериев, разбитые на 10 блоков. В том числе критерии включали в себя доступность парков для инвалидов
Оценка качества городского парка в контексте проблемы обеспечения устойчивого развития городских поселений	2018	Методика индексной оценки качества парков, в том числе включающая объекты доступности для инвалидов. Оценки были использованы для расчёта средних коэффициентов весомости и их нормирования, в результате чего все показатели приобрели собственные коэффициенты значимости
Методика паспортизации и классификации объектов и услуг с целью их объективной оценки для разработки мер, обеспечивающих их доступность	2013	Несколько типов анкет для различных объектов, в т.ч. парков. Анкеты содержат вопросы по: территории, входам/выходам, путям движения, санузелу, зоне целевого посещения. На основании этого объект признается полностью доступным, частично доступным, или недоступным и подлежащим адаптации
Методические рекомендации по организации деятельности по созданию условий для беспрепятственного доступа инвалидов и других маломобильных групп населения к объектам социальной инфраструктуры	2019	Содержит анкету по определению доступности объекта социальной инфраструктуры для инвалидов и других маломобильных групп населения, включает в себя параметры зон, техническое средство обустройства объекта, подходы к объекту, пути движения. Очень подробная методика оценки
Методика оценки уровня доступности объектов для инвалидов и маломобильных групп населения с использованием бальной системы	2014	Представлена методика оценки уровня доступности объектов с использованием бальной системы, рассчитываются коэффициенты доступности согласно алгоритму и с помощью формул

В нашем исследовании мы опирались на идеи социального конструктивизма, которые позволяют рассматривать контент социальных геоинформационных сетей как отражение реальных городских процессов. Был предложен оригинальный подход - оценка доступности городской среды по отзывам в социальных сетях.

В качестве основного источника информации, выбран сервис Google Maps. На первом шаге была изучена нормативная база, для использования данных. С развитием технологий появилась возможность использовать в качестве доказательств в судебном процессе сведения, полученные из геоинформационных систем. К ним можно отнести панорамы и фотоснимки с бесплатных сервисов Google Maps и «Яндекс.Карты» [2]. Подобные сведения в судебной практике последних лет довольно часто отвечают всем критериям оценки доказательств (хоть и не обладают легальным закреплением не только оценки, но и исследования), поэтому во многих случаях успешно подтверждают определенные факты. Однако, речь идёт об изображении, а не о комментариях пользователей. Судебных прецедентов, связанных с отзывами в Google Maps, обнаружено не было.

Дизайн исследования был основан на идеях Видодо Дж., который указывает, что для успешного улучшения инфраструктуры необходимо выстраивать диалог с жителями города, учитывать их культурный и цифровой опыт взаимодействия с материальными и нематериальными ценностями на разных уровнях городской среды [3]. Google Maps позволяет пользователям публиковать отзывы об объектах, которые имеются в базе данных системы, например, магазинах, но, помимо прочего, в этой системе также есть карты объектов отдыха - парков, садов и скверов, и пользователи также могут оставлять на них свои отзывы и оценки. Данное исследование проводилось на примере города Санкт-Петербург. Рекреационные объекты были собраны по районам, в городе 18 районов, для каждого из них с помощью внутреннего поиска Google Maps были выбраны объекты озеленения, для поиска использовались слова "парк", "сад", "площадь", "аллея", "лес", "поле", далее каждый объект сохранялся как ссылка на карточку объекта в системе. Все объекты были сохранены в итоговой таблице, дополнительно для каждой карты было записано общее количество отзывов и рейтинг. Рейтинг – это балл от 1 до 5. Для поиска отзывов от людей с ограниченной подвижностью был разработан скрипт на Python, содержащий базу данных поисковых слов. Кроме того, для анализа тона сообщений использовался сервис, основанный на методах машинного обучения. На рисунке 1 представлена общая модель сбора данных, которая проведена в ходе исследования.



Рис. 1. Общая схема сбора информации

В ходе анализа нам удалось собрать базу данных объектов отдыха, представленных в сервисе Google Maps, итоговая база данных включает 291 объект ландшафтного дизайна (парки, сады, скверы, аллеи). Было обработано более 360 000 отзывов. Все отзывы были обезличены. Тон был определен для отзывов, отдельно был проведен поиск отзывов, описывающих опыт использования парков маломобильными группами населения.

Дополнительно проводилось анкетирование. Для оценки потребностей жителей города в развитии инфраструктуры мы разработали анкету в виде вопросов-высказываний о личных потребностях в развитии города, а также коммуникационных стратегиях взаимодействия с городской средой в онлайн- и офлайн-пространстве: реальный опыт активного участия в городских сообществах, использование информационных ресурсов социальных сетей для взаимодействия с городом. Было обследовано 123 человека (80% женщин, 20% мужчин), из которых 62% составили представители маломобильных групп населения (родители с малолетними детьми, многодетные семьи, инвалиды, пенсионеры).

В статье также описывается модель сбора данных из геоинформационных сервисов, которая в будущем будет использоваться для определения проблем и ресурсов различных экологических зон города. Мы проанализировали субъективную оценку качества жизни в

городской среде жителей Санкт-Петербурга, чтобы выявить потребности в развитии инфраструктуры, описать "кризисные точки" городского благоустройства, повысить качество управления городскими ресурсами. На рисунке 2 представлены результаты опроса о доступности различных объектов городской инфраструктуры.

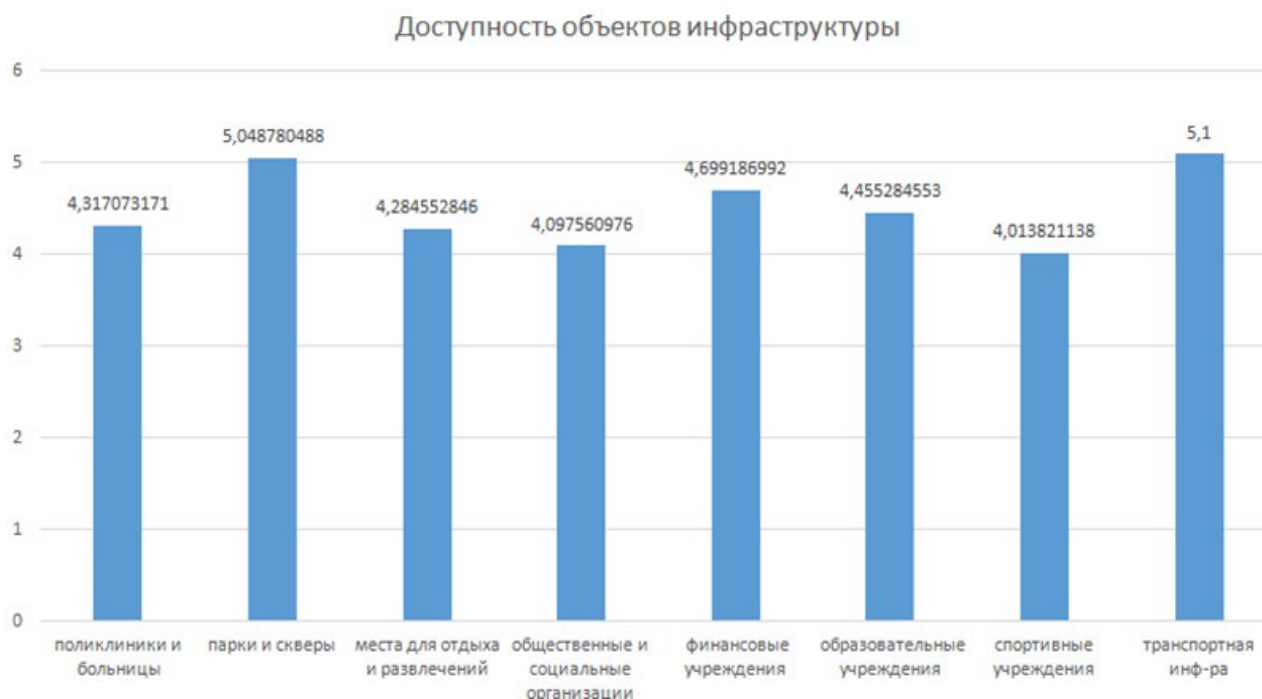


Рис. 2. Итоги опроса жителей

На следующем этапе исследования запланирован поиск корреляций между результатами опроса и данными, собранными в социальных сетях, результаты запланировано представить в виде картографического сервиса.

Исследовательский проект имеет потенциал и значимость для исследований и политики в области городского планирования и развития. Научная значимость решения проблемы заключается в междисциплинарном подходе, сочетающем традиционную методологию социальных исследований с возможностями современных автоматизированных систем на основе машинного обучения, которые позволяют оптимизировать процесс получения информации и могут способствовать формированию новых научных направлений.

Литература

1. Федеральный проект «Формирование комфортной городской среды» - 2020. [Электронный ресурс]. – Режим доступа: <https://www.minstroyrf.gov.ru/trades/gorodskaya-sreda/strategicheskoe-napravlenie-razvitiya-zhkkh-i-gorodskaya-sreda/> (дата обращения: 03.01.2022).
2. Шкурова П.Д. Использование сведений, полученных из сервисов Google Maps и «Яндекс.Карты», в качестве доказательств в практике арбитражных судов Российской Федерации // Сибирское юридическое обозрение. – 2021. – №18(1). – С. 83–91.
3. Widodo J. Urban Environment and Human Behaviour: Learning from History and Local Wisdom // Social and Behavioral Sciences 42:6–11. DOI: 10.1016/j.sbspro.2012.04.161.

УДК 004.91

ТЕХНОЛОГИИ ВЫЯВЛЕНИЯ В ТЕКСТАХ ИНДИКАТОРОВ СОЦИАЛЬНОГО САМОЧУВСТВИЯ

Огурцова Е.В.¹ (студент)

Научный руководитель – канд. социол. наук Видясова Л.А.¹

¹Университет ИТМО

e-mail: ek.ogurtsova@gmail.com

Работа выполнена в рамках НИР №621304 «Разработка сервиса тематической кластеризации корпуса текстов «Развитие цифрового государственного управления в Российской Федерации» на основе машинного обучения».

Статья посвящена технологиям выявления в текстах таких индикаторов социального самочувствия как ключевые слова и тональность. Рассматривается понятие социального самочувствия, его индикаторы, а также программное обеспечение необходимое для выявления описываемых индикаторов в текстах пользователей в социальных сетях.

Ключевые слова: социальное самочувствие, технологии сбора данных, анализ тональности текста, парсинг, инцидент-менеджмент, социальные сети.

На данный момент такое явление как социальное самочувствие является одним из важных индикаторов удовлетворенности населения жизнью в стране и действиями властей. Социальное самочувствие – это своего рода субъективная оценка людьми качества жизни на конкретной территории: в городе, регионе или стране. В последнее время все больше государственных программ направлено на повышение качества жизни и благосостояние населения, этому свидетельствуют, например, проект «Умный город» [1] национальной программы «Цифровая экономика» и одно из направлений внутри проекта, такое как «Благосостояние людей», а также цель государственной программы «Информационное общество», установленная в Постановлении Правительства РФ от 31 марта 2021 г. № 504–19 «О внесении изменений в государственную программу Российской Федерации «Информационное общество», которая состоит в «повышении качества жизни и работы граждан ... на основе использования информационных и телекоммуникационных технологий» [2].

Сегодня существует множество различных определений социального самочувствия, но мнения исследователей сходны относительно таких элементов социального самочувствия как: оценочность, эмоциональность и субъективность, а также зависимость оценки от окружающей среды. На основе анализа существующих определений [3–6] социального самочувствия это явление можно определить как субъективную оценку человеком окружающей его социальной реальности, своего социального положения, динамики экономической и политической ситуации, а также ее влияния на настроение индивида и принимаемые им решения. В таблице приведены определения социального самочувствия из научной литературы и статей в СМИ.

Индикаторы социального самочувствия и технологии их выявления в тексте

Для оценки социального самочувствия населения исследовательские организации проводят социологические опросы населения в формате личных интервью и телефонных опросов. Социальное самочувствие идентифицируется через различные показатели: удовлетворенность жизнью, социальный оптимизм, материальное положение, экономическое

положение страны, политическая обстановка, общий вектор развития страны, недовольство властями, готовность населения выйти на протестные акции.

Таблица. Определения социального самочувствия

Социальное самочувствие – это...	Автор	Год
базовый элемент социального настроения	Тощенко Ж.Т.	1998
целостная, относительно устойчивая эмоциональная реакция субъекта на воздействие социальной среды и условий его жизнедеятельности	Бочканова Е.Н.	2007
система субъективных ощущений, свидетельствующая о той или иной степени психологической и физиологической комфортности	Кобозева Н.Н.	2007
эмоционально-оценочное отношение индивидов к окружающей социальной реальности и своему месту в ней	Суняйкина О.Н.	2011
эмоциональная оценка людьми своего общественного положения, уровня удовлетворения социально-экономических и духовных потребностей	Слободян Е. («Аргументы и факты»)	2014

Отдельно стоит отметить подход Левада-Центра¹ к оценке социального самочувствия – в исследовании данной организации индикаторами социального самочувствия были непосредственно чувства трёх групп [7]:

- ожесточение, страх, обида, отчаяние, зависть;
- усталость, растерянность и одиночество;
- надежда, достоинство, уверенность, свобода, гордость, ответственность.

Говоря же об индикаторах социального самочувствия непосредственно в тексте, можно выделить два типа индикаторов:

- ключевые слова;
- тональность текста.

Ключевыми словами-индикаторами можно назвать, например, те слова, которые отражают те или иные аспекты жизни человека, относящиеся к сфере социального самочувствия. Так, в статье «Big Data в определении социального самочувствия населения России» авторы выделяют индикаторы социального самочувствия в виде ключевых слов по нескольким аспектам жизни человека в текстовых запросах Google, среди них: материальные условия – цены на жилье, банкротство, биржа труда, МРОТ и др.; физическое и ментальное здоровье – стресс, хроническая усталость, диета, отпуск, наука, смысл жизни, спорт и др. [8]; социальные условия и окружающая среда – воспитание ребенка, опека, поддержка семьи, оружие, социальная защита, нормативные акты, уровень жизни, охрана природы и др. Всего исследователями было отобрано 382 слова отражающих компоненты категорий субъективного благосостояния. В данном исследовании было применено web-приложение Google Trends, но также для решения подобных задач может быть использован сервис оценки пользовательского интереса к конкретным тематикам и подбора ключевых слов – Яндекс Wordstat. Также, нельзя не упомянуть такую мониторинговую систему как «Инцидент-менеджмент», с помощью которой органы власти и исследовательские организации могут проводить автоматический поиск и сбор комментариев пользователей в социальных сетях по ключевым словам. В дальнейшем, собранные комментарии могут быть проанализированы с точки зрения тональности текста и полученные данные также могут быть использованы для оценки социального самочувствия.

Как уже упомянуто выше, кроме ключевых слов индикатором социального самочувствия в тексте может служить его тональность. Для определения тональности текста применяется анализ тональности, или по-другому сентимент-анализ – класс методов контент-анализа в компьютерной лингвистике, основная задача которого заключается в классификации текста по его настроению [9]. Наиболее распространены двухчастный и трехчастный варианты классификации, где тексты классифицируются как негативный или позитивный, или же негативный, нейтральный и позитивный. Существуют также подходы, когда исследователи

¹ АНО «Левада-Центр» принудительно внесена Минюстом в реестр некоммерческих организаций, выполняющих функции иностранного агента

вычленяют эмоциональные состояния, которые ассоциируются с тем или иным текстом. Например, это могут быть страх, злость, счастье. В некоторых подходах текст располагается на шкале, в зависимости от эмоциональной окраски, от отрицательного числа, обозначающего крайнюю степень негативности, до положительного, наоборот, обозначающего позитивность текста. Соответственно социальное самочувствие может быть идентифицировано посредством анализа тональности текстов комментариев пользователей в социальных сетях, непосредственно обращений граждан на платформах обратной связи. Среди инструментов для сентимент-анализа можно выделить MonkeyLearn, поскольку данный сервис дает возможность создания своей модели анализа настроений для отраслевых текстов, а также предоставляет бесплатную базовую версию программы. Дополнительно инструменты для сентимент-анализа текстов можно изучить в каталоге ресурсов для обработки естественного языка NLPub [10].

Говоря о том, как именно тональность текста идентифицирует самочувствие человека следует привести пример: исследование связи статуса человека в Facebook и его субъективного благополучия показало, что люди, которые выражали больше негативных эмоций при обновлении статуса в Facebook, также сообщали о более низкой удовлетворенности жизнью, но в то же время, большее количество положительных эмоций, о которых сообщалось в Facebook, не было связано с более высокой удовлетворенностью жизнью [11]. Отсюда можно сделать вывод, что присутствует связь между сообщениями в социальных сетях человека и его самочувствием, а наиболее репрезентативные эмоции в социальных сетях – негативные.

При рассмотрении сентимент-анализа текстов встает вопрос о том, где и как брать текстовые данные. Отдельно хочется упомянуть веб-скрейпинг как практику сбора открытых данных с загруженных веб-страниц и форм, не предназначенных для этого [12]. Для таких задач существуют специальные программы, или парсеры, которые обращаются к веб-серверам для запроса данных и последующей их обработке – Scraper API, Octoparse, DataOx, ScrapingBot, Wintr и др.

Таким образом, под социальным самочувствием понимается субъективная оценка человеком окружающей его социальной реальности, своего социального положения, динамики экономической и политической ситуации, а также ее влияния на настроение индивида и принимаемые им решения. Можно выделить два типа индикаторов социального самочувствия в тексте: ключевые слова и тональность текста. Текстами для выявления индикаторов социального самочувствия могут быть запросы пользователей в поисковых системах Яндекс и Google, а также комментарии в социальных сетях. Основными технологиями и инструментами для выявления ключевых слов-индикаторов социального самочувствия в тексте можно назвать сервисы статистики поисковых запросов, такие как GoogleTrends и Яндекс Wordstat, а также различные системы автоматизированного мониторинга, как, например, «Инцидент-менеджмент». Для выявления такого индикатора социального самочувствия в тексте как его тональность основными инструментами будут сервисы для анализа тональности текстов, в том числе те, где сентимент-анализ одна из возможных функций, а также различные парсеры для сбора текстовых данных.

Литература

1. Умный город. [Электронный источник]. – Режим доступа: <https://russiasmartcity.ru/> (дата обращения: 01.03.2022).
2. Постановление Правительства РФ от 31 марта 2021 г. № 504–19 “О внесении изменений в государственную программу Российской Федерации “Информационное общество”. [Электронный источник]. – Режим доступа: <https://www.garant.ru/products/ipo/prime/doc/400466555/> (дата обращения: 01.03.2022).
3. Что такое социальное самочувствие и где в России оно лучше? | Вечные вопросы | Вопрос-Ответ | Аргументы и Факты. [Электронный источник]. – Режим доступа: <https://aif.ru/dontknows/eternal/1193899> (дата обращения: 01.03.2022).
4. Бочканова Е.Н. Социальное самочувствие горожан – ориентир муниципальной социальной политики // Известия Уральского государственного университета. – 2007. – №51. – С. 29–36.

5. Кобозева Н.Н. Социальное самочувствие как социологическая категория // Вестник Ставропольского государственного университета. – 2007. – №50. – С. 117–122.
6. Суняйкина О.Н. Понятие "социальное самочувствие" в социологии // Вестник Мордовского университета. – 2011. – №3. – С. 98–101.
7. Социальное самочувствие россиян. [Электронный ресурс]. – Режим доступа: <https://www.levada.ru/2013/12/11/sotsialnoe-samochuvstvie-rossiyan/> (дата обращения: 01.03.2022).
8. Фантащини Д., Шаклеина М.В., Юрас И.А. Big Data в определении социального самочувствия населения России // Прикладная эконометрика. – 2018. – №2(50). – С. 43–66.
9. Smetanin S. The Applications of Sentiment Analysis for Russian Language Texts: Current Challenges and Future Perspectives // IEEE Access. – 2020. – Vol.8. – Pp. 110693–110719.
10. Анализ тональности. [Электронный источник]. – Режим доступа: <https://nlpub.ru/> (дата обращения: 01.03.2022).
11. Do Facebook Status Updates Reflect Subjective Well-Being? | Cyberpsychology, Behavior, and Social Networking. [Электронный источник]. – Режим доступа: <https://www.liebertpub.com/doi/10.1089/cyber.2015.0022> (дата обращения: 01.03.2022).
12. Вилкова О.В. К вопросу о научной осмысленности применения веб-скрейпинга как метода сбора данных в социологических исследованиях // Вестник Томского государственного университета. Философия. Социология. Политология. – 2020. – №54. – С. 164.

УДК 004.422.81

ИССЛЕДОВАНИЕ ВОЗМОЖНОСТИ ОБЪЕДИНЕНИЯ СИСТЕМ "УМНЫЙ ДОМ" РАЗЛИЧНЫХ ПРОИЗВОДИТЕЛЕЙ

Парошин Я.Д.¹ (студент), Платонов Н.С.¹ (студент), Чёрный П.П.¹ (студент)

Научный руководитель – канд. техн. наук, доцент Маркина Т.А.¹

¹Университет ИТМО

e-mail: yqzarov@gmail.com, pavelchj@gmail.com

Работа выполнена в рамках темы НИР №620164 «Методы искусственного интеллекта для киберфизических систем».

В работе исследованы проблемы интеграции систем умных домов различных производителей, рассмотрены возможные способы и предложено комплексное решение, которое включает в себя три аспекта: интеграция устройств стороннего производителя в данную экосистему умного дома; интеграцию данной экосистемы умного дома в устройства стороннего производителя; создание цифрового двойника конечных устройств экосистемы. Дополнительно были разработаны прототипы каждого из описанных выше элементов решения.

Ключевые слова: Internet of Things, IoT, интернет вещей, умные системы, умный дом, интеграция, тестирование, цифровой двойник, обобщенный протокол.

«Интернет вещей» – это вычислительная сеть физических объектов, оснащенных встроенными технологиями сбора и передачи информации в совокупности с устройствами и технологиями хранения и интеллектуальной обработки информации, а также устройствами и алгоритмами генерации управляющих воздействий как на части системы, так и глобальных [1].

IoT-системы могут быть автономными или предоставлять пользователю возможность взаимодействовать с ними через настройку или предоставление доступа к собранным данным. Как правило такие системы работают в режиме реального времени.

Среди IoT-систем, или «умных» систем, можно выделить следующие направления: умный транспорт, умные промышленные системы, умное сельское хозяйство, умные города, умные дома. Последней области посвящена данная работа.

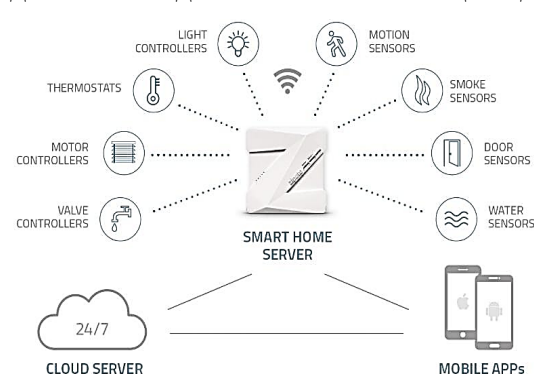


Рис. 1. Пример архитектуры системы умного дома

Под «умным домом» понимается комплекс технологий, которые позволяют экономить время на домашние дела посредством их автоматизации. Последнюю обеспечивают системы электронных и электронно-механических устройств, которые решают определенные домашние задачи без вмешательства человека. На рисунке 1 приведен пример архитектуры системы умного дома.

На данный момент существует большое количество производителей как отдельных устройств, так и целых экосистем, которые значительно отличаются по технологиям передачи данных и протоколам.

Из-за этого складывается ситуация, что потребитель, отдавая предпочтение той или иной системе или отдельным устройствам, вынужден жертвовать преимуществами других, которые не могут взаимодействовать с той, которую он выбрал, что доставляет большие неудобства. Для решения данной проблемы были исследованы возможные способы интеграции систем и устройств различных производителей друг с другом и разработан прототип.

Интегрировать в экосистему умного дома одного производителя устройства другого можно тремя способами:

- 1) конечный пользователь с помощью специальных плагинов;
- 2) использование единого протокола работы;
- 3) разработка промежуточного звена, которое будет выступать «переводчиком» с одного протокола на другой;

В данной работе были проанализированы первый и второй способы.

Для интеграции выбрана компания Apple, так как является одним из самых крупных игроков на рынке производства электроники на сегодняшний день. Но, несмотря на популярность компании, устройств умного дома, поддерживающих возможность подключения в их экосистему, не так много. Для того, чтобы интегрировать свои устройства, стороннему производителю необходимо получить специальную сертификацию. Если пользователь хочет интегрировать устройство, у которого нет сертификации Apple, он может собственноручно сделать это путем установки или написания с нуля соответствующего плагина.

Для того, чтобы иметь возможность управлять умным домом за пределами локальной сети Apple, необходим домашний центр. В качестве домашнего центра может выступить Apple TV, HomePod или iPad. Управление всеми устройствами умного дома происходит через приложение Home, предустановленное на устройствах Apple. Также существует фреймворк HomeKit, который позволяет настраивать, управлять и общаться с умными домашними устройствами.

Типы устройств умного дома Apple:

- HomeKit ready – устройства, которые добавляются напрямую в HomeKit и управляются через контроллер Apple;
- не поддерживающие HomeKit напрямую, но добавляются через сторонние концентраторы (например, Raspberry Pi, Athom Homey);
- устройства со специальным программным обеспечением;
- устройства, работающие через сторонний контроллер.

Для интеграции в экосистему устройств, которые изначально не поддерживаются в HomeKit, необходимы специальные серверные системы с возможностью подключения модулей и плагинов, которые позволят взаимодействовать с устройствами. Одной из таких систем является Homebridge.

Homebridge – это сервер NodeJS, который эмулирует API iOS HomeKit. Он поддерживает плагины, которые являются модулями, созданными сообществом. Эти модули обеспечивают мост между HomeKit и сторонними API, которые предоставляются производителями устройств для умного дома [2].

Для работы написанных плагинов на компьютере должен быть развернут сервер NodeJS. Пользователь может установить свой плагин в Homebridge, подключить устройство, для которого необходим этот плагин и после этого оно отобразится в Homebridge.

Этот метод был реализован на примере умной лампы от фирмы GS, для чего был разработан плагин на языке программирования JavaScript, связывающий устройство и Homebridge. Данный плагин позволяет управлять этой лампой, а именно: включать и выключать, регулировать яркость свечения. Результат представлен на рисунке 2.

Для каждого производителя «умных» устройств является актуальным вопрос взаимодействия со сторонними устройствами. В случае, если у производителя имеется своя полноценная экосистема умного дома, этот вопрос выходит на первый план. Для решения этой

проблемы он может интегрировать свою систему в экосистемы других крупных игроков, таких как Яндекс, Apple и т.д. Однако, использование только этого решения не является перспективным для компании, так как не способствует развитию внутренней экосистемы.

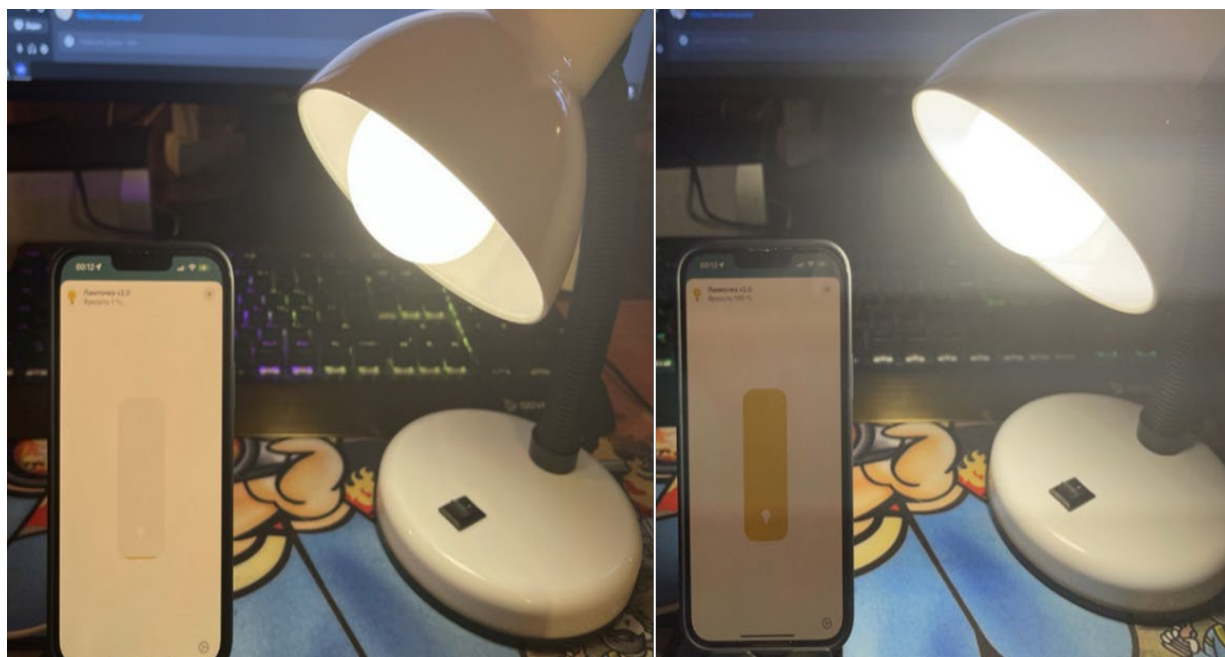


Рис. 2. Изменение яркости свечения лампочки с помощью плагина

Необходим универсальный инструмент, который позволит сторонним производителям полноценно интегрироваться в существующую экосистему, не затрачивая большие ресурсы. Для этого он должен иметь прозрачную структуру, подробную спецификацию и базироваться на одном из популярных существующих протоколов для упрощения процесса интеграции.

На сегодняшний день одними из самых популярных в России являются Zigbee и протокол умного дома Яндекс.

Dyson, Honor, LG, Panasonic, Xiaomi и еще более ста пятидесяти поставщиков устройств и систем умного дома имеют сертификацию Zigbee.

В сети ZigBee существует три типа устройств: координатор, роутер и конечные устройства. Набор свойств конечных устройств и роутеров определяется двумя параметрами – набором конечных точек и кластеров. При этом каждая конечная точка характеризуется набором кластеров. Каждый кластер имеет свой уникальный в пределах спецификации номер. Каждый кластер характеризуется специфическим набором атрибутов и команд. В устройстве может быть более одной конечной точки [3].

LG, Philips, Redmond, Samsung, Xiaomi и еще более восьмидесяти поставщиков устройств и систем умного дома поддерживают интеграцию с платформой умного дома Яндекс.

Последняя помогает объединять устройства от различных производителей в одну экосистему, которая обладает голосовым интерфейсом. Производителю достаточно обеспечить поддержку открытого протокола управления устройствами на стороне своего сервиса.

Открытый протокол управления устройствами разработан на базе HTTP REST API и его можно условно поделить на две части: серверную и клиентскую.

Серверная часть собирает данные устройств всех производителей и отправляет информацию об изменении их состояний. Клиентская часть уведомляет платформу умного дома Яндекс об изменении параметров и состояний устройств [4].

Решение обозначенной проблемы должно быть представлено в виде серверной части, которая встраивается в центральный контроллер умного дома и клиентской части, которая встраивается в устройства сторонних производителей и распространяется в виде библиотеки.

ZigBee имеет сложную структуру, а спецификация составляет более 1200 страниц. Для производителей устройств, которые построены на других протоколах будет затруднительно интегрироваться в систему, построенную на базе ZigBee.

Платформа умного дома Яндекс, в свою очередь, уже поделена на требуемые части и имеет простую структуру, а передача запросов происходит по HTTP с использованием JSON-документов, что облегчает реализацию обработчиков команд. Платформа умного дома Яндекс функционирует через облачное хранилище и разработчику остаётся незначительно переработать структуру запросов и JSON-документов. Исходя из изложенного выше, было принято решение использовать протокол умного дома Яндекс.

Первым шагом реализации универсального протокола стало написание GET запроса со стороны сервера, который получает конфигурацию девайсов клиента в JSON-формате. На рисунке 3 отображена процедура подключения клиента к серверу. На рисунке 4 показан результат отправки запроса GET devices.

The image shows two terminal windows side-by-side. The left window is the server terminal, and the right window is the client terminal. In the server terminal, a menu is displayed with options: 1 - Send a message to the client, 2 - Start server, 3 - Add Client, 4 - Close server, 10 - Exit, 0 - Show menu. The user enters '2' to start the server. The server then displays 'Waiting for the client.' and 'Enter the desired number in the menu (To display the menu, enter 0):'. The client terminal shows 'Connected to server!' and a cursor at the prompt.

Рис. 3. Подключение клиента

The image shows a terminal window with a dark background. At the top, there are tabs for 'ПРОБЛЕМЫ', 'ВЫХОДНЫЕ ДАННЫЕ', 'ТЕРМИНАЛ', and 'КОНСОЛЬ ОТЛАДКИ'. The terminal shows the server waiting for a client, then receiving the input '1'. The server prompts 'Enter the message you want to send:' and the user enters 'GET devices'. The server then returns a large JSON object representing the client's device configuration. The client terminal on the right shows 'Connected to server!' and 'Server: GET devices'.

Рис. 4. Запрос на передачу конфигурации устройств

Стандартный метод тестирования систем умного дома подразумевает использование физических устройств. При этом отсутствует возможность выявить проблемы на этапе разработки, появляется необходимость их производства (покупки), возрастают временные затраты, что увеличивает стоимость конечного продукта для компании и конечного пользователя.

Данный процесс возможно автоматизировать посредством разработки цифрового двойника системы.

Для разработки цифрового двойника системы умного дома необходимо проанализировать целевую аудиторию продукта. Были определены следующие группы потребителей:

- тестировщики систем умного дома;
- компании, производящие системы умного дома;
- пользователи систем умного дома.

Для формирования требований к продукту были проанализированы потребности этих групп, на основе чего сформулированы требования, согласно которым решение должно:

- иметь удобный интерфейс;
- быть кроссплатформенным;
- поддерживать универсальный протокол;
- иметь БД SQLite, содержащую необходимую информацию о ZigBee-устройствах;
- давать возможность импортировать, экспортировать и создавать сценарии работы цифровых двойников устройств;
- быть реализовано на базе фреймворка Qt;

Архитектура решения представлена на рисунке 5.

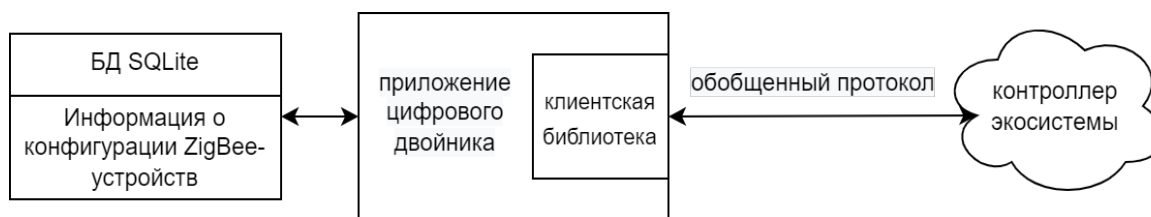


Рис. 5. Архитектура приложения цифрового двойника системы умного дома

Далее было разработано графическое приложение, обладающее следующими функциональными возможностями:

- добавление новых устройств, их типов и начальных состояний;
- создание сценариев поведения устройств;
- экспорт сценариев в формате универсального протокола.

На рисунках 6 представлены окна приложения, предназначенные для работы со сценариями (рис. 6, а) и добавления устройств (рис. 6, б).

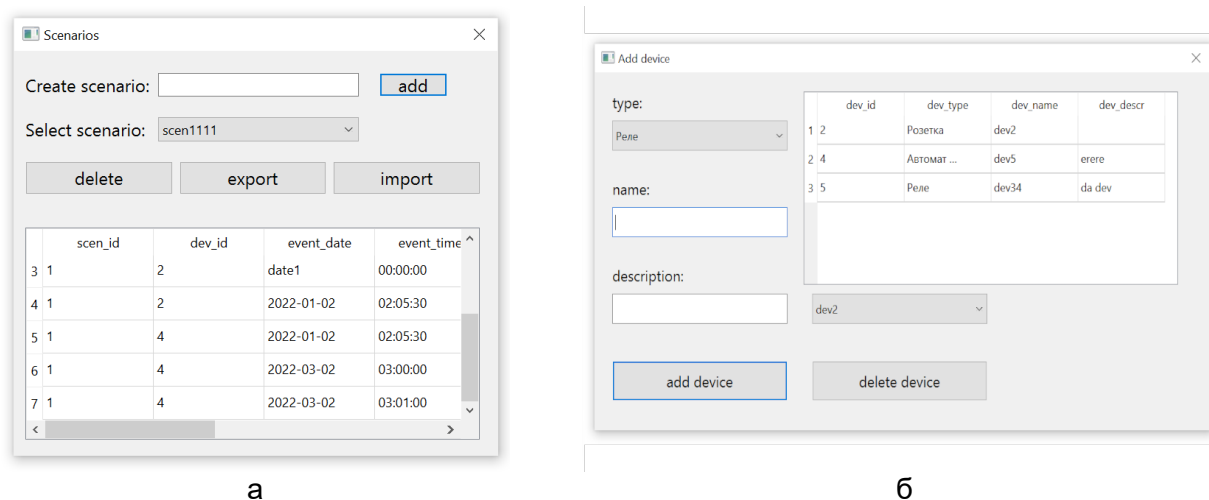


Рис. 6. Вид окна работы: а) со сценариями; б) с устройствами

За основу взяты данные спецификации протокола ZigBee как одного из самых используемых в устройствах интернета вещей, впоследствии может быть добавлена поддержка устройств, работающих по другим протоколам.

В дальнейшем в работе планируется завершить поддержку и совместимость с разрабатываемым универсальным протоколом, добавить часы для отсчета времени отправки команд и проверку корректности исполнения устройствами заданных сценариев с соответствующим логированием.

Литература

1. Промышленный интернет вещей. [Электронный ресурс]. – Режим доступа: <https://investmoscow.ru/media/3340535/03-промышленный-интернет-вещей.pdf> (дата обращения: 25.01.2022).
2. Homebridge Plugin Development. [Электронный ресурс]. – Режим доступа: <https://developers.homebridge.io/> (дата обращения: 10.11.2021)
3. Zigbee Cluster Library Specification. [Электронный ресурс]. – Режим доступа: <https://csa-iot.org/wp-content/uploads/2022/01/07-5123-08-Zigbee-Cluster-Library-1.pdf> (дата обращения: 25.01.2022).
4. Протокол работы сервиса уведомлений. [Электронный ресурс]. – Режим доступа: <https://yandex.ru/dev/dialogs/smart-home/doc/reference-alerts/resources-alerts.html> (дата обращения: 26.01.2022).

УДК 004.94

ПРИМЕНЕНИЕ СТРУКТУРНЫХ ДЕСКРИПТОРОВ ДЛЯ ПОИСКА ЗАКОНОМЕРНОСТЕЙ В 3D-СТРУКТУРАХ РЕЦЕПТОРА ВИТАМИНА D

Пац К.М.¹ (аспирант), Виноградова Е.А.¹ (студент)

Научный руководитель – канд. техн. наук Сергушичев А.А.¹

¹Университет ИТМО

e-mail: karina.m.pats@gmail.com

Работа выполнена в рамках темы НИР №621314 «Автоматизированный поиск переходных состояний в химических реакциях».

Работа посвящена исследованию возможности структурных дескрипторов выявлять закономерности как в статичных 3D-структурах рецептора витамина D, так и в структурах, полученных с помощью молекулярной динамики. Показана возможность дескрипторов выявлять видовые различия, структурные выбросы, а также учитывать влияние присоединяемых лигандов и коактиваторов на структуру рецептора.

Ключевые слова: структурные дескрипторы, ядерные рецепторы, молекулярное моделирование, аллостерические взаимодействия, анализ главных компонент, Мантел-тест.

Структурные дескрипторы широко используются в области компьютерного моделирования в естественных науках и представляют собой математическое описание каких-либо свойств структуры белка. Структурные дескрипторы зарекомендовали себя как надежный вычислительный инструмент при определении принадлежности белковых молекул к тому или иному семейству [1–2]. Данная работа направлена на исследование возможности применения таких дескрипторов при поиске структурных закономерностей, определяющих аллостерические взаимодействия рецептора витамина D. Рецептор витамина D (VDR) является высокоспециализированным рецептором, связывающимся с 1 α ,25-дигидроксивитамином D₃ (1,25D₃). Помимо функции поддержания уровня кальция и фосфора в процессе развития костей VDR может быть вовлечен в процессы клеточной пролиферации, формирования аутоиммунных реакций, воспалительные процессы, а также в механизмы развития рака. Как представитель семейства ядерных рецепторов, VDR состоит из двух доменов – ДНК- и лиганд-связывающего [3]. Последний играет важную роль в регуляции активности ядерных рецепторов. При связывании лиганда, структура рецептора подвергается конформационным изменениям, в том числе в регионе терминальной α -спирали H12, которая отвечает за присоединение к рецептору белковых корегуляторов. Ранее нами было показано, что с помощью дескрипторов, описывающих спирали 6–7 рецептора витамина D, а также процент заполняемости его сайта связывания, удалось выделить группы структур со схожими лигандами, как по структуре, так и по свойствам (рис. 1). Кроме этого, мы разработали собственный набор структурных дескрипторов, описывающих вторичную структуру белковой молекулы, а также участки связывания с белковыми корегуляторами, специфичные именно для рецептора витамина D. Поэтому целью работы было исследовать способность разработанного нами набора выявлять аналогичные закономерности.

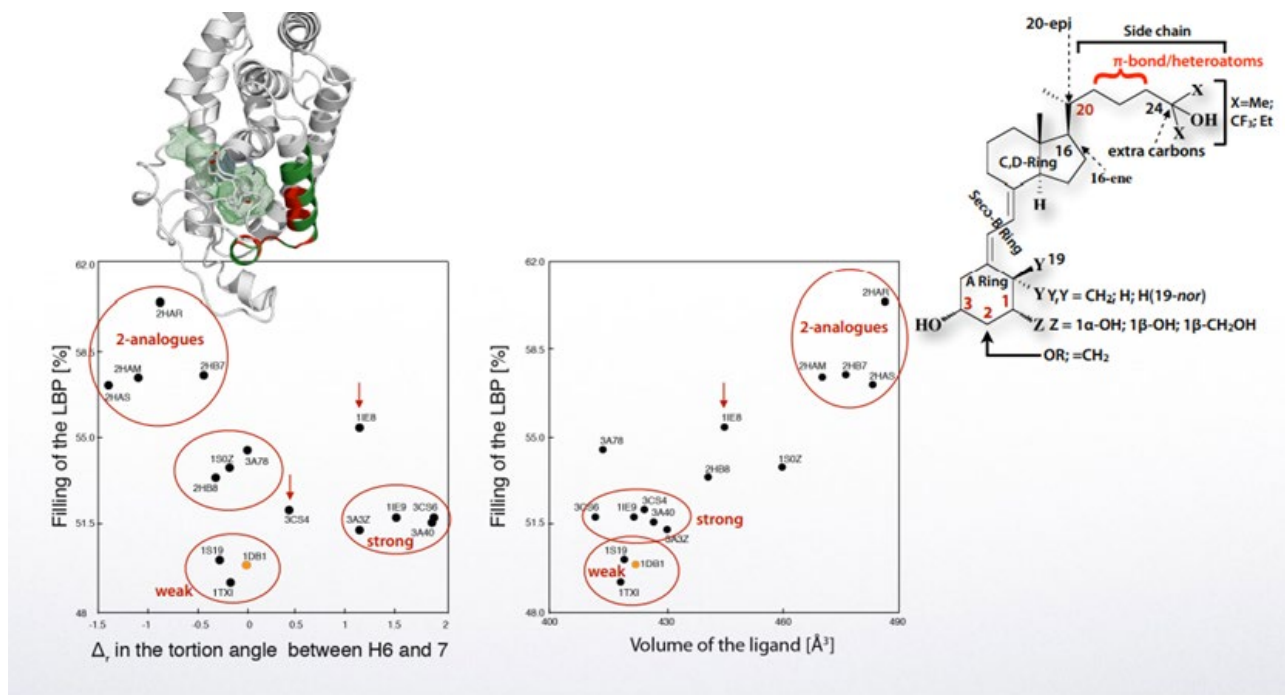


Рис. 1. Предварительное исследование способности ряда структурных дескрипторов выявлять закономерности в структурах рецептора витамина D

В ходе работы исследовалось два типа структур – статичные структуры, полученные из базы данных PDB, а также структуры, полученные с помощью молекулярной динамики. В первом случае использовались структуры трех видов – человека (hVDR), серой крысы (rVDR) и данио-рерио (zVDR). В случае анализа структур, полученных в ходе молекулярной динамики, исследовалось 18 траекторий hVDR в комплексе с тремя различными лигандами (VD3 – кальцитриол или $1\alpha,25$ -дигидроксивитамин D₃; GEM – $1\alpha,25$ -дигидрокси-21-(3-гидрокси-3-метилбутил)витамин D₃; TX527 – 19-нор-14,20-бис-эпи-23-ин- $1\alpha,25$ -дигидроксивитамин D₃) и двумя коактиваторами (SRC1 и PGC1a), содержащих по три мотива, связывающихся с VDR (L1, L2, L3). Время симуляции составило 10нс. Для всех используемых структур были рассчитаны 10 основных дескрипторов, использующих предсказания инструмента KPAХ [4] для определения вторичной структуры. Помимо основных дескрипторов, вторичная структура была также определена при помощи алгоритма DSSP, и разница предсказаний этих двух инструментов была использована как отдельный дескриптор. Таким образом, в результате расчетов количество признаков составляет порядка 250 в зависимости от конкретной структуры, т.к. нулевые значения удаляются.

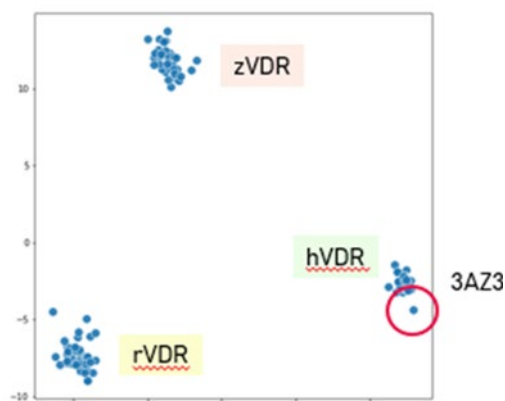


Рис. 2. Разделение структур по видам при помощи дескрипторов (PCA)

При анализе структур из базы PDB, нам удалось выявить способность разработанного нами набора дескрипторов выделять группы структур, принадлежащие различным видам (рис. 2). Помимо этого, при помощи дескрипторов удалось выявить структуру-выброс с PDB ID 3AZ3. При более детальном исследовании этой структуры выяснилось, что аминокислоты на участке 290-299 отличаются от остальных структур hVDR.

Во время анализа траекторий молекулярной динамики мы исследовали возможность замены структурными дескрипторами параметра RMSD (root mean square deviation), который является одним из наиболее популярных параметров для оценки качества молекулярной динамики и анализа полученных траекторий. Для этого мы выбрали комплексы VDR с лигандами VD3 и GEM, а также с коактиватором PGC1a.

Для сравнения результатов расчета дескрипторов и RMSD для одной и той же траектории мы извлекли 100 кадров и построили матрицы RMSD, отражающие отклонение каждого кадра относительно всех остальных. Также для этих 100 кадров мы рассчитали значения структурных дескрипторов и отобразили в двухмерное пространство с использованием метода анализа главных компонент, после чего вычислили попарные расстояния между всеми точками в этом пространстве, получив еще одну матрицу расстояний. Такое представление данных дает возможность использовать тест Мантела для сравнения двух матриц. В результате данного анализа мы видим, что в случаях присутствия кластеров на матрицах RMSD, мы также наблюдаем соответствующие группы точек на графиках PCA, а корреляция подтверждается значением Мантел-теста (рис. 3).

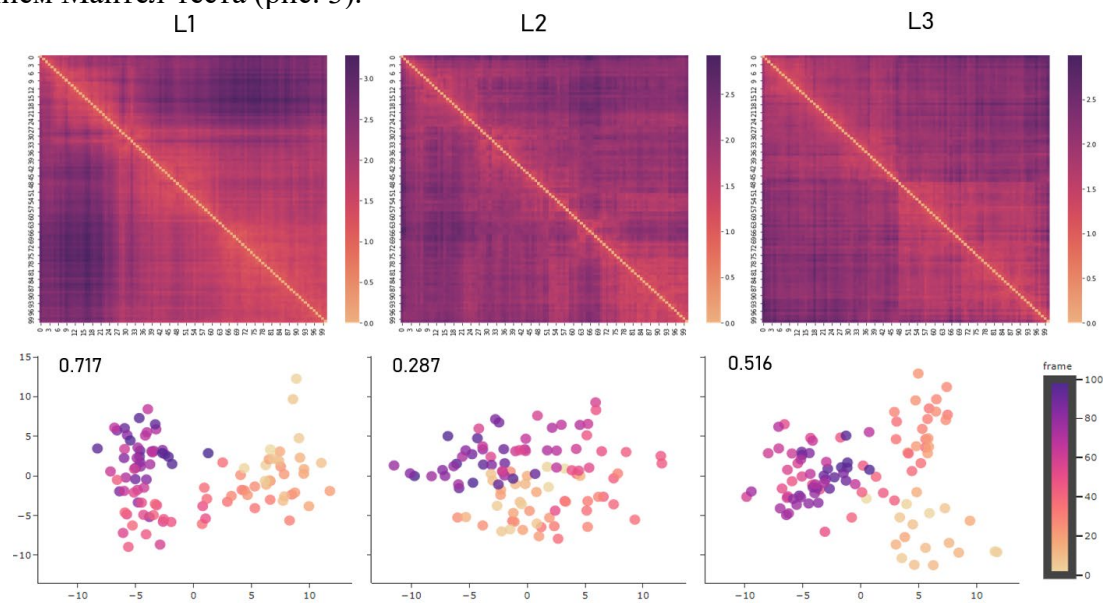


Рис. 3. Сравнение матриц RMSD и структурных дескрипторов для траекторий VDR-VD3-PGC1a (p-value при расчете Мантел-теста = 0,0001)

Дополнительно мы разбили все 18 траекторий (суммарно 1800 кадров) по коактиваторам и лигандам. Так, на рисунке 4 показана группировка кадров по коактиваторам и выделение различных лигандов цветом. В результате мы видим, что в обоих случаях явно выделен кластер структур, имеющих в своем составе лиганд GEM, однако, в структурах с коактиватором PGC1a, те из них, которые содержат лиганд TX527, также выделяются в отдельную группу.

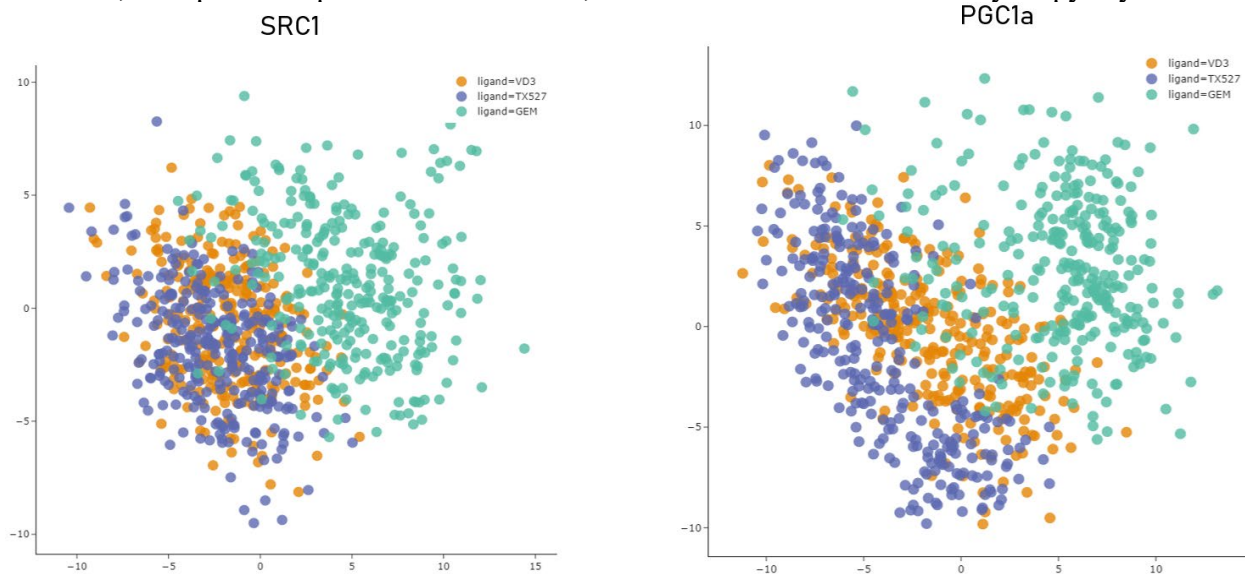


Рис. 4. Сравнение распределения структур в зависимости от различных коактиваторов и лигандов

Это также подтверждается и использованием kNN-классификатора на этих данных – в случае PGC1a точность классификации немного возрастает – 75,8% против 70,3% для SRC1. Такое разделение свидетельствует в пользу одной из гипотез об аллостерической регуляции VDR, а именно связи TX527 с последующим связыванием PGC1a.

Таким образом, в ходе работы мы применили разработанные нами дескрипторы к двум типам трехмерных структур рецептора витамина D. В обоих случаях была показана возможность находить структуры-выбросы с отличающимися регионами, была исследована возможность использования структурных дескрипторов в качестве более быстрой альтернативы расчета RMSD, а также способность дескрипторов выделять также различия в структурах при присоединении к ним различных молекул.

Литература

1. Lindstrom A., Pettersson F., Linusson A. Quantitative protein descriptors for secondary structure characterization and protein classification // *Chemometrics and Intelligent Laboratory Systems*. – 2009. – V.95. – Pp. 74–85.
2. Jain P., Hirst J. D. Automatic structure classification of small proteins using random forest // *BMC Bioinformatics*. – 2010. – V.11(364). – Pp. 1–14.
3. Molnar F. Structural considerations of vitamin D signaling // *Frontiers in Physiology*. – 2014. – V.5(191). – Pp. 1–22.
4. Ritchie D.W. et al. Fast Protein Structure Alignment using Gaussian Overlap Scoring of Backbone Peptide Fragment Similarity // *Bioinformatics* – 2012. – №28(24). – Pp. 3274-3281.

УДК 004.891.2

СРАВНЕНИЕ МЕТОДОВ ВЫДЕЛЕНИЯ ПРИЗНАКОВ ИЗ МЕТАГЕНОМОВ ДЛЯ РЕШЕНИЯ ЗАДАЧИ КЛАССИФИКАЦИИ ВОСПАЛИТЕЛЬНЫХ ЗАБОЛЕВАНИЙ КИШЕЧНИКА

Попов В.В.¹ (студент), Иванов А.Б.² (аспирант)

Научный руководитель – канд. техн. наук Ульянов В.И.²

¹*Санкт-Петербургский государственный университет*

²*Университет ИТМО*

e-mail: popov99vlad@gmail.com

Данная работа посвящена сравнению методов выделения признаков из данных метагеномного секвенирования для решения задачи классификации групп образцов микробиоты кишечника. В ходе работы протестированы подходы к выделению признаков, основанные на поиске уникальных коротких подпоследовательностей генома (к-меров), либо более длинных участков – прочтений. Для решения задачи классификации использованы различные методы машинного обучения, а также программы для выделения признаков.

Ключевые слова: метагеномика, к-меры, воспалительные заболевания кишечника, выделение признаков, машинное обучение.

Наследственная информация каждого организма зашифрована в геномах. Геномика – раздел генетики, посвященный изучению геномов живых организмов. Геномы определяют с помощью секвенирования пар оснований ДНК. Ярким примером является «Human Genome Project», в ходе которого были секвенированы 3 миллиарда пар оснований. Но остаются ещё тысячи организмов всех трёх доменов, геномы которых не были секвенированы. Метагеномика же занимается изучением геномов, полученных из разных организмов, обитающих в общей среде. С её помощью может быть получено представление не только о структуре сообщества, но и о его функциональных возможностях. Примерами интересующих исследователей сред выступают почва, морская вода, а также желудочно-кишечные тракты позвоночных.

Особенно важным и актуальным является изучение кишечника человека. Микробиота кишечника участвует в регуляции обмена веществ и изменяется при развитии заболеваний. Однако, механизм возникновения и влияния на структуру микробиоты многих заболеваний остаётся неисследованным.

Воспалительные заболевания кишечника (ВЗК) – это хронические заболевания, при которых поражается слизистая оболочка части желудочно-кишечного тракта (ЖКТ). ВЗК в основном представлены болезнью Крона и язвенным колитом. Болезнь Крона может поражать все отделы ЖКТ, характеризуется образованием язв и рубцов на стенках кишки. Язвенный колит в основном поражает слизистую оболочку толстой кишки. ВЗК – это комплексные заболевания с разнообразными симптомами, что затрудняет раннюю диагностику, которая крайне важна для успешного лечения. Возникает задача поиска методов для диагностики и понимания причин развития таких заболеваний.

Целью работы является поиск лучших методов для анализа данных микробиоты для диагностирования заболеваний. В ходе работы будут использованы такие методы для обработки метагеномных данных, как MetaFast [1], Commet [2] и КМС-3 [3], а также программа SHAP [4].

MetaFast для выделения метагеномных признаков

Программа MetaFast использует k -мерный подход для выделения компонент. Принцип работы метода следующий:

1. Сначала для каждого образца подсчитываются k -меры: получаются бинарные файлы для дальнейшей обработки.
2. Извлекаются k -меры, присутствующие хотя бы в некотором заданном числе образцов из одной группы и полностью отсутствующие в других (уникальные).
3. Строится граф де Брейна вокруг выделенных k -мер. Далее он разбивается на компоненты, которые используются как характерные для определенной категории.
4. Для каждого индивида считается покрытие каждой компоненты. Для этого вычисляется доля символов из подстрок в компоненте, для которых существует k -мер из индивида, совпадающий с соответствующей подстрокой в компоненте. Таким образом, получается информация о представленности компонент, которая может быть использована для сравнения образцов между собой (с помощью машинного обучения).

Проведено исследование методов на данных метагеномах пациентов с воспалительными заболеваниями кишечника. Для тестирования и сравнения описанных ранее методов были взяты четыре набора секвенирования метагеномных данных: Schirmer, He, Franzosa и HMP2. Образцы в наборах были секвенированы на платформе Illumina HiSeq, были получены парноконцевые прочтения длиной 100 нуклеотидов. В каждой выборке представлены три группы – пациенты с язвенным колитом (ЯК), болезнью Крона (БК) и контрольная группа. В таблице 1 представлены наборы данных и соответствующие числа наблюдений в группах.

Таблица 1. Число наблюдений в группах в используемых наборах метагеномных данных

	Болезнь Крона	Язвенный колит	Контроль	Всего
Schirmer	23	17	14	54
He	63	0	53	116
Franzosa	88	76	56	220
HMP2	65	38	27	130
всего	239	131	150	520

С помощью программы MetaFast для каждой группы были подсчитаны k -меры, присутствующие хотя бы в n наблюдений из группы. Затем на основе этих k -меров построен граф де Брейна, из которого выделены компоненты. В таблице 2 представлено число компонент M для каждой из групп в зависимости от n .

Таблица 2. Число компонент для разных групп набора данных Schirmer, красным цветом выделены компоненты, использованные для дальнейшего анализа

n / M	Язвенный колит	Болезнь Крона	Контроль
4	22834	120961	55285
5	3743	48969	13741
6	508	23832	3764
7		11263	

Для каждой из групп было выбрано определенное число компонент, выбрана сбалансированная комбинация. Далее с помощью MetaFast для каждого наблюдения было подсчитано покрытие компонент k -мерами, получено 54 векторов покрытия. На рисунке 1, а представлены индивиды в плоскости первых двух главных компонент, группы выделены цветом. Как видим, даже по двум компонентам (суммарная информативность 34%) группы хорошо отделяются. Затем посчитано покрытие полученных компонент индивидами из других наборов данных (He, Franzosa и HMP2). Значения были спроектированы (рис. 1, б) на полученную плоскость. Заметно, что классы гораздо хуже отделяются друг от друга.

Вектора покрытия индивидов Schirmer и их метки были взяты в качестве обучающей выборки. Затем была предпринята попытка предсказания классов индивидов из других наборов данных с помощью методов машинного обучения XGboost и Random Forest. Точность предсказания представлена в таблице 3. Метод машинного обучения Random Forest показал более высокую точность, но в целом качество для трехклассовой классификации невысокое.

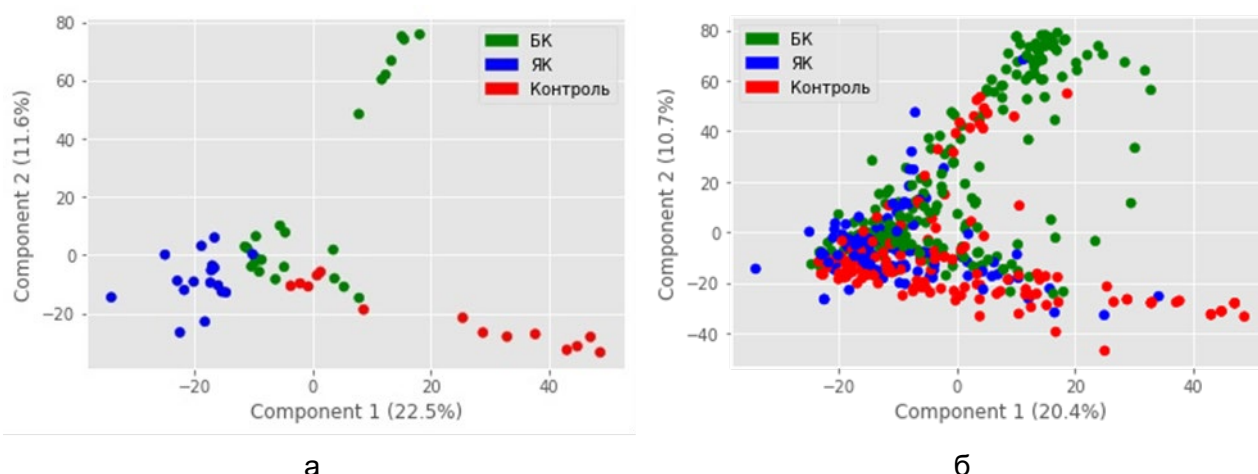


Рис. 1. Распределение по плоскости Schirmer: а) индивиды Schirmer в плоскости первых двух главных компонент по Schirmer; б) индивиды всех наборов данных в той же плоскости

Таблица 3. Точность классификации в зависимости от метода машинного обучения

Точность, %	HMP2	Franzosa	He
Random Forest	62,4 ± 1,3	55 ± 0,5	57,4 ± 1,6
XGboost	55	50	50,9

Выделение информативных компонент с помощью SHAP

Компонент (столбцов в матрице обучающей выборке) получилось много – 47837. Эти данные зашумлены, многие признаки малоинформативны и лишь замедляют процесс машинного обучения. В связи с этим было решено воспользоваться одним из инструментов для выделения наиболее информативных признаков – программой SHAP [4].

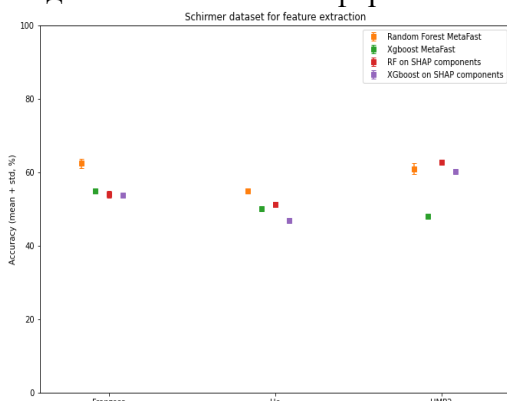


Рис. 2. Сравнение точности предсказания по всем признакам и признакам по SHAP

Методы Commet и КМС-3 для выделения метагеномных признаков

Commet [2] – метод для сравнения метагеномов, представленных в виде файла с прочтениями. Он производит поиск прочтений из одного метагенома, которые покрыты k-мерами из другого метагенома хотя бы заданное число раз. Программа Commet была запущена на небольших тестовых данных, но при запуске на данных Franzosa возникли проблемы с нехваткой оперативной памяти (в размере 256 Гб). Затем было решено взять меньший набор Schirmer, однако, завершение вычислений заняло бы несколько недель. В исследовании [5] было проведено сравнение методов Commet и MetaFast и показана высокая степень сходства результатов на используемых данных, поэтому было принято решение приостановить подсчёты.

В программе MetaFast заложено ограничение – возможная длина k-меров ограничена и не может быть больше 31. Сделано предположение, что построенные на более длинных k-мерах компоненты позволят улучшить точность предсказания. Для проверки этого предположения была использована программа КМС-3 [3], которая позволяет считать k-меры длиной более 31.

На ранее описанном наборе данных Schirmer было подсчитано число уникальных (отсутствующих в других группах) k -меров, присутствующих хотя бы в n образцах из группы. На рисунке 3 показана число уникальных k -меров в каждой из групп в зависимости от длины k . Были рассмотрены длины $k = 31, 41, 51$.

Для каждой из групп из набора данных Schirmer произведен следующий процесс получения признаков:

- выделены уникальные k -меры с помощью КМС-3, причем отброшены те, что встретились лишь в одном файле;
- выполнен поиск прочтений, содержащих уникальные k -меры во всех образцах;
- полученные прочтения собраны в контиги с помощью программы SPAdes [6];
- контиги поданы на вход программе MetaFast, и для всех образцов получены вектора покрытия;
- для предсказания были подсчитаны вектора покрытия полученных контигов и для всех других наборов данных. Далее для каждой длины k была обучена модель Random Forest и посчитана точность предсказания на He, Franzosa и HMP2. Сравнение точности предсказания для различных значений k представлено на рисунке 4.

Как видим, точность предсказания с увеличением длины k -меров не растёт, а значит, наше предположение не подтвердилось.

Выводы

- задача классификации метагеномных образцов для диагностирования воспалительных заболеваний кишечника является актуальной и сложной — точность существующих методов остается невысокой;
- программа Commet неприменима для больших данных;
- предсказания на длинных k -мерах не дали значимого улучшения качества;
- использование информативных компонент по SHAP позволяет уменьшить объем данных без значительной потери информации.

Показана перспективность развития программы MetaFast. Однако, в текущей реализации алгоритма выделения уникальных k -меров введен полный запрет на присутствие k -меров в образцах других категорий, что является недостатком программы. Поскольку в образцах могут быть незначительные загрязнения, возникает идея замены этого ограничения на некоторые другие условия, например, с использованием статистических тестов. Полученные специфичные k -меры могут дать лучшее представление о каждом из классов и повысить качество предсказания. Одним из возможных направлений дальнейшего развития является реализация данной идеи в качестве модификации программы MetaFast.

Литература

1. Ulyantsev V.I., Kazakov S.V., Dubinkina V.B. et al. MetaFast: fast reference-free graph-based comparison of shotgun metagenomic data // Bioinformatics. – 2016. – 06. – Vol.32. – №18. – Pp. 2760–2767.
2. Maillat N., Collet G., Vannier T. et al. Commet: Comparing and combining multiple metagenomic datasets // 2014 IEEE International Conference on Bioinformatics and Biomedicine (BIBM). – 2014. – Pp. 94–98.

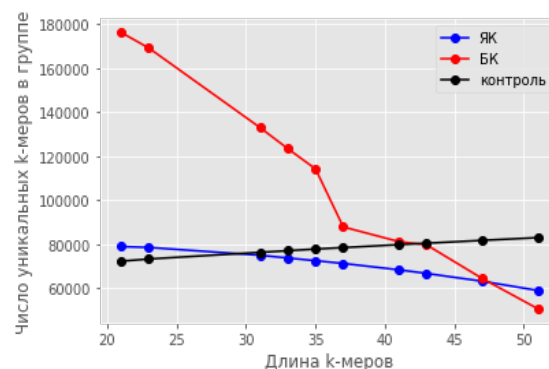


Рис. 3. Зависимость числа уникальных k -мер в различных группах от длины k

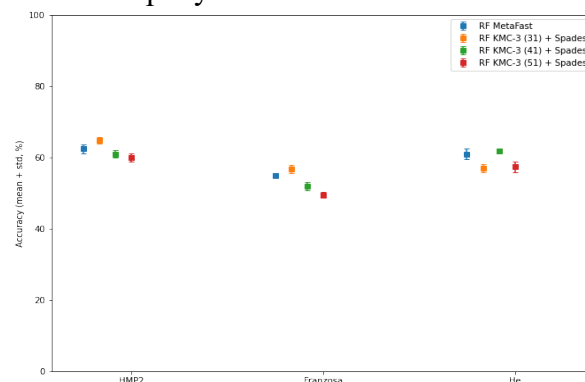


Рис. 4. Сравнение точности предсказания методом Random Forest при различных способах выделения признаков

3. Kokot M., Dlugosz M., Deorowicz S. KMC 3: counting and manipulating k-mer statistics // *Bioinformatics*. – 2017. – 05. – Vol.33. – №17. – Pp. 2759–2761.
4. Lundberg S.M., Lee S.-I. A unified approach to interpreting model prediction // *Advances in Neural Information Processing Systems* 30. – 2017. – Pp. 4765–4774.
5. Galand P.E., Pereira O., Hochart C. et al. A strong link between marine microbial community composition and function challenge the idea of functional redundancy // *The ISME Journal*. – 2018. – Jun. – Vol.12. – №10. – Pp. 2470–2478.
6. Bankevich A., Nurk S., Antipov D. et al. SPAdes: A new genome assembly algorithm and its applications to single-cell sequencing // *Journal of Computational Biology*. – 2012. – May. – Vol.19. – №5. – Pp. 455–477.

УДК 004.934

АНАЛИТИЧЕСКИЙ ОБЗОР МЕТОДОВ ИЗВЛЕЧЕНИЯ ТЕКСТОВЫХ ТРАНСКРИПЦИЙ ИЗ РЕЧЕВЫХ ВЫСКАЗЫВАНИЙ

Ролинский С.О.¹ (студент), Двойникова А.А.^{1,2} (аспирант)

Научный руководитель – д-р техн. наук, доцент Карпов А.А.^{1,2}

¹*Университет ИТМО*

²*Санкт-Петербургский Федеральный исследовательский центр Российской академии наук*

e-mail: sergrol@gmail.com

В работе рассмотрены основные существующие подходы к автоматическому распознаванию речи, а также проводится сравнительный анализ открытых компьютерных систем распознавания речи. Для экспериментальных исследований эффективности работы рассматриваемых систем используется речевой корпус LibriSpeech.

Ключевые слова: распознавание речи, ASR, нейронные сети, аудиомодальность, WER.

В настоящее время как в сети Интернет, так и в электронных архивах организаций, существует большое количество наборов речевых данных по различным тематикам, представленных в текстовом виде. Если некоторое время назад, когда технологии еще не были достаточно развиты, сбор подобных данных производился преимущественно вручную, то сейчас современный мир стремится к автоматизации сбора и генерации данных. В настоящее время существуют различные методы и системы, позволяющие извлекать из речевых данных текстовые транскрипции автоматически. Автоматическое распознавание речи (англ. Automatic Speech Recognition, ASR) применяется не только для сбора данных, но и для естественного взаимодействия человека с машиной. Так, распознавание речевых сигналов может применяться во многих системах: в голосовых помощниках, приложениях на смартфонах или в системах «Умный дом». Отсюда следует, что тема автоматического распознавания речи для различных языков в настоящее время является актуальной.

Аналитический обзор методов распознавания речи. Несмотря на существование различных подходов и методов ASR, этапы анализа речевых данных для распознавания текстовых транскрипций универсальны для всех систем. Данные этапы показаны на рисунке ниже. Первым этапом является оценка качества сигнала. На этом этапе определяется уровень шума и помех, пока аудиосигнал проходит через модуль, отделяющий полезный речевой сигнал от тишины и шума (англ. Voice Activity Detection, VAD). На следующем этапе осуществляется параметрическое представление речевого сигнала, в ходе которого каждый сегмент речи преобразуется в вектор признаков. Наконец, последний и самый главный этап – декодер сигнала, который определяет последовательность слов с наибольшей вероятностью и создает итоговый результат распознавания. Декодер использует акустическую модель – она строится на речевых данных заранее и описывает произнесение звука в речи; также в ней рассчитываются параметры речи, необходимые для распознавания. Также декодером используется языковая модель, которая содержит большой объем информации о правилах и взаимодействиях во фразах языка, о частях речи и о наиболее статистически вероятных последовательностях слов.

В области распознавания речи выделяются 3 основных подхода: марковские цепи, динамическое программирование и искусственные нейронные сети [1]. Скрытые марковские модели (СММ) строятся для каждой фонемы, после чего они объединяются в цепочки для

построения слов, фраз и предложений. Для обучения сначала проводится спектральный анализ речи, затем последовательность спектральных векторов сравнивается с другими максимально похожими векторами, синтезированными на основе словаря. Например, в работе [2] представлена ASR на основе СММ с учетом фонемных наборов биграмм. Динамическое программирование проводит сравнение фрагмента речи с заранее записанным эталоном, анализируя участки, соответствующие тем или иным звукам и суммируя эти расстояния с некоторыми коэффициентами. Задача сводится к поиску оптимального нелинейного согласования двух отрезков речи. Самый используемый из подобных алгоритмов – алгоритм динамического выравнивания времени (англ. Dynamic Time Warping, DTW) [3]. Искусственные нейронные сети представляют собой математическую модель, построенную по принципу организации и функционирования человеческого мозга. Они состоят из большого количества нейронов, взаимосвязанных между собой. Сначала через сеть пропускается обучающая выборка, а затем на её основе обученные нейроны осуществляют распознавание новых данных. На вход нейронной сети, как правило, подаются векторы признаков речевых сигналов, а на выходе получаются вероятности возможных слов. Этот метод может применяться как для акустического, так и для языкового моделирования. Нейронные сети наиболее часто используются современными учеными для автоматического распознавания речи [4].



Рисунок. Общая архитектура системы автоматического распознавания речи

Нейронные сети также часто применяются в интегральных (end-to-end) системах автоматического распознавания речи. В отличие от многих других методов, представляющих собой множество объединённых в одну рабочую цепь компонентов, обучение интегральных моделей выполняется так, что одна модель выдаёт результат распознавания в одиночку, без других компонент. End-to-end модели, как правило, проще в реализации, демонстрируют лучшую производительность и требуют меньший вычислительный объём для использования. Наиболее распространёнными интегральными моделями обычно служат глубокие нейронные сети [5].

Важно заметить, что для большинства людей, не обладающих углублёнными знаниями в области речевых технологий, реализация вышеописанных методов распознавания речи является трудозатратным процессом, требующим вычислительных ресурсов. Однако, в современном мире также существует большое количество автоматических систем распознавания речи в открытом доступе, которые может использовать для своей работы любой человек. Ниже приводится подробное описание наиболее популярных систем автоматического распознавания речи.

Yandex SpeechKit [6] – система распознавания речи от компании Яндекс, которая может применяться для создания голосовых помощников, автоматизации контакт-центров, контроля качества сервиса и других задач. Разработчики утверждают, что, помимо многих своих других функций, система способна извлекать текстовые транскрипции из любой аудио- и видеозаписи за достаточно короткое время. Система способна обрабатывать текст после распознавания речи,

например, расставляя некоторые знаки препинания или преобразуя числительные в цифры. Всего доступно 3 вида распознавания: потоковое в режиме реального времени, синхронное для коротких записанных одноканальных файлов и асинхронное для длинных многоканальных файлов. Аудиосигнал подается в формате LPCM (Linear Pulse-Code Modulation), с частотой дискретизации 8, 16 или 48 кГц и разрядностью квантования 16 бит. В основе системы распознавания речи Яндекса лежит глубокая нейронная сеть, которая постоянно дообучается, используя массивы данных из приложений и сервисов Яндекса. Система может работать с русским, английским, турецким и казахским языками.

Google Speech Recognition [7] – система распознавания речи от компании Google. Она используется в голосовых помощниках, для аутентификации клиентов по голосу, в голосовом обслуживании и т.д. Система, как и в Yandex SpeechKit, может работать в 3 режимах распознавания речи – потоковом, синхронном и асинхронном. У сервиса распознавания речи Google существует огромное количество отличительных преимуществ: большой словарь с поддержкой более 125 языков, в том числе русского, адаптация речи под определенные тематики, многоканальность, фильтрация шума. Система Speech Recognition использует интегральные глубокие нейронные сети, применяя принципы нейросетевой темпоральной классификации (англ. Connectionist Temporal Classification — CTC) – выделение в аудиопотоке конкретных фонем и расставление их по порядку произнесения.

Kaldi [8] – фреймворк с открытым исходным кодом, созданный специально для распознавания аудиосигналов и речи. Он работает на основе скрытых марковских моделей, моделей смеси гауссовских распределений (англ. GMM – Gaussian Mixture Model), дискриминативного обучения, а также глубоких нейронных сетей с задержкой по времени (англ. Time-Delay Neural Network, TDNN) свёрточных (англ. Convolutional Neural Network, CNN) и рекуррентных (англ. Recurrent Neural Network, RNN) нейронных сетей. Считается, что Kaldi предлагает пользователю наиболее богатый выбор алгоритмов распознавания. Помимо извлечения признаков MFCC для TDNN, Kaldi также включает в себя конечные преобразователи состояния (англ. Finite-State Transducer, FST) для языкового моделирования – они помогают в создании языковых моделей и вероятностного предсказания следующего слова.

CMU Sphinx [9] – библиотека для распознавания речи, написанная на языке программирования Java. Помимо автоматического распознавания речи, она также способна идентифицировать дикторов по речи и находить ключевые слова. Инструменты Sphinx были специально разработаны для низкоресурсных платформ. Для распознаваний на небольших данных используется пакет rocketsphinx. Отличие CMU Sphinx от большинства других (облачных) систем в том, что он может работать в офлайн-режиме. В этой системе встроена только одна базовая модель английского языка, однако, при желании можно свободно использовать модели из других языков.

Vosk [10] – сервис для автоматического распознавания речи, который, как и rocketsphinx, способен работать в автономном режиме при наличии языковой модели. В Vosk поддерживается около 20 языков и их диалектов. Основными плюсами этого инструмента являются легкий вес моделей при достаточно высокой точности, удобная поддержка потоковой записи и реконфигурации словаря. В своей архитектуре Vosk использует проекты грамматики Kaldi. Он используется для создания чат-ботов, виртуальных ассистентов, субтитров и т.д.

Silero [11] – сервис для распознавания речи, написанный командой программистов из России. Разработчики позиционируют Silero как наиболее подходящий инструмент для извлечения текстовых транскрипций из аудиофайлов небольшого объёма. Являясь простыми в использовании, модели Silero, по словам разработчиков, имеют схожую с технологиями Google точность. Доступные в данный момент языки: русский, английский, испанский и немецкий.

Voco [12] – программное обеспечение для распознавания речи, разработанное ООО «Центр речевых технологий». Оно не является открытым, но может работать без подключения к сети, достаточно только скачать модели и словари. Встроенные команды значительно упрощают процесс распознавания. Программа может распознавать только русскую речь, и содержит 2 основных словаря – юридический и финансовый.

Для проведения экспериментальных исследований были выбраны все вышеописанные системы ASR, кроме Yandex SpeechKit, т.к. у неё платный доступ; Kaldi, т.к. она плохо совместима с Windows; и Vosco, т.к. она анализирует только русскоязычную речь. Эффективность данных систем была проанализирована на корпусе аудиоданных LibriSpeech [13]. Этот набор содержит 1000 часов аудиокниг, взятых из состава LibriVox – некоммерческого проекта по созданию аудиокниг. Речь записана на английском языке с частотой 16кГц, и разделена на «чистую» и «другую» речь, в зависимости от того, какой уровень ошибок распознавания слов (англ. Word Error Rate, WER) для каждого диктора выдавала внутренняя система тестирования. Это значит, что в классе «другой» речи находятся записи дикторов с плохим или неудачным произношением. LibriSpeech также содержит встроенные языковые модели с использованием трех и четырех грамм.

Для распознавания были выбраны подвыборка случайным образом из класса «другой» речи, общим объемом 1,4 ч. Каждая описанная выше система автоматического распознавания речи при извлечении текстовых транскрипций из аудиоданных корпуса LibriSpeech выдавала ошибки в текстовых данных. В таблице ниже представлен сравнительный анализ систем распознавания речи: их точность извлечения текстовых транскрипций, выраженная показателем WER, и время работы.

Таблица. Сравнительный анализ систем ASR

Система ASR	WER, %	Показатель реального времени
Google Speech Recognition	15,9	0,155
Vosk	4,8	0,185
Silero	14,9	0,174
CMU Sphinx	74,3	0,661

Из таблицы видно, что вероятность ошибок распознавания слов у большинства рассмотренных систем является низкой. Значения WER показывают, что основную часть всех слов системы ASR распознают верно. Наиболее высокую точность распознавания англоязычной речи показала система Vosk, показатель WER = 4,8%. Это можно объяснить тем, что некоторые модели этой системы обучались именно на корпусах аудиокниг. Гораздо более низкую точность распознавания речи показала система Sphinx, это обусловлено возможной непригодностью базовой модели к речи в староанглийском художественном стиле, содержащемся во многих аудиокнигах LibriSpeech – многие из них были написаны в 19 веке и ранее. Среднее время работы систем распознавания составило приблизительно 24 мин для анализа 1 ч 24 мин используемой подвыборки базы данных LibriSpeech.

Исходя из проведенных экспериментальных исследований извлечения текстовых транскрипций из аудиоданных, можно сделать вывод о том, что современные системы распознают речь с более высокой точностью по сравнению с системами, актуальными несколько лет назад. Скорость выполнения работы также достаточно высокая для всех систем, несмотря на большие объемы данных, которые им приходится анализировать: каждый час звуковых данных распознается приблизительно за 20 минут. Также стоит учитывать то, что любая модель машинного обучения (в том числе и нейронные сети) зависит от обучающих данных, что может значительно повлиять на точность распознавания речи. Аналитический обзор методов и систем автоматического распознавания речи показывает, что именно нейронные сети наиболее эффективны и наиболее востребованы в нынешнее время.

Литература

1. Грачев А.М. Статистические подходы к автоматическому распознаванию речи // Вестник Нижегородского университета им. Н.И. Лобачевского – 2016. – №2(2). – С. 376–379.
2. Важенина Д.А., Кипяткова И.С., Марков К., Карпов А.А. Методика выбора фонемного набора для автоматического распознавания речи // Труды СПИИРАН. – 2014. – Т.5. – №36. – С. 92–113.

3. Ковтун В.В. Разработка компьютерной системы распознавания речи // Современные научные исследования и инновации. – 2015. – №1. – Ч.1. [Электронный источник] – Режим доступа: <https://web.snauka.ru/issues/2015/01/45410> (дата обращения: 16.02.22).
4. Мамырбаев О.Ж., Тұрдалыұлы М., Мекебаев Н.О., Тұрдалықызы Т., Шаяхметова А.С. Автоматическое распознавание казахской речи с использованием DNN // Вестник Казахстанско-Британского технического университета – 2019 – Т.16(2) – С. 134–142.
5. Марковников Н.М., Кипяткова И.С. Аналитический обзор интегральных систем распознавания речи // Труды СПИИРАН. – 2018. – №58(3). – С. 77–110.
6. Yandex.Cloud – Распознавание речи. [Электронный ресурс]. – Режим доступа: <https://cloud.yandex.ru/docs/speechkit/stt/> (дата обращения: 19.01.22).
7. GoogleCloud Speech-to-Text: Accurately convert speech into text using an API powered by Google’s AI technologies. [Электронный ресурс]. – Режим доступа: <https://cloud.google.com/speech-to-text> (дата обращения: 19.01.22).
8. Povey D. et al. The Kaldi speech recognition toolkit // IEEE 2011 workshop on automatic speech recognition and understanding. – IEEE Signal Processing Society. – 2011. – 4 p.
9. Huggins-Daines D. et al. Pocketsphinx: A free, real-time continuous speech recognition system for hand-held devices // 2006 IEEE International Conference on Acoustics Speech and Signal Processing Proceedings. – 2006. – Т.1. – 4 p.
10. Vosk – a speech recognition toolkit. [Электронный ресурс]. – Режим доступа: <https://alphacephei.com/vosk/> (дата обращения: 19.01.22).
11. Silero – решаем реальные проблемы с использованием Speech-To-Text, Text-To-Speech, NLP и других технологий машинного обучения. [Электронный ресурс]. – Режим доступа: <https://www.silero.ai/about/> (дата обращения: 16.02.22).
12. VOCO - Windows-приложение для преобразования речи в текст знаков. [Электронный ресурс]. – Режим доступа: <https://www.speechpro.ru/product/programmy-dlya-raspoznavaniya-rechi-v-tekst/voco> (дата обращения: 16.02.22).
13. Panayotov V. et al. Librispeech: an asr corpus based on public domain audio books //2015 IEEE international conference on acoustics, speech and signal processing (ICASSP). – 2015. – С. 5206–5210.

УДК 004.934.8

РАЗРАБОТКА МЕТОДИКИ ИССЛЕДОВАНИЯ ОШИБОК ДИАРИЗАЦИИ ПО ЛЕКСИКО-ГРАММАТИЧЕСКИМ ПАРАМЕТРАМ

Рудалева Е.А.¹ (студент)

Научный руководитель – в.н.с. Попов Д.В.²

¹*Университет ИТМО*

²*ООО «ЦРТ-Инновации»*

e-mail: rudaleva.e@gmail.com, ppdmitry@gmail.com

В настоящей работе применяется метод составления лексических словарей для идентификации дикторов в рамках разработки методики детектирования ошибок диаризации с извлечением лингвистической информации, а именно прагматико-сематической и синтаксической областей, для которых были использованы подходы дискурсивного анализа и грамматики конструкций. Применение описываемого метода может служить независимой метрикой в проверке качества диаризации через анализ дискурсивных и лексических аспектов словоупотребления.

Ключевые слова: диаризация, определение диктора, лексические параметры, грамматика конструкций, дискурсивный анализ.

В настоящее время в области речевых технологий активно развивается сфера детектирования дикторов. Большое внимание в исследованиях диаризации уделяется разработке методов для повышения ее качества, в которых освещаются проблемы шума, неречевых аудиосегментов, наложений речи, а также влияния лингвистической информации. В работе широко освещается тезис о первичности последней для решения задач в области речевых технологий.

Говорящий, он же диктор, в речи — участник коммуникативной ситуации, вписанной в определенные дискурсивные рамки и отличающейся уникальными конвенциональными лексическими особенностями, которые напрямую связаны с языковой средой. Ввиду неразрывности лингвистики с речевыми технологиями, в данной работе процесс диаризации осмысливается через функциональное направление лингвистики и анализ дискурса.

В транскриптах аудиозаписей содержится большое количество лингвистической информации, маркирующей уникальных дикторов в сегментах речи. Главной целью работы является разработка методики использования лингвистической и дискурсивной информации, вложенных в записи диалога, чтобы на ее основе проводить различие между дикторами в регистре формальной речи.

В качестве теоретической базы исследования были выделены два направления лингвистики, которые охватывают несколько уровней языка (текстовой, синтаксический, прагматический и семантический): грамматика конструкций и дискурсивный анализ.

Грамматика конструкций (далее СxG) рассматривает язык как совокупность конструкций — языковых выражений, у которого есть аспект плана выражения (форма) и плана содержания, и которое не композиционально — не производимо из сложения его составляющих элементов [1]. СxG предполагает, что онтологический статус конструкции — комплекс из всех языковых уровней: от просодического до прагматического. Конструкции обладают вариативностью по длине от одной морфемы до словосочетания или предложения.

Основное положение СxG о моделируемости языка через конструкции является необходимым внутри данного исследования для формирования методики выделения лексических паттернов, свойственных уникальным дикторам ввиду формализованности и одновременной вариативности конструкций.

Дискурсивный анализ принимается как первичный элемент в исследованиях из области речевых технологий. Он включает в себя рассмотрение форм, хорошо узнаваемых носителями языка, и предполагает конвенциональность и повторяемость некоторых символических форм, внутри одного или разных текстов устного или письменного модуса одного или разных авторов [2, 3]. Последний пункт имеет связь с теорией СxG, изложенной выше, поскольку допускает структурную фрагментарность в порождении и восприятии речи.

Создание лексической базы, привязанной к уникальному диктору, тесно сопряжено с регистром. Под регистром в данной работе понимается план содержания языка; иными словами—язык в действии [3, с. 286], в конкретной коммуникативной ситуации.

В данной работе в качестве такой таксономии была принята эксплицитная аналитическая рамка регистров, предложенная в [3]: адресант (количество), адресат (количество), домен (частный, рабочий, академический), общее место-время, модус (устный, письменный), порождение (спонтанное, проверенное, подготовленное), фактичность (основанное на фактах, спекулятивное, воображаемое, смешанное), цели: убеждение, информация, развлечение, самовыражение (высокий, средний, низкий уровень), уровень (специальный, популярный), объект (сфера).

Разработанная Байбером [3] аналитическая рамка сочетает в себе функциональный (ситуативное языковое использование) и конвенциональный (языковая норма) языковые уровни, которые охватывают как письменный, так и устный модус речи. Особенность обсуждаемой работы также заключается в не дискретном представлении регистра, что позволяет комбинировать дискурсивные параметры и рассматривать их в градуальной плоскости. Это помогает решать задачи неоднородности регистров в их признаках, изменчивости модели коммуникации (смещение или смена регистров) и изменчивости самих говорящих.

Проведенная теоретическая часть составляет основу лингвистического подхода в исследовании, применяемого к диаризации речи. В качестве основных сложностей диаризации можно выделить [4, с. 22]:

1. Различия в домене данных.
2. Наложение речи нескольких дикторов.
3. Взаимодействие с автоматическим распознаванием речи.

Предполагается, что принятый метод составления лексических словарей в рамках лингвистического подхода может способствовать частичному преодолению обозначенных выше проблем [5, 6].

Методика с опорой на таксономию регистров и классификацию конструкций в СxG позволяет, во-первых, определить домен (регистр) данных, построив в дальнейшем словари для дикторов; во-вторых, выделить области без наложений реплик и потенциально конструкции и позиции в предложении, предшествующие возникновению наложений; в-третьих, конструкции в СxG, допуская вариативность, могут предсказать семантико-синтаксическую роль неизвестного, потенциального объекта в речи.

В исследовании были использованы данные внутренней проприетарной базы исполнителя из офиса обслуживания длиной в 826,658 часов, состоящей из 520 записей. Базово сессия обслуживания происходит между двумя дикторами: работником фронт-офиса (далее, оператор) и посетителем (далее, клиент). Число операторов строго определено, в то время как число клиентов может варьироваться от записи к записи (в среднем 2 за запись).

С использованием аналитической рамки регистров [3] было произведено детальное представление регистра сессий офиса обслуживания, изложенного далее:

1. *Коммуникативные характеристики участников*: адресант – клиент, адресат – оператор. Вместе с тем, оператор первым вступает в диалог, например, с приветствия. Затем

- обращается с просьбой клиент. Часто в записи могут присутствовать третьи лица, не относящиеся содержательно к сессии (нецелевые дикторы).
2. *Отношения между адресантом и адресатом*: адресант координирует диалог: задает вопросы, выдвигает требования; следовательно, оператор обладает большей коммуникативной властью как специалист. Интерактивность между дикторами средняя: взаимодействие обычно происходит только в рамках клиентского запроса, отталкивается от него.
 3. *Ситуация общения*: Место коммуникации – общественное, офис обслуживания. Дикторы преимущественно находятся в одном пространственно-временном континууме.
 4. *Канал связи*: преимущественно устный модус, преходящий диалог: общение лицом к лицу, телефонный разговор.
 5. *Отношение участников к тексту*: адресант – эмоционально отстранен от текста, содержит фрагменты запланированной, шаблонной лексики; адресант – чаще более эмоционально вовлечен, спонтанная речь по запросу.
 6. *Цели, намерения, задачи*: коммуникативная ситуация основана на фактах; высокая по уровню цель: убедить или продать; средняя по уровню цель: передать информацию; низкая по уровню цель: самовыражение, развлечение.
 7. *Тема/Объект*: специализированный уровень обсуждения, формальный жанр.

Представление регистра, в котором находятся диалоги анализируемой базы данных (БД), позволило четко разграничить роли оператора и клиента, выделив явные дискурсивные различия. На их основе, а также с учетом специфики закрытой проприетарной БД офиса обслуживания и эмпирической проверки записей были составлены словари для операторской и клиентской лексики, включающие в себя лексическую и семантико-синтаксическую информацию.

Словари уникальной операторской и клиентской лексики были отобраны по количественным характеристикам токенов в исследуемой базе данных. Лексема считалась включенной в словарь, если количество ее употреблений в корпусе больше 30, а доля вхождений в канал диктора — не менее 0,9 от общей подвыборки по данной лексеме.

Проверка лексем на исключительность использования одним диктором осуществлялась эмпирически с фиксированием количественных данных по количеству записей и количеству употреблений в канале клиента и оператора по результату диаризации и по истинной роли, определяемой при прослушивании. Также была рассчитана доля пропущенных вхождений в реплики искомого диктора (далее *ошибка*).

Для оператора было определено 24 уникальных лексических паттерна и 8 композиционных конструкций (см. ниже) для оператора, для клиента – 10. Такая разница в количестве паттернов связана с дискурсивными отличиями речи обоих дикторов: операторской – более формальной и подготовленной, клиентской – более спонтанной и не универсализированной.

В таблицах 1 и 2 наглядно представлены первые 5 паттернов (конструкций) для каждого из словарей.

Таблица 1. Пример паттернов речи оператора

Паттерн	N записей	N употреблений	Оператор: результат диаризации	Истинный оператор	Ошибка
согласие на обработку данных	161	219	0,909	1	0,091
[за] пределами Российской Федерации	113	149	0,960	1	0,040
Поставить подпись	143	148	0,885	1	0,115
запрос [...] истории	110	132	0,864	0,985	0,121
получение карты	83	103	0,835	0,942	0,107

Таблица 2. Пример паттернов речи клиента

Паттерн	№ записей	№ употреблений	Клиент: результат диаризации	Истинный клиент	Ошибка
девушка	182	348	0,595	0,925	0,330
спасибо [вам] большое	114	155	0,355	1	0,645
алло	117	148	0,311	0,953	0,642
мне сказали	93	129	0,659	0,969	0,310
у меня есть	77	102	0,569	0,990	0,422

При сравнении лексических паттернов оператора и клиента выделяются длина и сложность конструкций, употребляемых обоими дикторами: операторские паттерны лингвистически являются более сложными и наполненными. Ввиду определенной ранее специфики формального делового регистра было проверено также наличие не только шаблонных, сложно наполненных конструкций, но и их сочетаний.

Так, были выявлены *композиционные конструкции*, состоящие из последовательности единичных паттернов, которые сохраняются с учетом вариативности дикторов и коммуникативных ситуаций.

Таблица 3. Композиционные конструкции

Паттерн	№ записей	№ вхождений	Оператор: результат диаризации	Истинный оператор
Подпись фамилия	148	199	0,905	0,899
Подпись фамилия + инициалы	71	83	0,880	0,904
Подпись фамилия инициалы + имя отчество	6	7	1	0,714
Подпись фамилия + сегодняшнее число	32	35	1	1
Подпись фамилия инициалы + сегодняшняя дата	42	46	0,957	0,913

Практическая ценность полученных дикторских словарей проверялась посредством ряда метрик: точности, полноты и F1-значения. Для их расчета были построены *матрицы спутываний* для каждого лексического паттерна (таблица 4).

Таблица 4. Матрицы спутываний и метрики по лексике дикторов

	TP	FP	TN	FN	Точность	Полнота	F1-значение
Оператор	1825	105	43	148	0,946	0,925	0,935
Оператор: Композиционные конструкции	953	86	7	101	0,917	0,904	0,911
Клиент	572	31	73	507	0,949	0,530	0,680

На исследуемой базе данных алгоритм диаризации продемонстрировал высокие значения точности и полноты на выделении лексики оператора, в то время как лексика клиента при высоких значениях точности отличается низкими показателям полноты, что свидетельствует о понижении способности угадывать истинно клиентскую лексику (ср. recall по операторской – 0,925 и клиентской лексике – 0,53).

Проведенный расчет подтверждает возможность использования лексических словарей для БД в качестве независимой метрики для оценки качества диаризации, а также потенциально способствовать улучшению определения дикторов.

Обнаружение разницы проверки дикторов, где определение оператора было сравнительно более успешно, вызывает необходимость применения СxG-подхода для анализа паттернов операторской и клиентской лексики с точки зрения языковой структуры. Так, словари комплексно рассмотрены на синтаксическом, семантическом и прагматическом уровнях.

Разметка на уровне синтаксиса показала преобладание именных конструкций в операторской лексике. Ярким тому примером являются композиционные конструкции (рисунок). С точки зрения СxG, такие типы конструкций, несмотря на свою шаблонность,

являются не заученными, но образующимися в результате операций над более простыми конструкциями [7]. Клиентские лексические паттерны, напротив, отличаются преобладанием глагольных конструкций

Таблица 5. Наиболее частотные синтаксические конструкции операторской и клиентской лексики

Синтаксический паттерн	%
N (A) + [Pr NP (B)]	0,187
N (A)+ N (A')	0,156
N (A) + [Pr NP (B)] + N (A')	0,125
VP (A)	0,3
V (A)+ DP (B)	0,2

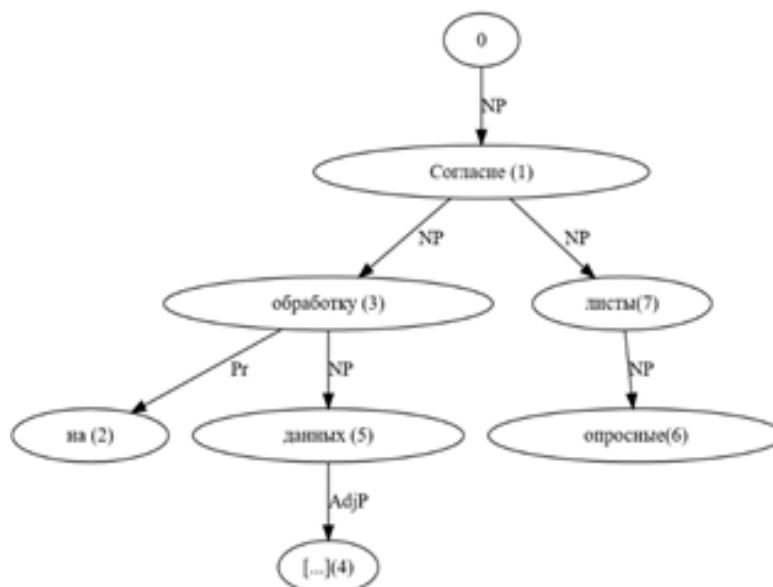


Рисунок. Синтаксическая разметка примера из словаря композиционной лексики оператора

Семантическая и прагматическая составляющая уникальной лексики оператора направлены на указание и побуждение к действию, например, предоставить что-либо (65% директивов) и меньше – на запрос информации (32%). Чаще всего (71,8%) в репликах утверждается истинность информации, то есть говорящий предполагает истинность содержания реплики относительно собеседника (клиента).

Клиентские конструкции чаще предполагают выразительную прагматико-семантическую функцию (40%) констатации фактов (40%) (например, констатация проблемы) вместо побудительной (10%). При этом клиентская уникальная лексика равно наполнена ассертивными и пропозициональными конструкциями (30%), тогда как в операторской наблюдается явное преобладание пропозициональных (71,8% и 15,6%, соответственно), что свойственно формальной и узкой специфике регистра.

Так, обнаруживается разрыв внутри самого определяемого регистра. «Сильная», более строгая его версия относится к уникальной лексике оператора, «мягкая» – к клиентской. Полученные результаты лингвистической разметки позволили выделить синтаксические структуры, которые в дальнейшем возможны к применению в фильтрации по частям речи для задачи определения диктора, а семантико-прагматические паттерны предоставляют материал для автоматической семантической обработки текста для расширения словарей/паттернов дикторов.

Таким образом, настоящая работа является репрезентацией метода, объединяющего лингвистическую теорию и речевые технологии, а изложенные примеры лексических словарей могут быть применимы для независимой оценки качества диаризации, а также служить языковым набором для улучшения ее качества. Польза метода лексических словарей, несмотря

на большую погрешность по сравнению с другими метриками (пословная разметка, сплошная разметка), метод лексических словарей имеет большое преимущество в виде прямой доступности и скорости применения на новом трафике корпуса.

Литература

1. Fillmore C.J., Kay P., O'connor M.C. Regularity and idiomaticity in grammatical constructions: The case of let alone // *Language*. – 1988. – Pp. 501–538.
2. Kempson R. et al. Language as mechanisms for interaction // *Theoretical linguistics*. – 2016. – V.42. – №3-4. – Pp. 203–276.
3. Байбер Д. Аналитическая рамка для изучения языковых регистров // *Социолингвистика и социология языка: хрестоматия*/отв. ред. НБ Вахтин. СПб.: Изд-во Европ. ун-та в Санкт-Петербурге. – 2012. – С. 257–91.
4. Park T.J. et al. A review of speaker diarization: Recent advances with deep learning // *Computer Speech & Language*. – 2022. – V.72. – Pp. 101–317.
5. Canseco-Rodriguez L., Lamel L., Gauvain J.L. Speaker diarization from speech transcripts // *Proc. ICSLP*. – 2004. – V.4. – Pp. 3–7.
6. Bigot B., Fredouille C., Charlet D. Speaker role recognition on tv broadcast documents // *First Workshop on Speech, Language and Audio in Multimedia*. – 2013. – Pp. 66–71.
7. Рахилина Е.В. Лингвистика конструкций. – Общество с ограниченной ответственностью «Издательский центр «Азбуковник»». – 2010. – С. 1–285.

УДК 004.5

ПРОЕКТИРОВАНИЕ ПОЛЬЗОВАТЕЛЬСКОГО ИНТЕРФЕЙСА ПРИЛОЖЕНИЯ С ИСПОЛЬЗОВАНИЕМ АДАПТИВНОЙ ДОПОЛНЕННОЙ РЕАЛЬНОСТИ ДЛЯ ПРОИЗВОДСТВА МЕДИАКОНТЕНТА

Рождественский Д.А.¹ (студент)

Научный руководитель – канд. филос. наук Смолин А.А.¹

¹*Университет ИТМО*

e-mail: dima.christmas98@gmail.com

В данной работе исследуются способы управления дополненной реальностью, применение их на практике, а также описываются первые этапы разработки интерфейса приложения с дополненной реальностью.

Ключевые слова: дополненная реальность, медиа, интерфейс, мобильное приложение.

Среди специалистов по графическим эффектам (VFX- и CGI- художники) существует спрос на ускорение и облегчение продакшена. VFX в своем развитии движется к рендерингу в реальном времени, технология AR – к фотореализму. Эти направления может объединить в себе система, с помощью которой можно изменять реальность в AR, отрисовывать реалистичные объекты и эффекты в реальном времени и максимально сократить этап пост-продакшена. Такой системой является проект «Expose Area», который объединяет в себе мобильное приложение с AR для съемки контента и профессиональный редактор.

Перед началом проектирования интерфейса были изучены статьи, в которых описаны различные методы управления AR: управление взглядом с помощью AR-шлема, передвижение по сцене с помощью сенсорного экрана и передвижение по сцене с помощью передвижения смартфона в пространстве. Для разрабатываемого приложения наиболее удобным и простым способом взаимодействия с AR является перемещение в пространстве со смартфоном и управление объектами AR с помощью сенсорного экрана. Также в других исследованиях было обозначено, что с увеличением экрана устройства снижалась рабочая нагрузка, что доказывает удобство использования AR не только на смартфонах, но и на планшетах.

Проектирование велось в парадигме дизайн-мышления, что позволило собрать как можно больше полезной информации от потенциальных пользователей, выявить их цели, задачи и боли, определить функционал приложения, которые покроет их, и построить первые макеты пользовательского интерфейса с минимально необходимым функционалом, определенным на предыдущих этапах.

Целевая аудитория проекта делится на несколько сегментов:

1. Использующие AR в творческих целях (VFX- и CGI-специалисты, кинооператоры, клипмейкеры, блогеры).
2. Использующие AR в прикладных целях (дизайнеры интерьеров, строители, работники конвейеров, учителя).

На первых этапах проекта было решено сфокусироваться на первом сегменте – пользователи, которые создают медиаконтент в творческих целях. Был проведен опрос среди профессионального сообщества по выделению ключевого сегмента аудитории, который подтвердил спрос среди создателей визуального контента. Также была выявлена готовность к покупке предлагаемой системы. Самый большой сегмент по возрасту – 20 – 25 лет (49,2%).

Среди профессионального сообщества был проведен Customer Development – тестирование идеи или прототипа будущего продукта на востребованность с помощью потенциальных потребителей. Особенно большой спрос проявили VFX-специалисты – они готовы использовать предлагаемую систему, с учётом того, что им будут доступны весь привычный функционал по созданию и редактированию 3D-сцены, привычные паттерны управления интерфейсом. Они также проявили заинтересованность к рендерингу в реальном времени на мобильном устройстве, что сильно сократит объем работы, затрачиваемое время и стоимость производства.



Виталий

25 лет

VFX-художник, клипмейкер

Имеет большой опыт работы, хорошее портфолио в сфере VFX и CGI.

Хорошо знает профессиональное ПО (Blender, After Effects).

Ценит своё время и время заказчика.

“Я быстро строю сцену с нужными мне эффектами и объектами, но рендеринг занимает кучу времени”

“Вот бы снять готовый клип сразу со всеми эффектами и ничего потом не редактировать!”

Цели

- Быть востребованным специалистом
- Делать свою работу быстрее и легче
- Больше зарабатывать
- Создавать качественный контент

Задачи

- Эффектно показывать свое портфолио
- Изучать новые технологии и ПО
- Зарабатывать на продаже своих эффектов и моделей
- Быстро отдавать проекты

Страхи

- Тратить много времени на работу
- Терять время из-за долгого рендеринга
- Не освоить новую технологию
- Не сдать проект вовремя
- Не получать новые проекты

Критерии успеха

- Быстро сдавать проекты
- Качественно делать проекты
- Собрать хорошее портфолио
- Интуитивно осваивать новые технологии и ПО
- Эффективно использовать новые технологии и ПО
- Заработать много денег

Рис. 1. Персона

На основе имеющихся данных был составлен такой артефакт, как персона (рис. 1). Он позволяет отразить общие цели, задачи, страхи и критерии успеха ключевого сегмента аудитории в одном человеке, чтобы при дальнейшем проектировании было проще и точнее выявить нужный функционал, проявив эмпатию к будущему пользователю.

Руководствуясь данными персоны и результатами опросов, можно составить основные пользовательские сценарии:

1. Разместить объект на сцене и настроить его.
 - a. Открыть приложение.
 - b. Поставить кадр.
 - c. Открыть галерею объектов.
 - d. ЕСЛИ объект есть в готовом виде.
 - i. ДА – Выбрать объект.
 - ii. НЕТ – Создать объект, используя примитивы.
 - e. Разместить объект на сцене.
 - f. Выделить объект.
 - g. Выбрать нужную текстуру.
 - h. Поставить источник света.
 - i. Настроить свет.
 - j. Сохранить проект.

2. Записать видеоролик с построенной сценой.
 - a. Иметь построенную сцену.
 - b. Перейти в режим съемки.
 - c. Начать запись видео.
 - d. Завершить запись видео.
3. Сохранить пресет.
 - a. Иметь построенную сцену.
 - b. Сохранить данные о сцене в пресет.
4. Применить пресет.
 - a. Открыть приложение.
 - b. Открыть галерею пресетов.
 - c. ЕСЛИ пресет хранится локально.
 - i. ДА – Выбрать пресет из локального хранилища.
 - ii. НЕТ – Выбрать пресет из онлайн-каталога.
 - d. Поставить кадр.

На основе имеющихся артефактов были построены низкоуровневые макеты интерфейса. Следующим этапом является проектирование высокоуровневых макетов (рис. 2 – 7): создание визуального стиля, применение сетки, доработка функционала, учет состояний элементов интерфейса.

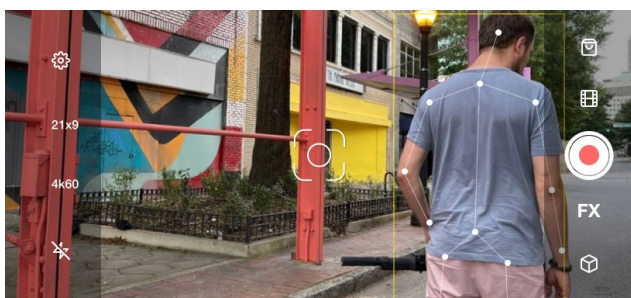


Рис. 2. Высокоуровневые макеты

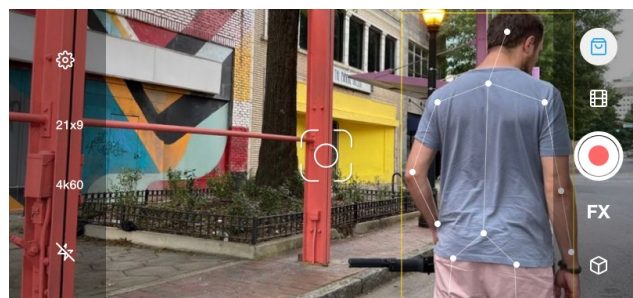


Рис. 3. Высокоуровневые макеты

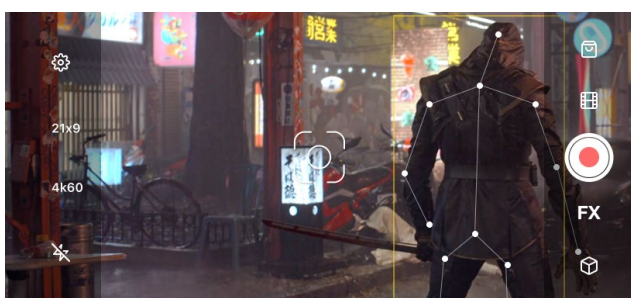


Рис. 4. Высокоуровневые макеты

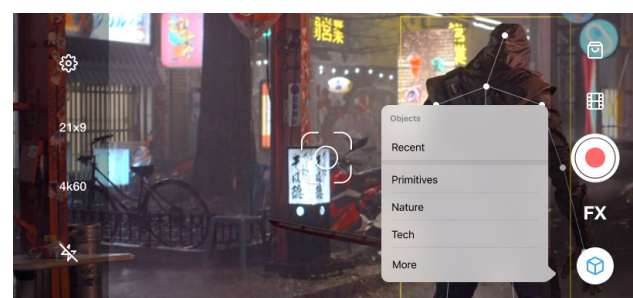


Рис. 5. Высокоуровневые макеты

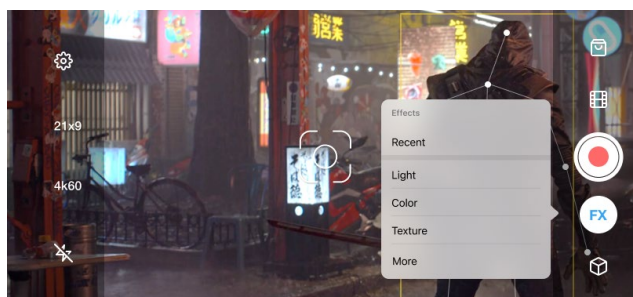


Рис. 6. Высокоуровневые макеты

Основные методы управления интерфейсом – передвижение в пространстве со смартфоном для передвижения по сцене, взаимодействие с объектами сцены с помощью сенсорных жестов – нажатие для выделения объекта, долгое нажатие для редактирования объекта, перетаскивание для перемещения объекта, щипок для изменения размеров объекта. Также доступны стандартные элементы управления камерой смартфона – съемка фото и видео, пауза и остановка записи, приближение.

Был проведен анализ источников на тему дополненной реальности, чтобы сравнить разные методы управления AR и применить лучшие на практике. Некоторые методы легли в основу взаимодействия с приложением. На данный момент спроектированы макеты интерфейса, которые покрывают основные пользовательские сценарии. Планируется продолжать разработку, тестировать прототипы интерфейса на пользователях, проверять существующий функционал и внедрять новые элементы.

Литература

1. Arifin Y., Sastria T.G., Barlian E. User Experience Metric for Augmented Reality Application: A Review // *Procedia Computer Science*. – 2018.– Vol.135. – Pp. 648–656. DOI: 10.1016/j.procs.2018.08.221.
2. Yavuz M., Çorbacıoğlu E., Başoğlu A.N., Daim T.U., Shaygan A. Augmented reality technology adoption: Case of a mobile application in Turkey. *Technology in Society*. – 2021. – Vol.66. – 101598. DOI: 10.1016/j.techsoc.2021.101598.
3. Law E.L.-C., Heintz M. Augmented reality applications for K-12 education: A systematic review from the usability and user experience perspective // *International Journal of Child-Computer Interaction*. – 2021. – Vol.30. – 100321. DOI: 10.1016/j.ijcci.2021.100321.
4. Design Thinking: Study Guide – Nielsen Norman Group. [Электронный ресурс]. – Режим доступа: <https://www.nngroup.com/articles/design-thinking-study-guide/> (дата обращения: 20.01.2022).

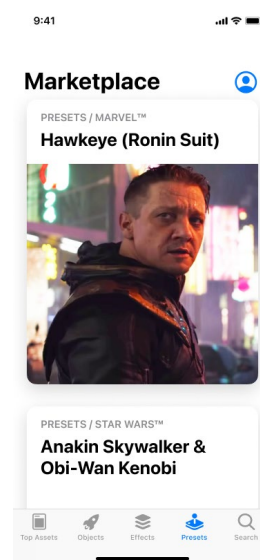


Рис. 7. Высокоуровневые макеты

УДК 535.15

ИССЛЕДОВАНИЕ ОПТИЧЕСКИХ ХАРАКТЕРИСТИК ОБРАЗЦОВ Hg_{0.7}Cd_{0.3}Te, ВЫРАЩЕННЫХ МЕТОДОМ МПЭ

Ружевич М.С.¹ (аспирант)

Научный руководитель – д-р физ.-мат. наук Мынбаев К.Д.¹

¹*Университет ИТМО*

e-mail: max.ruzhevich@niuitmo.ru

Работа выполнена в рамках темы НИР №620164 «Методы искусственного интеллекта для киберфизических систем».

Твердые растворы кадмий-ртуть-теллур (HgCdTe, КРТ) состава $x \approx 0,3$ применяются для создания фото- и оптоэлектронных устройств, работающих в средней инфракрасной области спектра. Технология выращивания материала менялась на протяжении времени, а вместе с этим менялись оптические характеристики материала. В работе рассмотрены особенности оптических свойств образцов КРТ, выращенных методом молекулярно-пучковой эпитаксии (МПЭ), связанные с разупорядочением твердого раствора.

Ключевые слова: кадмий-ртуть теллур, твердые растворы, инфракрасная оптоэлектроника, оптическое пропускание, фотолюминесценция, молекулярно-пучковая эпитаксия, флуктуации состава.

В работе проводится исследование оптических особенностей образцов HgCdTe (КРТ), выращенных методом молекулярно-пучковой эпитаксии (МПЭ). Проводится сравнение полученных данных с результатами по изучению образцов КРТ, полученных ранее на образцах, выращенными различными группами, а также выращенных другими методами.

Твердые растворы кадмий-ртуть-теллур применяются для создания фото- и оптоэлектронных приборов, работающих в инфракрасной области спектра. КРТ состава $x \approx 0,3$ обладает шириной запрещенной зоны, соответствующей длине волны $\lambda = 3 - 5$ мкм. Приборы на основе данного материалы необходимы для передачи/приема сигнала в окнах прозрачности атмосферы, а также для определения концентрации ряда важных газов.

В работе исследуются образцы КРТ, представляющие собой как объемные кристаллы, выращенные методом направленной кристаллизации, так и эпитаксиальные слои. Эпитаксиальные слои, созданные методом ЖФЭ (жидкофазная эпитаксия), были выращены на подложке CdTe и выполняют в работе большей степени роль референсных образцов, так как в настоящее время этим способом материал выращивается в ограниченном объеме. Основной акцент в работе сделан на образцах выращенных МПЭ, так как данный метод — один из наиболее современных и перспективных. В таблице представлена информация о параметрах исследованных образцов.

Для исследования характеристик образцов КРТ применялись оптические методы. В первую очередь это измерение спектров оптического пропускания образцов после выращивания. Для этой цели использовался Фурье-спектрометр Infracum-801, позволяющий получать спектры при комнатной температуре. Кроме этого, проводились измерения спектров ФЛ (фотолюминесценция) в широком диапазоне температур $T = 4,2 - 300$ К. Для этого применялся решеточный монохроматор МДР-23, ФЛ возбуждалась в импульсном режиме полупроводниковым лазером с рабочей длиной волны 1,03 мкм, сигнал регистрировался

приемником на основе InSb (КРТ фотоприемник использовался для более длинноволновых образцов). Для уменьшения влияния внешних факторов применялась техника синхронного детектирования сигнала.

Таблица. Параметры исследованных образцов

Образец	Состав, x	Подложка	Толщина пленки, мкм
C1	0.310	–	–
L1	0.319	CdTe	16
L2	0.280	CdTe	20
L3	0.290	CdTe	16
M1	0.320	Si	8.1
M2	0.320	Si	8.1
M3	0.320	Si	8.1
M4	0.320	Si	8.1
M5	0.320	Si	8.1
M6	0.300	GaAs	8.2
M7	0.309	Si	7.7
M8	0.303	Si	8.4
M9	0.293	Si	8.3
M10	0.297	Si	7.8
M11	0.295	Si	7.6
M12	0.293	Si	7.9
M13	0.296	Si	7.5
M14	0.288	Si	8.2
M15	0.306	Si	8.8

Полученные спектры оптического пропускания позволяют оценить степень однородности состава исследуемых образцов. Так, было установлено, что объемный кристалл обладает наибольшей однородностью состава, в то время как ЖФЭ образцы характеризуются наименьшей степенью однородности, что выражается в «завале» края поглощения на спектре. Это может быть связано с особенностями технологии выращивания данных образцов, а именно, в наличии широкого слоя переменного состава (от подложки CdTe до эпитаксиального слоя малого состава), который вносит свой вклад в спектр пропускания. Для МПЭ образцов спектры пропускания имеют одинаковый наклон, край поглощения сдвигается с изменением состава. Кроме этого, на спектрах пропускания для этих образцов имеет место интерференционная картина, которая может свидетельствовать о хорошей планарности эпитаксиальных пленок. Спектры пропускания для различных образцов КРТ приведены на рисунке 1.

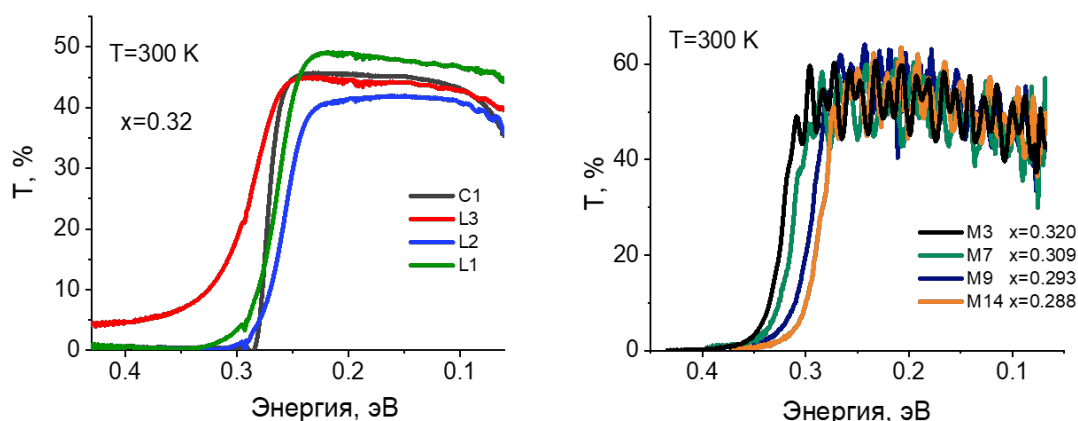


Рис. 1. Спектры пропускания при комнатной температуре

Массив данных, полученный при измерении ФЛ, позволяет оценить некоторые особенности оптических характеристик исследуемых образцов. Форма спектра и его полуширина типична для всех образцов, независимо от способа их выращивания. Однако, стоит отметить, что получить спектры ФЛ во всем температурном интервале $T = 4,2 - 300$ К удавалось только для образцов, выращенных МПЭ. Это может быть объяснено тем, что при

высокой температуре интенсивность излучения МПЭ образцов больше, чем для материала, выращенного иным способом. На рисунке 2 представлены спектры ФЛ различных образцов, полученные при температурах жидкого гелия и жидкого азота.

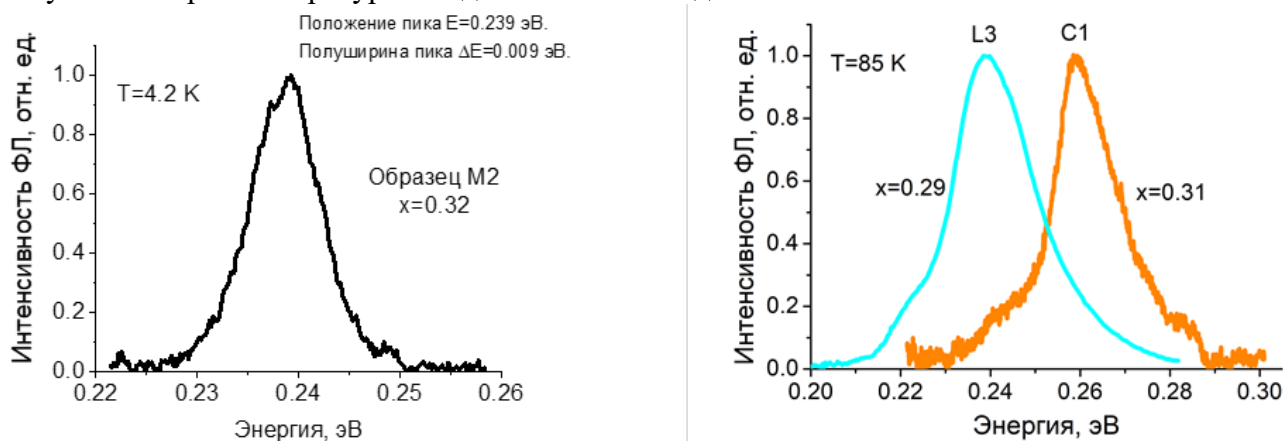


Рис. 2. Спектры ФЛ образцов КРТ, выращенных различными методами

При высокой температуре ФЛ может наблюдаться из-за особенностей, характерных для материала, выращенного МПЭ. На построенных температурных зависимостях на рисунке 3 видно, что энергии пика ФЛ сильно меньше (разница может достигать 25%) расчетной ширины запрещенной зоны E_g . Данное явление связывается с разупорядочением твердого раствора. А именно, предполагается, что носители заряда (экситоны) локализируются на флуктуациях состава и играют определяющую роль в излучательной рекомбинации. С увеличением температуры носители заряда делокализируются и энергия пика ФЛ приближается к величине E_g . Однако, наблюдаемое при комнатной температуре излучение дает основание предполагать, что носители заряда делокализируются не полностью. К похожим выводам уже приходили различные научные группы [1, 2].

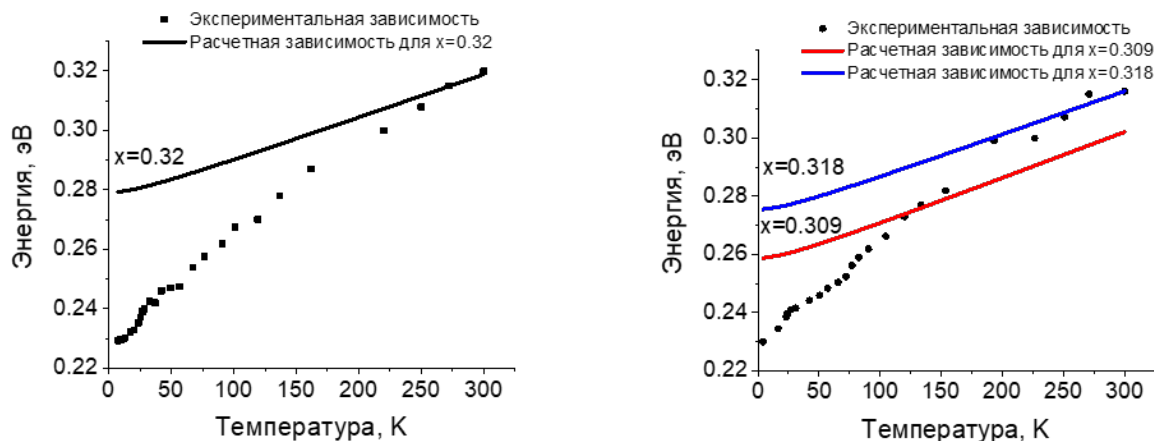


Рис. 3. Температурные зависимости пика ФЛ

Полученные в ходе исследований результаты демонстрируют перспективность КРТ состава $x \approx 0,3$ как материала для создания оптоэлектронных приборов. Выявлены особенности оптических характеристик современных образцов, выращенных МПЭ. Изучение свойств этого материала может оказать положительное влияние как на технологию выращивания КРТ, так и на приборостроение на основе данного материала.

Литература

1. Fang-Yu Yue, Su-Yu Ma, Jin Hong, Ping-Xiong Yang, Cheng-Bin Jing, Ye Chen, and Jun Hao Chu. Optical characterization of defects in narrow-gap HgCdTe for infrared detector applications // Chin. Phys. B. – 2019. – V.28. – №1. – Pp. 1–13.
2. Андрущенко Д.А., Ружевич М.С., Смирнов А.М., Баженов Н.Л., Мынбаев К.Д., Ремесник В.Г. Оптические и структурные свойства эпитаксиальных слоев Hg_{0.7}Cd_{0.3}Te // Физика и техника полупроводников. – 2021. – №55(11). – С. 1040–1044.

УДК 004.62

АВТОМАТИЗАЦИЯ АГРЕГАЦИИ АНАЛИТИЧЕСКИХ ДАННЫХ О ЛЕЧЕНИИ ОНКОЛОГИЧЕСКИХ ЗАБОЛЕВАНИЙ

Рыжков Н.М.¹ (студент), Суренков Н.С.¹ (студент)

Научный руководитель – Береснев А.Д.¹

¹Университет ИТМО

e-mail: n.ryzhkov@corp.mail.ru

Работа выполнена в рамках тема НИР «Формирование требований к системе сбора и анализа данных для НМИЦ онкологии им. Н. Н. Петрова»

В работе рассмотрено проектирование архитектуры информационной системы для автоматизации агрегации и проверки данных, генерируемых в процессе деятельности НМИЦ онкологии и других медицинских организаций и собираемых с целью проведения анализа и построения прогнозов по ключевым метрикам и показателям.

Ключевые слова: онкология, анализ данных, проектирование информационной системы, агрегация данных, качество данных, проверка данных, построение прогнозов.

Федеральное государственное бюджетное учреждение "Национальный медицинский исследовательский центр (НМИЦ) онкологии имени Н.Н. Петрова" Министерства здравоохранения Российской Федерации является одним из крупнейших федеральных научно-лечебных центров России, оказывающих специализированную, в том числе высокотехнологичную медицинскую помощь на уровне мировых стандартов. Научно-исследовательская деятельность центра направлена на разработку методов профилактики, диагностики и лечения онкологических заболеваний [1].

В процессе деятельности медицинского центра создается и передается огромное количество различной информации. Данные передаются как внутри медицинской информационной системы (МИС) НМИЦ онкологии, так и в наружные МИС. Например, МИС популяционного ракового регистра (ППР) или МИС фонда обязательного медицинского страхования (ФОМС) (рис. 1).

Центр контролирует работу онкологических медицинских организаций (МО) всего северо-западного региона. Для организации контроля и поддержки региональных центров внутри структуры НМИЦ сформирован отдел по работе с регионами. В этом отделе была образована группа обработки статистических данных. Группа занимается сбором и проверкой данных и анализом работы регионов. Передача данных между центром по работе и региональными учреждениями регулируется приказом об информационном взаимодействии между ТФОМС, МО и СМО при осуществлении персонифицированного учета оказанной медицинской помощи. Он описывает структуру и формат файлов обмена данными. Регионы передают информацию в пределах внутренней МИС в виде реестра счетов об оказании медицинских услуг за равные периоды времени.

Специалисты группы обработки статистических данных вручную проверяют целостность собранных архивов данных и корректность содержания файлов в архиве. Далее эти данные загружаются в устаревшее программное обеспечение для хранения данных и анализируются. Со временем НМИЦ данных становится все больше, и важность их анализа для центра растет. Группа поставила перед собой цель - анализ более широкого и объемного массива данных для

усовершенствования результатов анализа ранее рассматриваемых метрик и создания возможности анализа новых показателей, анализировать которые с текущим набором данных не предоставляется возможным. Главным образом, расширить набор данных требуется за счет сбора информации из других медицинских подразделений.

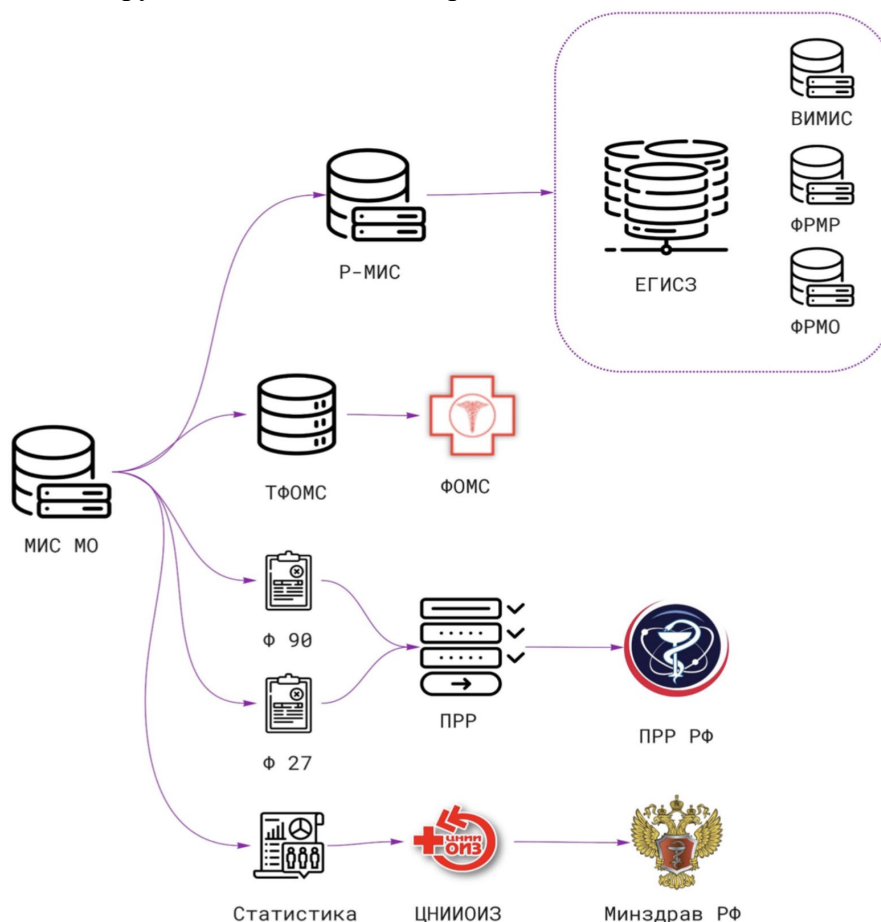


Рис. 1. Схема передачи данных в информационные системы других МИС

Цель данной работы - предоставить возможности расширенного анализа группе из отдела по обработке данных. Для достижения указанной цели необходимо разработать информационную систему. В результате интервьюирования специалистов были выявлены различные требования к системе.

Сформированное бизнес-требование (т.е. цель, которую организация хочет достичь [2]) – автоматизация проверки данных, оптимизация внутренних процессов центра и построение прогнозов ключевых метрик и показателей с помощью анализа расширенного набора данных. Среди пользовательских требований на начальном этапе проектирования есть возможность загрузить файл реестра счетов в общую базу данных и получить отчет автоматизированной проверки на числовую и семантическую корректность, возможность проводить анализ имеющегося расширенного набора данных.

Если численная проверка с точки зрения здравого смысла вполне понятна и реализуема программистами на уровне кода, то семантические проверки довольно сложны и требуют особенных знаний предметной области. Ввиду такой особенности проверок было решено реализовать инструмент для врачей-специалистов, который даст им возможность самим формировать семантические правила валидации данных, специфичные для онкологической области.

Среди системных были выявлены требования по взаимодействию проектируемой системы с внешней МИС в виде подключения к удаленной базе данных под управлением СУБД Firebird. Нефункциональных требований, ограничений в проектировании, реализации и выборе технологий на текущий момент не выдвинуто.

Для удобства взаимодействия с информационной системой специалистам НМИЦ, не являющимися продвинутыми техническими специалистами, принято решение реализовать систему как веб-приложение. Архитектурно приложение будет выполнено в классическом клиент-серверном стиле с сервером приложений (рис. 2).

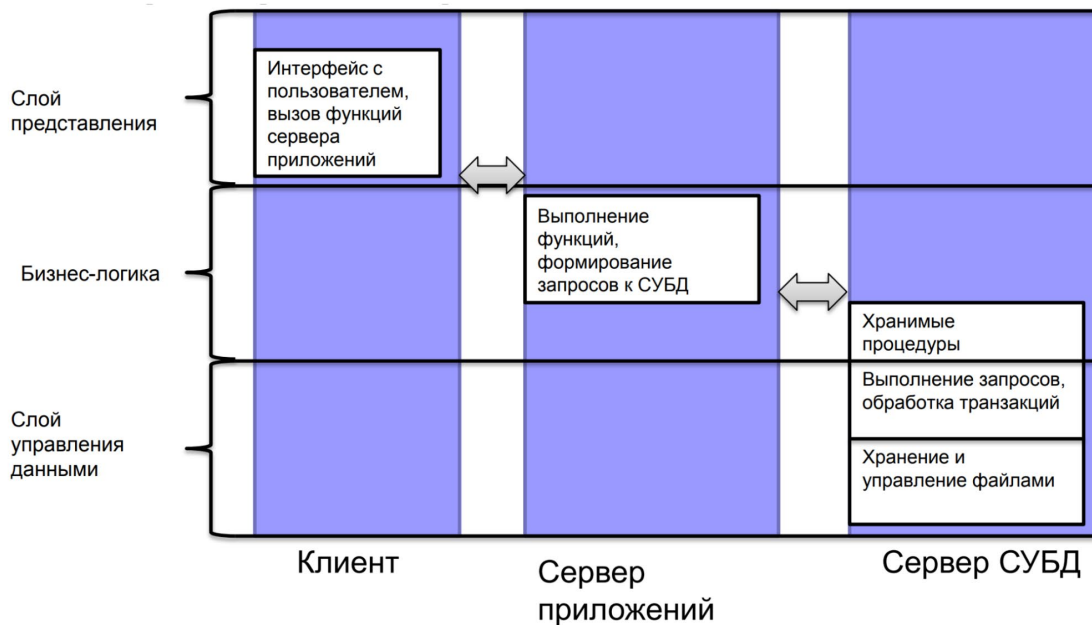


Рис. 2. Графическое представление клиент-серверного приложения

Преимущества такой архитектуры: небольшое количество кода и логики на клиенте, передается минимальное количество информации между клиентом и сервером - аргументы функций и результат выполнения, сервер приложения можно развернуть в любом количестве экземпляров на любом количестве машин, менее трудозатратное и более быстрое обновление функциональности сервера приложения. Из минусов отмечается высокая сложность на администрирование и разработку сервера и наладку взаимодействия между сервером приложений и СУБД [3].

На рисунке 3 представлена схема потока данных от агрегируемых данных до их визуализатора. Общий пайп выглядит так: данные собираются сервером НМИЦ, сохраняются в общую базу данных. Приложение для анализа берет эти данные, проводит анализ и выдает в визуализатор результатов.

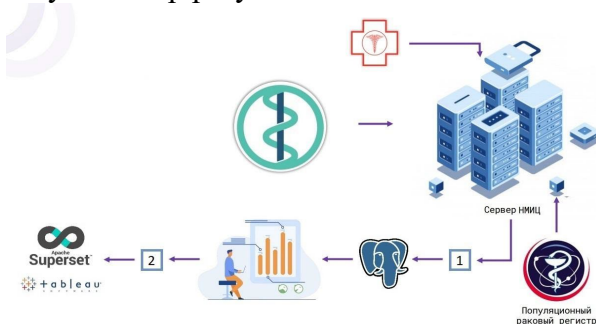


Рис. 3. Схема потока данных в МИС НМИЦ

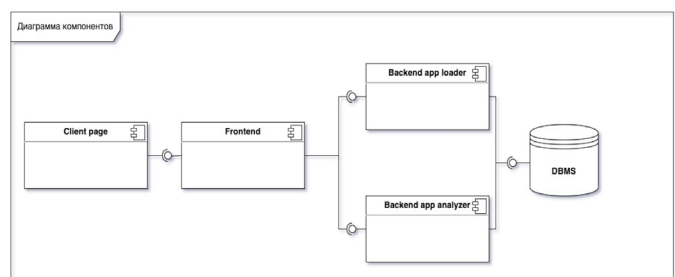


Рис. 4. UML диаграмма размещения компонентов

Всего в процессе проектирования было выделено 5 компонентов: серверное приложение по агрегированию и проверке данных; серверное приложение, проводящее анализ и строящее прогнозы; веб-сервер; веб-страница; база данных.

Схема компонентов приложения и их взаимодействия приведена на рисунке 4.

Были проанализированы и выбраны технологии для реализации описываемой информационной системы. Клиентская часть будет реализована с применением фреймворка React и использованием всего технологического стека веб-разработчика: HTML, CSS,

JavaScript. Серверное приложение для загрузки файлов реестра счетов будет реализовано на языке Python с использованием фреймворка FastAPI. Фреймворк был выбран на основе уже проведенного анализа [4]. Файлы реестра счетов об оказанных медицинских услугах поступают в НИМЦ из регионов в формате XML. Как прямой наследник SGML, он сохранил многие его функции, и в то же время он более простой и человекочитаемый. XML стал основой для многих современных языков разметки [5]. Для преобразования XML в объекты языка Python на основе анализа и тестов производительности была выбрана библиотека lxml.etree.ElementTree. Для организации хранения большого количества информации в качестве СУБД была выбрана PostgreSQL, так как она оптимизирована для обработки больших массивов данных и имеет инструментарий для импорта из других баз данных.

На текущий момент проведено эскизное проектирование веб-сервиса, разработана схема базы данных, создан дизайнерский прототип для страницы формирования правил проверки. В дальнейшие планы входит разработка описанной информационной системы и ее развертывание, подключение внешних источников данных, программная реализация существующих методов анализа и разработка новых с учетом расширенного набора данных.

Литература

1. Официальный сайт НИИЦ онкологии им. Петрова. [Электронный ресурс]. – Режим доступа: [https:// www.niioncologii.ru/institute](https://www.niioncologii.ru/institute). (дата обращения: 24.12.2021).
2. Wiegers K., Beatty J. Software Requirements Third edition, Microsoft Press. – 2013. – 48 p.
3. Проектирование и эксплуатация информационных систем в медиаиндустрии. Лекция 3. Конспект, Выломова Е.Ф. – 15 с.
4. A Comparative Guide to Python Web Frameworks. [Электронный ресурс]. – Режим доступа: <https://analyticsindiamag.com/django-vs-flask-vs-fastapi-a-comparative-guide-to-python-web-frameworks/> (дата обращения: 01.01.2022).
5. Shklar L., Rosen R. Web application architecture principles, protocols and practice. – 2003. – 357 p.

УДК 004.932.2

АНАЛИТИЧЕСКИЙ ОБЗОР КОРПУСОВ ДЛЯ АВТОМАТИЧЕСКОГО ОЦЕНИВАНИЯ ПСИХОФИЗИЧЕСКИХ СОСТОЯНИЙ ЧЕЛОВЕКА

Рюмина Е.В.^{1,2} (аспирант)

Научный руководитель – д-р техн. наук, доцент Карпов А.А.¹

¹Санкт-Петербургский Федеральный исследовательский центр Российской академии наук

²Университет ИТМО

e-mail: ryumina_ev@mail.ru, karpov@iias.spb.su

Исследование выполнено при поддержке Совета по грантам Президента РФ (грант № НШ-17.2022.1.6).

В работе представлен аналитический обзор корпусов CREMA-D, AFEW, Firs Impression V2 и AViD-Corpus пригодных для задачи автоматического оценивания психофизических состояний человека. Актуальность темы обусловлена тем, что разработка интеллектуальных систем автоматического оценивания психофизических состояний человека способствует более естественному человеко-машинному взаимодействию, что наглядно отражается в областях их применения, а именно в: робототехнике, медицине, образовании, маркетинге и т.д.

Ключевые слова: психофизические состояния, гендер, возраст, эмоции, депрессия, персональные качества личности.

На протяжении последних нескольких десятилетий большое внимание научным сообществом уделяется изучению психофизических состояний человека. Психофизические состояния человека – это свойства, которые закономерно отражаются в акустических особенностях речи и голоса, мимике лица, жестах и т.д. Психофизические состояния подразделяются на психические и биофизические [1]. На рисунке 1 представлена классификация психофизических состояний.

В отличие от биофизических состояний, которые остаются неизменчивы достаточно долгое время, психические состояния классифицируются по продолжительности их протекания (рис. 2).



Рис. 1. Классификация психофизических состояний человека



Рис. 2. Классификация психических состояний человека по продолжительности их протекания

Стоит отметить, что автоматическое оценивание психофизических состояний выполняется с помощью информации, передаваемой вербальным и невербальным способами. Так, информация, передаваемая различными способами, формирует систему коммуникаций, представленную на рисунке 3.



Рис. 3. Система коммуникаций

Из рисунка 3 следует, что 93% передаваемой информации приходится на невербальный способ коммуникации, т.е. информация передается посредством аудио и видео модальностей, и лишь 7% приходится на текстовую модальность. Так, для создания надежной системы оценивания психофизических состояний необходимо учитывать несколько модальностей. Следовательно, цель данной работы – изучить наиболее репрезентативные корпуса для оценивания психофизических состояний человека.

Мультимодальные корпуса

На сегодняшний день собрано много разнообразных корпусов для анализа поведения человека, однако, не все они подходят для создания надежных систем анализа психофизических состояний человека. Ниже представлены некоторые проблемы не репрезентативности корпусов:

- недоступность корпусов (например, крупномасштабный корпус UDIVA и т.д.);
- унизадачность корпусов (AffWild-2 (эмоции), VoxCeleb (пол) и т.д.);
- унимодальность корпусов (RusNeuroPsych (аудио), и SFEW (видео) т.д.);
- низкий возрастной диапазон (RAMAS, IEMOCAP и т.д.);
- небольшой объем исследовательских данных (MULTISIMO, SAVEE, GEMER и т.д.);
- лабораторные условия записи (RAVDESS и т.д.).

Несмотря на многообразие корпусов, позволяющих анализировать поведение человека, для оценивания психофизических состояний подходят только 4 корпуса: CREMA-D [2], AFEW [3], First Impression V2 [4], AViD-Corpus [5].

Корпус CREMA-D [2] собирался для задачи распознавания эмоциональных состояний, которых насчитывается 6, а именно: «страх», «печаль», «злость», «счастье», «нейтральное состояние». Корпус разнообразен по количеству дикторов, их возрастному диапазону (от 20 до 74 лет, средний возраст 36), распределению гендера (43 женщины и 48 мужчин), этнической принадлежности (испанец и не испанец) и расе (афроамериканец, азиат, европеец). Данный корпус был записан в лабораторных условиях, это означает, что на лицах дикторов отсутствует окклюзия лица, дикторы скованы в движениях рук, тела и головы, освещение лица и сцены однотипное, отсутствие зашумленности и посторонних звуков на аудиосигналах. Следовательно, условия записи корпуса не являются его преимуществом, поскольку в неконтролируемых условиях, система, обученная на данном корпусе, будет менее эффективна.

Корпус AFEW [3] собирался для задачи распознавания эмоциональных состояний, которых насчитывается 7 (дополнительно к предыдущим 6 состояний рассматривается «удивление»). Корпус разнообразен по количеству дикторов, их возрастному диапазону (от 5 до 76 лет, средний возраст 35) и распределению гендера (139 женщин и 210 мужчин). В отличие от предыдущего корпуса, данный корпус собирался в неконтролируемых условиях. Аудиовидеозаписи для корпуса взяты из фильмов, следовательно видеок cadры лиц имеют широкий диапазон позы головы, условий освещения, окклюзий, а аудио кадры не лишены зашумленности.

Корпус First Impression V2 [4] собирался для задачи распознавания персональных качеств личности человека и оценивания успешности принятия его на работу. В корпусе насчитывается 5 персональных качеств личности, а именно: «открытость опыту», «доброжелательность», «добросовестность», «эмоциональная стабильность», «экстраверсия». Корпус разнообразен по количеству дикторов, их возрастному диапазону (от 11 до 60 лет, средний возраст 27), распределению гендера (1312 женщин и 1748 мужчин), расы (афроамериканец, азиат,

европеец). Корпус собирался в неконтролируемых условиях посредством видеохостинга YouTube, аудиовидеозаписи обладают теми же преимуществами, что и в корпусе AFEW.

Корпус AViD-Corpus [5] собирался для задачи распознавания депрессии. Корпус разнообразен по количеству дикторов, их возрастному диапазону (от 18 до 63 лет, средний возраст 31,5), распределению гендера. Корпус собирался в лабораторных условиях при взаимодействии с компьютером.

В таблице представлена сводная информация об анализируемых корпусах.

Таблица. Параметры анализируемых корпусов

Корпус	Модальность	Кол-во дикторов	Кол-во записей	Задача
CREMA-D [2]	Аудио, видео	91	7442	Оценивание гендера, возраста, эмоций, расы, этнической принадлежности
AFEW [3]	Аудио, видео	349	1163	Оценивание гендера, возраста, эмоций
First Impression V2 [4]	Аудио, видео, текст	3060	10000	Оценивание гендера, возраста, пяти персональных качеств личности, расы
AViD-Corpus [5]	Аудио, видео	292	340	Оценивание гендера, возраста, депрессии

Как можно заметить из таблицы, все корпуса могут быть использованы для оценивания таких биофизических состояний, как гендер и возраст. Психические состояния в корпусах различны, можно выделить три психических состояния: корпуса для оценивания эмоций (CREMA-D и AFEW), персональных качеств личности (First Impression V2) и депрессии (AViD-Corpus).

В ходе исследования был проведен сравнительный анализ корпусов, пригодных для задачи автоматического оценивания психофизических состояний.

В данной работе были выделены 4 корпуса: CREMA-D, AFEW, First Impression V2 и AViD-Corpus. Помимо того, что авторы привлекали для сбора корпусов разных дикторов, анализируемое психическое состояние в корпусах также различно. Поэтому при выборе корпуса для исследования необходимо в первую очередь исходить из того, какое психическое состояние необходимо анализировать. Таким образом, использование данных корпусов позволит построить мультизадачную систему оценивания психофизических состояний человека, при этом точность оценивания каждого отдельно взятого состояния будет превосходить точность, полученную при построении унизадачной системы.

Литература

1. Александров И.О. и др. Психология XXI века: Учебник для вузов / Под ред. В.Н. Дружинина. – П 57 М. ПЕР СЭ. – 2003. – 863 с.
2. Cao H. et al. CREMA-D: Crowd-sourced emotional multimodal actors dataset // IEEE Trans. Affect. Comput. – 2014. – V.5. – №4. – Pp. 377–390.
3. Dhall A. et al. Acted facial expressions in the wild database // Australian National University, Canberra, Australia, Technical Report TR-CS-11. – 2011. – V.2. – Pp. 1–16.
4. Escalante H.J. et al. Explaining first impressions: Modeling, recognizing, and explaining apparent personality from videos // arXiv preprint arXiv:1802.00745. – 2018. – Pp. 1–46.
5. Valstar M. et al. AVEC 2013: the continuous audio/visual emotion and depression recognition challenge // The 3rd ACM international workshop on Audio/visual emotion challenge (AVEC'13). Association for Computing Machinery. – 2013. – Pp. 3–10.

УДК 004.932.2

ОБЗОР МУЛЬТИМОДАЛЬНЫХ КОРПУСОВ ДЛЯ ИССЛЕДОВАНИЯ ВЛИЯНИЯ ЭМОЦИОНАЛЬНОГО СОСТОЯНИЯ ДИКТОРА НА АВТОМАТИЧЕСКОЕ РАСПОЗНАВАНИЕ ФРАЗ ПО ГУБАМ

Рюмина Е.В.^{1,2} (аспирант), Иванько Д.В.¹

Научный руководитель – д-р техн. наук, доцент Карпов А.А.^{1,2}

¹Санкт-Петербургский Федеральный исследовательский центр Российской академии наук

²Университет ИТМО

e-mail: ryumina_ev@mail.ru, denis.ivankov11@gmail.com, karpov@iias.spb.su

Исследование выполнено при поддержке Совета по грантам Президента РФ (грант № МК-42.2022.4).

В работе представлен аналитический обзор мультимодальных корпусов CREMA-D, RAVDESS, SAVEE и eNTERFACE'05 пригодных для задачи автоматического распознавания эмоциональной произвольной речи диктора. Известно, что мультимодальные данные дополняют друг друга, это помогает лучше воспринимать речь говорящего. Именно поэтому использование анализируемых корпусов позволит построить надежные аудиовизуальные системы для автоматического распознавания фраз по речи и губам.

Ключевые слова: чтение по губам, распознавание фраз, эмоции человека, мультимодальные корпуса, эмоционально окрашенная речь.

Современные системы автоматического распознавания речи по аудио- и видеoinформации активно используются в нашей повседневной жизни. Однако, эффективность работы таких систем значительно ухудшается из-за ряда факторов, одним из которых является эмоционально окрашенная речь диктора. Так, в зависимости от эмоции диктора изменяется:

- тембр, высота и громкость голоса;
- длительность произнесения звуков;
- длительность пауз;
- артикуляция.

На рисунке 1 представлены осциллограммы сигналов для одной фразы в зависимости от эмоции диктора.

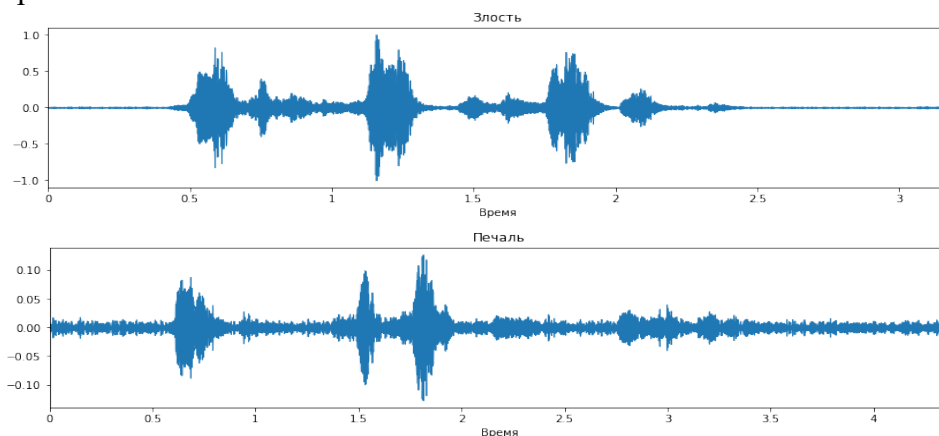


Рис. 1. Осциллограммы сигналов для одной фразы

Сравнение двух осциллограмм сигналов для одной фразы при различных эмоциях диктора показывает, что: 1) при эмоции «злость» амплитуда сигнала примерно в 10 раз превышает амплитуду сигнала при эмоции «печаль»; 2) при эмоции «злость» паузы между словами значительно короче, чем при эмоции «печаль». Аналогичные изменения можно наблюдать не только по аудиоинформации, но и по видеоинформации. На рисунки 2 представлен пример изменения артикуляции губ в зависимости от эмоции диктора.

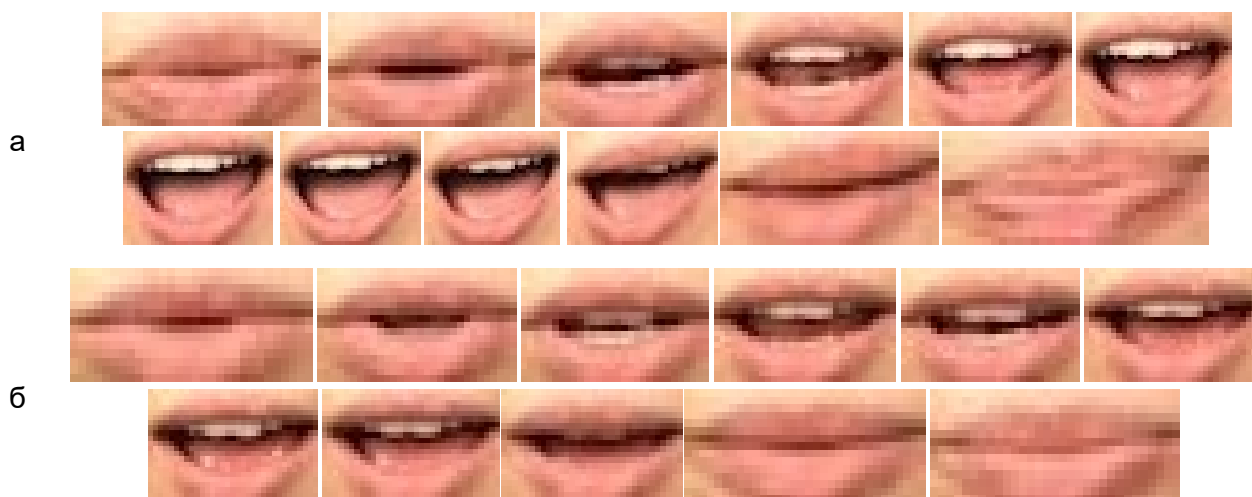


Рис. 2. Пример изменения артикуляции губ в зависимости от эмоции диктора: а) эмоция «злость» (с 50 по 61 кадры), б) эмоция «печаль» (с 84 по 95 кадр)

Сравнение двух примеров изменения артикуляции губ для одной фразы при различных эмоциях диктора показывает, что: 1) при эмоции «злость» деятельность области рта выше, чем при эмоции «печаль»; 2) при эмоции «злость» начало движения губ начинается с 50 кадра, тогда как при эмоции «печаль» с 84 кадра. Наблюдаемое временное смещение кадров характерно тому, что при эмоции «злость» диктор старается быстрее изложить свою мысль, при этом паузы между словами сокращаются.

Сегодня во всем мире активно исследуются и разрабатываются системы автоматического распознавания речи по аудио- и видеоинформации. При этом, научных исследований, направленных на анализ влияния эмоционального состояния диктора на автоматическое распознавание фраз по речи и губам, не проводилось ни в России, ни за рубежом. В связи с этим актуальность данной темы исследований не может быть переоценена и требует детальной проработки.

Мультимодальные корпуса

Мультимодальные корпуса для исследования влияния эмоционального состояния диктора на автоматическое распознавание фраз по речи и губам отличаются в первую очередь тем, какие эмоциональные состояния они охватывают.

Эмоциональное состояние (эмоции/аффект) – это психофизический процесс, отражающий внутренние намерения человека. В научной литературе известны две модели описания эмоций: категориальная и пространственная (непрерывная) [1]. Категориальная модель основана на фиксированной единице, дискретной категории, которая отделяет одну категорию эмоций от других. Так, Полом Экманом было предложено [2] деление эмоций на 6 базовых категорий: «удивление», «счастье», «злость», «страх», «отвращение» и «печаль». Данные категории эмоций проявляются практически идентично у людей со всего мира в независимости от расы и социальной принадлежности. Другой взгляд на теорию эмоций – это пространственная модель (например, модель Рассела [3]), которая рассматривает эмоцию как числовую оценку в двумерном или трехмерном пространстве, а не как отдельные категории. Большинство пространственных моделей включают два измерения, таких как валентность (от англ. Valence) и активация (от англ. Activation) или интенсивность (от англ. Arousal), а в

некоторых случаях добавляется третье – доминирование (от англ. Dominance). Знание как категориальных, так и пространственных эмоций позволяет более точно определять эмоциональное состояние человека [3].

Помимо эмоциональных состояний (количества эмоций и уровней интенсивности), которые охватывают мультимодальные корпуса, они также отличаются друг от друга количеством дикторов, записей, фраз и задач, которые с помощью данных корпусов можно решить.

Несмотря на многообразие корпусов с эмоционально окрашенными данными для решения поставленной задачи подходят только 4 корпуса: CREMA-D [4], RAVDESS [5], SAVEE [6], eNTERFACE'05 [7].

Корпус CREMA-D [4] содержит 7 442 аудиовизуальных записей, на которых 91 диктор (43 женщины и 48 мужчин) с разным уровнем интенсивности (низкий, средний, высокий и неопределенный уровни) имитируют 6 категорий эмоций: «счастье», «печаль», «злость», «страх», «отвращение» и «нейтральность». Каждый диктор произносил 12 фраз с различным эмоциональным окрасом. Одна фраза было составлена с тремя уровнями интенсивности, остальные 11 фраз – с неопределенной интенсивностью. Основной уникальностью данного корпуса является различная этническая принадлежность, гендерное распределение и возрастной диапазон (от 20 до 74 лет) дикторов, это позволит создать более устойчивую систему распознавания фраз в независимости от этнической принадлежности, гендера и возраста дикторов. Что также является одной из основных проблем уменьшения эффективности распознавания речи по аудио- и видеоинформации.

Корпус RAVDESS [5] содержит 4 904 аудиовизуальных записей для речевого и мелодичного воспроизведения эмоциональных фраз, на которых 24 диктора (по 12 мужчин и женщин) с различным уровнем интенсивности (нормальный, высокий уровни) имитируют 8 категорий эмоций (дополнительно к 6 эмоциям из предыдущего корпуса рассматриваются «спокойствие» и «удивление»). Каждый диктор произносил 2 фразы с различным эмоциональным окрасом. Данный корпус уникален тем, что имеет как речевое, так и мелодичное воспроизведение фраз, это позволит исследовать влияние не только эмоционального состояния диктора, но и также стиля воспроизведения на точность распознавания фраз.

Корпус SAVEE [6] содержит 480 аудиовизуальных записей, на которых 4 мужчины-диктора имитируют 6 базовых эмоций и «нейтральное состояние». Для каждого эмоционального состояния записано 15 фраз: 3 общих и 12 эмоциональных. Для нейтрального состояния помимо 15 фраз, повторно были записаны 3 общие фразы и 12 эмоциональных фраз (по 2 фразы от 6 других эмоциональных состояний). Также стоит отметить, что при записи видео авторы корпуса показывали дикторам изображения с типичной мимикой для каждого эмоционального состояния и 270 текстовых подсказок на дисплее. От дикторов требовалось максимально точно передать эмоции, представленные на экране. Кроме того, чтобы зафиксировать ключевые особенности каждого выражения лица, 60 синих маркеров были нанесены на лоб, брови, щеки, губы и челюсть участников. Еще одним плюсом корпуса является фонетическая временная разметка, которая позволит исследователям, помимо фраз, распознавать n-gramm (последовательности из n слов).

Корпус eNTERFACE'05 [7] содержит 1166 аудиовизуальных записей, на которых 42 диктора (34 мужчины и 8 женщин) имитируют 6 базовых эмоций. Для каждого эмоционального состояния записывалось по 5 фраз, однако, не все дикторы сумели правдоподобно симпровизировать эмоции, поэтому часть видео были исключены из корпуса. Для получения более естественной имитации эмоции авторы корпуса ставили дикторов в определенные ситуации, за которыми должна была следовать определенная эмоциональная реакция.

В таблице представлена сводная информация об анализируемых корпусах. Как можно заметить из таблицы, первые два корпуса (CREMA-D и RAVDESS) пригодны для задачи распознавания фраз в зависимости от эмоции и интенсивности их воспроизведения. Корпус

RAVDESS также подходит для распознавания фраз в зависимости от стиля речи. С помощью корпуса SAVEE можно распознавать фразы, слова и n-грамм в зависимости от эмоции диктора. Корпус eNTERFACE'05 меньше других подходит для анализа распознавания фраз в зависимости от эмоционального состояния диктора, поскольку фразы в корпусе изолированы для каждой конкретной эмоции.

Таблица. Параметры анализируемых корпусов

Корпус	Кол-во эмоций	Кол-во классов (фраз)	Задача
CREMA-D [4]	6 (- «удивление») + 4 уровня интенсивности	12	Распознавание фраз в зависимости от эмоций и их интенсивности
RAVDESS [5]	8 (+ «спокойствие») + 2 уровня интенсивности	2	Распознавание фраз в зависимости от эмоций, их интенсивности и стиля речи
SAVEE [6]	7 + фонетическая разметка	3 (общих) + 12 (эмоциональных фраз для «нейтрального состояния»), 84 (по 12 фраз для каждой эмоции)	Распознавание фраз, слов, n-грамм в зависимости от эмоций
eNTERFACE'05 [7]	6 (- «нейтральное состояние»)	35 (по 5 фраз для каждой эмоции)	Распознавание фраз индивидуально для каждой эмоции

В ходе исследования был проведен сравнительный анализ мультимодальных корпусов, пригодных для задачи автоматического распознавания эмоциональной непроизвольной речи диктора.

В исследовании были выделены 4 корпуса: CREMA-D, RAVDESS, SAVEE и eNTERFACE'05. Корпус CREMA-D является более перспективным для использования в рамках решаемой проблемы, поскольку в сборе корпуса принимал участие 91 диктор в возрастном диапазоне от 20 до 74 лет. Кроме того, в нем представлено 12 общих (для каждого эмоционального состояния) фраз, что значительно превышает общее количество фраз в других корпусах. Ожидается, что с помощью данного корпуса будет получен более точный анализ влияния эмоционального состояния диктора на автоматическое распознавание фраз по речи и губам.

Литература

1. Ляксо Е.Е., Фролова О.В., Гречаный С.В., Матвеев Ю.Н., Верхоляк О.В., Карпов А.А. Голосовой портрет ребенка с типичным и атипичным развитием. – СПб.: Издательско-полиграфическая ассоциация высших учебных заведений. – 2020. – 204 с.
2. Ekman P. Are there basic emotions? // Psychol Rev. – 1992. – V.99. – №3. – Pp. 550–553.
3. Russell J.A. A circumplex model of affect // Journal of Personality and Social Psychology. – 1980. – V.39. – №6. – Pp. 1161–1178.
4. Cao H., Cooper D. G., Keutmann M. K., Gur R. C., Nenkova A. and Verma R. CREMA-D: Crowd-sourced emotional multimodal actors dataset // IEEE Trans. Affect. Comput. – 2014. – V.5. – №4. – Pp. 377–390.
5. Livingstone S.R. and Russo F.A. The ryerson audio-visual database of emotional speech and song (RAVDESS): A dynamic, multimodal set of facial and vocal expressions in north American english // PLoS One. Article e0196391. – 2018. – V.13. – №5. DOI:10.1371/journal.pone.0196391.
6. Haq S., Jackson P. and Edge J. Audio-visual feature selection and reduction for emotion classification // AVSP. – 2008. – Pp. 185–190.
7. Martin O., Kotsia I., Macq B. and Pitas I. The eNTERFACE'05 audio-visual emotion database // 22nd International Conference on Data Engineering Workshops (ICDEW'06). – 2006. – Pp. 1–8.

УДК 004.934

ВЛИЯНИЕ АКУСТИЧЕСКИХ УСЛОВИЙ НА РАБОТУ СИСТЕМЫ ОБНАРУЖЕНИЯ АТАК ПОВТОРНОГО ВОСПРОИЗВЕДЕНИЯ ПРИ ГОЛОСОВОЙ АУТЕНТИФИКАЦИИ

Самородова М.Э.¹ (студент)

Научный руководитель – Чирковский А.Д.²

¹*Университет ИТМО*

²*ООО «ЦРТ»*

e-mail: beautiful.wolf.fury@gmail.com

В данной работе исследуется влияние внешних акустических условий на качество работы системы обнаружения атак повторного воспроизведения при голосовой аутентификации и рассматривается возможный вариант повышения надёжности системы: аугментация данных для обучения модели импульсными откликами различных помещений и устройств записи. Результатом работы является анализ особенностей функционирования системы антиспуфинга в зависимости от акустических условий. Актуальность работы обусловлена тем, что в настоящее время слабо исследованы способы адаптации модели под разные факторы физической среды, негативно влияющие на качество системы антиспуфинга.

Ключевые слова: биометрические системы, антиспуфинг, атака повторного воспроизведения, реверберация, акустические условия, нейронные сети.

Существует два традиционных метода аутентификации личности. Первый основан на знании пароля или PIN-кода, в то время как второй – на владении идентификационным объектом (или «токеном»): удостоверение личности, ключ, пропуск и т.д. Эти два метода аутентификации можно использовать взаимодополняющим образом для повышения безопасности, как в банковских картах. Однако, у них есть свои недостатки. В первом случае пароль может быть забыт или угадан третьим лицом. Во втором же случае «токен» может быть утерян или украден. Биометрические системы являются альтернативным решением двух предыдущих методов аутентификации, которые всё чаще используются для защиты от несанкционированного доступа. Несомненным преимуществом таких систем является использование биологических характеристик, которые являются уникальными для каждого человека (отпечатки пальцев или ладоней, сетчатка или радужная оболочка глаза, геометрия рук, голос). В голосовой биометрии распознавание говорящего осуществляется с помощью акустических особенностей голоса человека, таких как физиологические и поведенческие.

С каждым годом всё больше растёт интерес к технологиям голосовой аутентификации. В первую очередь это обусловлено доступностью данного подхода в век цифровизации: многие персональные устройства такие, как смартфоны, персональные компьютеры, обладают встроенным микрофоном. Кроме того, использование глубоких нейронных сетей, позволяющих работать с большими объёмами данных, повышает внутриклассовую и межклассовую вариативность современных систем аутентификации голоса [1].

Так, в настоящее время уже существует множество областей, где голосовая аутентификация более чем успешно реализована для обеспечения безопасности. Это и доступ к смартфону или персональному компьютеру, и управление удалённым доступом, и осуществление банковских операций, и пересечение государственных границ, а также любое взаимодействие с государственными органами, требующее аутентификации.

Однако, с ростом востребованности голосовой аутентификации для бизнеса возрастает и количество мошенников, желающих обойти систему безопасности. При этом системы аутентификации по голосу успешно способны справляться со злоумышленниками, которые не пытаются изменить свой голос и сделать его похожим на целевого диктора. Но они уязвимы к более сложным спуфинг-атакам [2–4].



Рис. 1. Работа тандемной системы

Спуфинг-атака – это попытка обмануть злоумышленником систему аутентификации личности, посредством представления поддельного или скопированного биометрического образца или путем умышленного изменения собственных биометрических характеристик. Поэтому обычно система аутентификации работает в тандеме с системой антиспуфинга (рис. 1).

Система антиспуфинга является бинарным классификатором, который делит входные данные на два класса: живая речь (класс genuine или bonafide) и спуфинг-атака (класс spoof). Проблема обнаружения спуфинг-атак вызывает высокий интерес в исследовательском сообществе биометрии [3].

Как продемонстрировано на рисунке 2 (а – логический доступ, б – физический доступ), спуфинг-атака может быть совершена в логическом и физическом канале. В первом случае предполагается, что злоумышленник каким-то образом взломал систему и в обход микрофона подал поддельную запись в систему. Во втором случае – атака совершается повторным воспроизведением записи и в отличие от логического канала включает распространение звука в акустической среде.



Рис. 2. Изображение логического и физического доступа в канал связи

Атаки через физический канал более вероятны, поскольку они требуют меньше технических знаний. Внешние условия записи, например, такие, как реверберация, возникающие в физическом канале, могут негативно сказываться на работе системы антиспуфинга и ухудшать качество всей системы аутентификации, что недопустимо в приложениях безопасности, например, таких как доступ к банковским счетам. Поэтому в рамках данной работы исследуется влияние акустических условий на работу системы обнаружения атак повторного воспроизведения при голосовой аутентификации.

В рамках настоящей работы основное внимание уделяется не архитектуре нейронной сети, на основе которой осуществляется обучение модели антиспуфинга, а анализу влияния внешних условий, таких как реверберация, на работу системы и поиску возможных путей повышения устойчивости системы к акустическим условиям. В качестве архитектуры была рассмотрена архитектура LFCC-LCNN [5]. Выбор был обусловлен тем, что она показала достаточно высокую производительность в разные годы в рамках конкурса ASVspoof [6]. В качестве набора данных рассматривался набор данных ASVspoof2021-PA [33] (трек физический доступ), который был позаимствован из конкурса [6].

Живая речь (genuine) в ASVspoof2021-PA записывалась в реальном физическом пространстве, в то время как спуфинг-атаки (spoof) сначала записывались и затем повторно воспроизводились с использованием устройств различного качества. Помещения для записи и повторного воспроизведения могли различаться. Соответственно, данные содержат реверберацию и аддитивный шум как от окружающей среды, так и от устройств воспроизведения. Подача записей в систему производилась в девяти разных комнатах,

в которых были расположены микрофоны трех разных типов на каждом из шести различных расстояний от говорящего до системы аутентификации. Следовательно, существует $162 (= 9 \times 3 \times 6)$ различных сценариев внешних условий для записи. Как подлинные, так и поддельные записи были подвергнуты понижению частоты дискретизации до 16 кГц и предварительно обработаны для удаления начальных и конечных сегментов тишины, превышающих 0,3 секунды.

Набор данных ASVspoof2021-PA включает в себя две части: «progress» и «eval». При этом первая использовалась для обучения модели, а вторая – для тестирования. Также для обучения модели применялась кросс-валидация с разбиением на $k=5$ частей, то есть модель на $k-1$ частей обучалась, а оставшаяся часть использовалась для валидации. Такой подход позволяет наиболее равномерно использовать данные для оценки эффективности модели. Для объективной оценки надёжности работы системы антиспуфинга используется метрика оценки качества бинарных систем принятия решения – равный уровень ошибок (EER). Это коэффициент, при котором ошибки первого и второго рода равны друг другу. Чем ниже EER, тем выше точность системы.

Разметку данных по $rt60$ произведём с помощью нейросетевой модели QE [7], где $rt60$ – это время реверберации, которое необходимо для спада плотности звуковой энергии на 60 дБ от первоначального уровня. При обучении моделей отнесём записи живой речи к классу 0, то есть $genuine = 0$, а записи атаки к классу 1, то есть $spoof = 1$. Гистограммы распределения $rt60$ по классам представлены на рисунке 3.

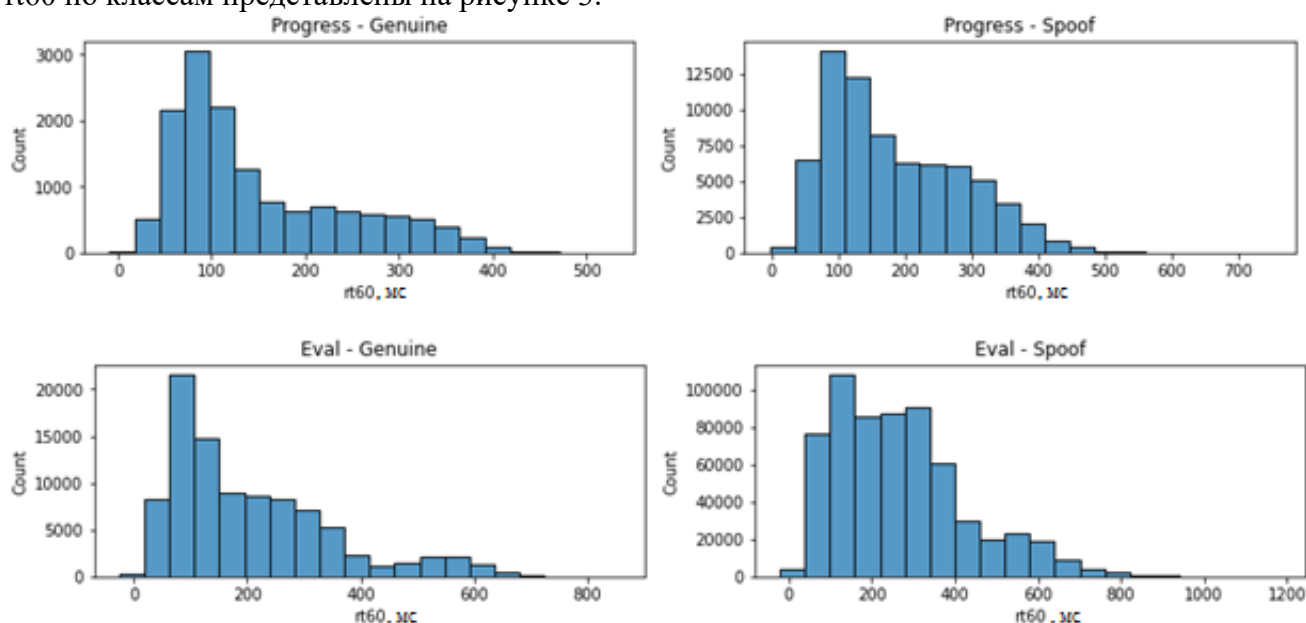


Рис. 3. Гистограмма распределения $rt60$ набора данных ASVspoof2021

Оценки EER и порогов принятия решения (thr) были рассчитаны для progress и eval данных. В случае progress части использовалось кросс-валидационное предсказание, для eval данных EER и порог рассчитаны усреднением предсказаний всех 5 моделей (табл. 1).

Таблица 1. EER и порог принятия решения на progress и eval части

	EER	thr
progress	0.0781	0.6227
eval	0.1969	0.6864

Также были рассчитаны оценки по квартилям $rt60$, которые представлены в таблице 2. С увеличением реверберации наблюдается увеличение порога, то есть модель склонна считать оригинальные записи в помещении с высокой реверберацией ближе к спуфинг-атакам. Склонность модели относить записи с большим значением $rt60$ к спуфинг-атакам обосновывается тем, что на обучающих данных было больше атак с высокой реверберацией, чем оригинальной речи (рис. 3).

Таблица 2. EER и порог принятия решения на progress и eval части по квартилям rt60

Квартиль	EER progress	EER thr progress	EER eval	EER thr eval
1	0.0867	0.5025	0.1829	0.6043
2	0.0715	0.6477	0.1879	0.6744
3	0.0820	0.6715	0.2129	0.7090
4	0.0768	0.6997	0.2160	0.7036

Возможным решением данной проблемы может являться аугментация данных обучения, с помощью которой удастся сбалансировать классы genuine и spoof так, чтобы уменьшить степень влияния реверберации. Аугментация будет осуществляться добавлением реверберации к исходным записям с помощью свёртки импульсного отклика помещения с сигналом.

Используемая база для аугментации состоит из импульсных откликов реальных помещений с частотой дискретизации 16 кГц. Данную базу условно можно поделить на 3 категории по реверберации: 100, 300 и 600 мс.

Для анализа итоговой реверберации после аугментации ко всем исходным данным для обучения были в отдельности добавлены импульсные отклики из всех трёх категорий фильтров. Графики зависимости изменения значения $rt60$ от его начального значения представлены на рисунке 4 для случая аугментации файлами в 100, 300 и 600 мс. Коэффициент детерминации равен 0.15, 0.35 и 0.64 соответственно, то есть, о линейной зависимости можно говорить только в случае импульсных откликов в 600 мс. С повышением исходного значения $rt60$ наблюдается понижение прироста реверберации, то есть аугментировать целесообразнее файлы с меньшими начальными значениями реверберации.

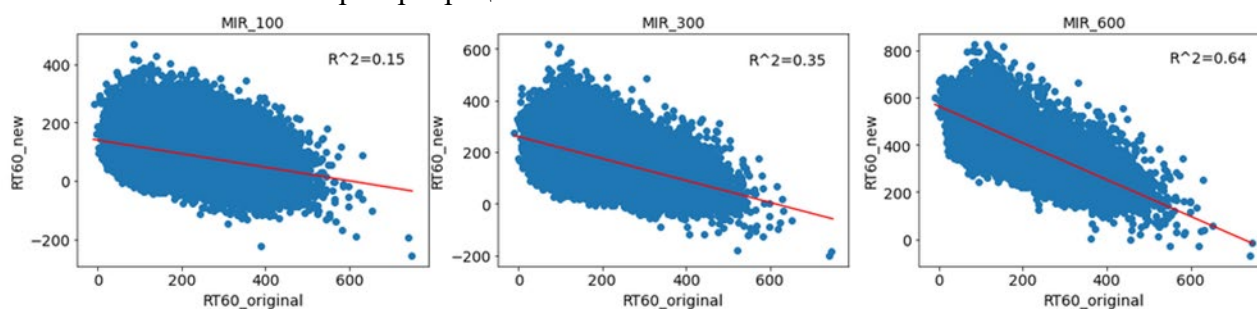


Рис. 4. Зависимость $rt60$ аугментированного файла от начального $rt60$ исходного файла в случае аугментации категорией в 100 мс, 300 мс и 600 мс

Как уже упоминалось ранее, распределение живой речи и спуфинг-атак немного отличаются, поэтому была выдвинута гипотеза: аугментация данных позволит приблизить распределения классов и убрать их несбалансированность, тем самым уменьшив зависимость от внешних условий. Для проверки данной гипотезы была осуществлена балансировка обучающей выборки по реверберации тремя способами: только на основе исходных данных, с добавлением аугментации только для класса живой речи, с добавлением аугментации для обоих классов.

Так же как и ранее была произведена кросс-валидация и для каждого сбалансированного набора данных было обучено 5 моделей. Оценка моделей осуществлялась на 4 наборах данных: тестовая часть «progress» данных, «eval» данные, а также сформируем новые тестовые выборки из данных без аугментации и только из аугментированных данных. Полученные результаты представлены в таблицах 3–6.

Таблица 3. EER и порог принятия решения на progress и eval части

Модель	EER _{progress}	thr _{progress}	EER _{eval}	thr _{eval}
Без аугм.	0.0791	0.5513	0.2086	0.6453
Аугм. обоих классов	0.1439	0.4297	0.2327	0.5462
Аугм. класса живой речи	0.0302	0.4973	0.2076	0.7032

Таблица 4. EER и порог принятия решения на progress и eval части по квартилям rt60 для сбалансированной модели без аугментации

Квартиль	EER _{progress}	thr _{progress}	EER _{no aug}	thr _{no aug}	EER _{aug}	thr _{aug}	EER _{eval}	thr _{eval}
1	0.0759	0.5689	0.0695	0.6152	0.4166	0.8287	0.1870	0.5951
2	0.0754	0.5247	0.0993	0.5717	0.4626	0.8297	0.1891	0.6335
3	0.0789	0.5812	0.0923	0.6619	0.4637	0.8304	0.2044	0.6833
4	0.0865	0.5298	0.1108	0.5858	0.4889	0.8309	0.2252	0.6492

Таблица 5. EER и порог принятия решения на progress и eval части по квартилям rt60 для сбалансированной модели с аугментацией обоих классов

Квартиль	EER _{progress}	thr _{progress}	EER _{no aug}	thr _{no aug}	EER _{aug}	thr _{aug}	EER _{eval}	thr _{eval}
1	0.1278	0.4205	0.0791	0.4349	0.1496	0.3926	0.2114	0.5621
2	0.1441	0.4385	0.1059	0.4695	0.1418	0.4120	0.2439	0.5634
3	0.1394	0.4377	0.0897	0.4806	0.1619	0.4092	0.2401	0.5386
4	0.1634	0.4223	0.1107	0.4544	0.1653	0.3872	0.2385	0.4741

Таблица 6. EER и порог принятия решения на progress и eval части по квартилям rt60 для сбалансированной модели с аугментацией только класса живой речи

Квартиль	EER _{progress}	thr _{progress}	EER _{no aug}	thr _{no aug}	EER _{aug}	thr _{aug}	EER _{eval}	thr _{eval}
1	0.0503	0.4774	0.0559	0.2998	0.5869	0.1721	0.1999	0.6246
2	0.0285	0.4949	0.0993	0.3513	0.5826	0.1702	0.2221	0.7195
3	0.0223	0.4994	0.0266	0.5193	0.5615	0.1686	0.2423	0.7378
4	0.0148	0.5125	0.0715	0.7452	0.5447	0.1681	0.2165	0.7518

Общая ошибка при обучении на сбалансированных наборах данных на «eval» части вышла чуть больше в случае балансировки без аугментации и с аугментацией только класса живой речи по сравнению с обучением на исходных данных. Наибольшая общая ошибка наблюдается при аугментации обоих классов. Относительное изменение порог по квартилям реверберации представлено в таблице 7.

Таблица 7. Относительное изменение порогов на progress и eval части по квартилям реверберации

Обучение	progress	eval
Исходные данные	27,5%	14,72%
Балансировка без ауг.	8,84%	12,91%
Балансировка с ауг. обоих классов	4,11%	15,85%
Балансировка с ауг. только класса живой речи	6,85%	16,92%

Как видно из представленных результатов на progress части данных, сбалансированное обучение во всех случаях в меньшей степени зависит от реверберации. На eval части только балансировка без аугментации в меньшей степени зависит от реверберации. Это может быть связано с тем, что реальные данные отличаются от аугментированных, и модель склонна переобучаться на них.

Таким образом, гипотеза о снижении влияния реверберации на качество работы системы антиспуфинга за счёт применения балансировки данных с помощью аугментации была подтверждена. При этом аугментация вносит в данные свои особенности, которые модель выучивает, что снижает её обобщающую способность.

Литература

1. Taigman Y. et al. Deepface: Closing the gap to human-level performance in face verification // Proceedings of the IEEE conference on computer vision and pattern recognition. – 2014. – С. 1701–1708.
2. Wu Z. et al. Spoofing and countermeasures for speaker verification: A survey // Speech communication. – 2015. – Т. 66. – С. 130–153.

3. Ergünay S.K. et al. On the vulnerability of speaker verification to realistic voice spoofing // 2015 IEEE 7th International Conference on Biometrics Theory, Applications and Systems (BTAS). – 2015. – С. 1–6.
4. Alegre F., Amehraye A., Evans N. Spoofing countermeasures to protect automatic speaker verification from voice conversion // 2013 IEEE International Conference on Acoustics, Speech and Signal Processing. – 2013. – С. 3068–3072.
5. ASVspoof 2021 Baseline Systems. [Электронный ресурс]. – Режим доступа: <https://github.com/asvspoof-challenge/2021> (дата обращения: 14.01.2022).
6. ASVspoof2021: Automatic Speaker Verification Spoofing and Countermeasures Challenge. [Электронный источник]. – Режим доступа: <https://www.asvspoof.org/> (дата обращения: 14.01.2022).
7. Lavrentyeva G., et al. Blind Speech Signal Quality Estimation for Speaker Verification Systems // INTERSPEECH. – 2020. – С. 1535-1539.

УДК 004.8

МОДЕЛИ И МЕТОДЫ ОБЕСПЕЧЕНИЯ КИБЕРБЕЗОПАСНОСТИ СИСТЕМ ЦИФРОВОЙ ЭКОНОМИКИ НА ОСНОВЕ МАШИННОГО ОБУЧЕНИЯ

Селезнева Я.М.¹ (студент), Зенкин А.М.¹ (студент)

Научный руководитель – канд. техн. наук Капитонов А.А.¹

¹Университет ИТМО

e-mail: ymselezneva@itmo.ru, a.zenkin@itmo.ru

В статье приведен обзор существующих алгоритмов и методов обеспечения кибербезопасности с применением различных методов машинного обучения с целью анализа и разработки более усовершенствованных алгоритмов и моделей для повышения эффективности и работоспособности систем защиты от атак систем цифровой экономики с целью обеспечения их безопасности.

Ключевые слова: машинное обучение, цифровая экономика, кибербезопасность, сетевые атаки, нейронные сети.

В современном мире цифровизация проникает практически во все сферы деятельности человека. В условиях стремительно развивающихся технологий она также коснулась и экономического сектора. Все большую популярность набирает ведение бизнеса в рамках цифровой экономики. Это означает, что все бизнес-процессы переходят в электронный формат деятельности. Такие условия влекут за собой вероятность повышения кибератак и возможность утери важной информации, а также – финансов.

На основании этого кибербезопасность является одним из ключевых направлений обеспечения безопасности в секторе цифровой экономики. В связи с тем, что данная отрасль является достаточно молодой, то в настоящее время отсутствует единый надежный инструмент контроля и обеспечения безопасности таких систем. Поэтому в настоящее время проводятся различные исследования. Особенно интересным направлением является применение технологий искусственного интеллекта.

Машинное обучение широко применяется при изучении процессов обнаружения аномалий и различных угроз целостности информации при ее обработке, передачи и хранении. Тем самым технологии машинного обучения являются перспективным направлением для реализации систем обеспечения безопасности в секторе цифровой экономики. Целью данной работы является анализ существующих решений и возможностей применения технологий машинного обучения, направленных на определение и защиту систем цифровой экономики.

В одном из существующих исследований в качестве наблюдаемого параметра экономической безопасности авторы определили вероятностный показатель – вероятность безопасного состояния или защищенности системы в данный момент [1].

Авторы утверждают, что основное преимущество обеспечения экономической безопасности на основе данного вероятностного показателя заключается в том, что он нормализован, удобен для оценки сложной схемы событий, имеет четкое, однозначное, функциональное объяснение, отражающее опасность угроз и возможность противодействия им средствами безопасности.

В данном исследовании вероятности скрытых внешних угроз рассчитываются по наблюдаемым аномалиям состояния внешней среды. Для построения вероятностных моделей используются сочетания байесовских сетей и цепей Маркова. На вход байесовской сети

подаются значения наблюдаемых параметров внешней среды, а на выходе получаются вероятности категориальных переменных, которые характеризуют внешние угрозы. Для предсказания угроз данные состояния категориальных переменных связывают с цепью Маркова, которая в свою очередь предсказывает вероятности возможных угроз.

Внутренние угрозы прогнозируются функционально-ресурсным методом в сочетании с логистической регрессией путем выявления аномалий состояния и расчета их вероятности. Обучающая выборка строится с одним или несколькими индикаторами угроз и индикаторами фазового состояния. С помощью однозначной или множественной логистической регрессии рассчитываются вероятности внутренних угроз.

Разработанная система приема, идентификации и обработки внешних и внутренних угроз представлена на рисунке 1.

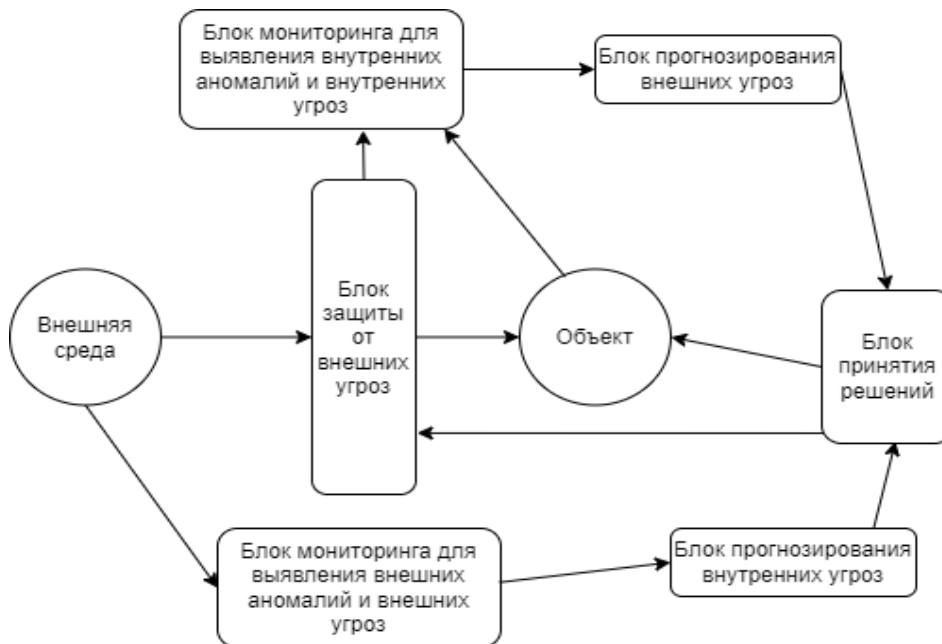


Рис. 1. Структура и компоненты системы контроля безопасности

Следующим рассмотренным методом является метод статического анализа вредоносных программ [2]. Для классификации были взяты признаки, представляющие собой коды операций исполняемых файлов. Для проверки предлагаемых методов авторами был использован наивный байесовский классификатор, который может обрабатывать произвольное количество независимых переменных. Суть алгоритма заключается в том, что на вход подаются предварительно обработанные данные, а именно: количество безопасных исполняемых файлов, количество исполняемых файлов вредоносных программ, количество необходимых функций. Для всех доброкачественных данных происходит добавление частот каждого кода операции и их нормализация относительно количества безопасно исполняемых файлов. То же самое происходит и для файлов вредоносных программ. Затем находится разность нормализованных частот каждого кода операции. В конце возвращается количество кодов операций с наивысшим значением разности нормализованных частот.

Блок-схема обнаружения неизвестных вредоносных программ путем разделения исполняемых файлов на группы представлена на рисунке 2.

Учитывая набор функций (кодов операций), наивный байесовский классификатор дает апостериорную вероятность доброкачественной или вредоносной программы. Окончательная классификация делается путем сравнения апостериорной вероятности между обеими моделями классов. Если апостериорная вероятность тестового исполняемого файла вредоносного класса высока, тогда он классифицируется как вредоносное ПО, в противном случае – как доброкачественное.

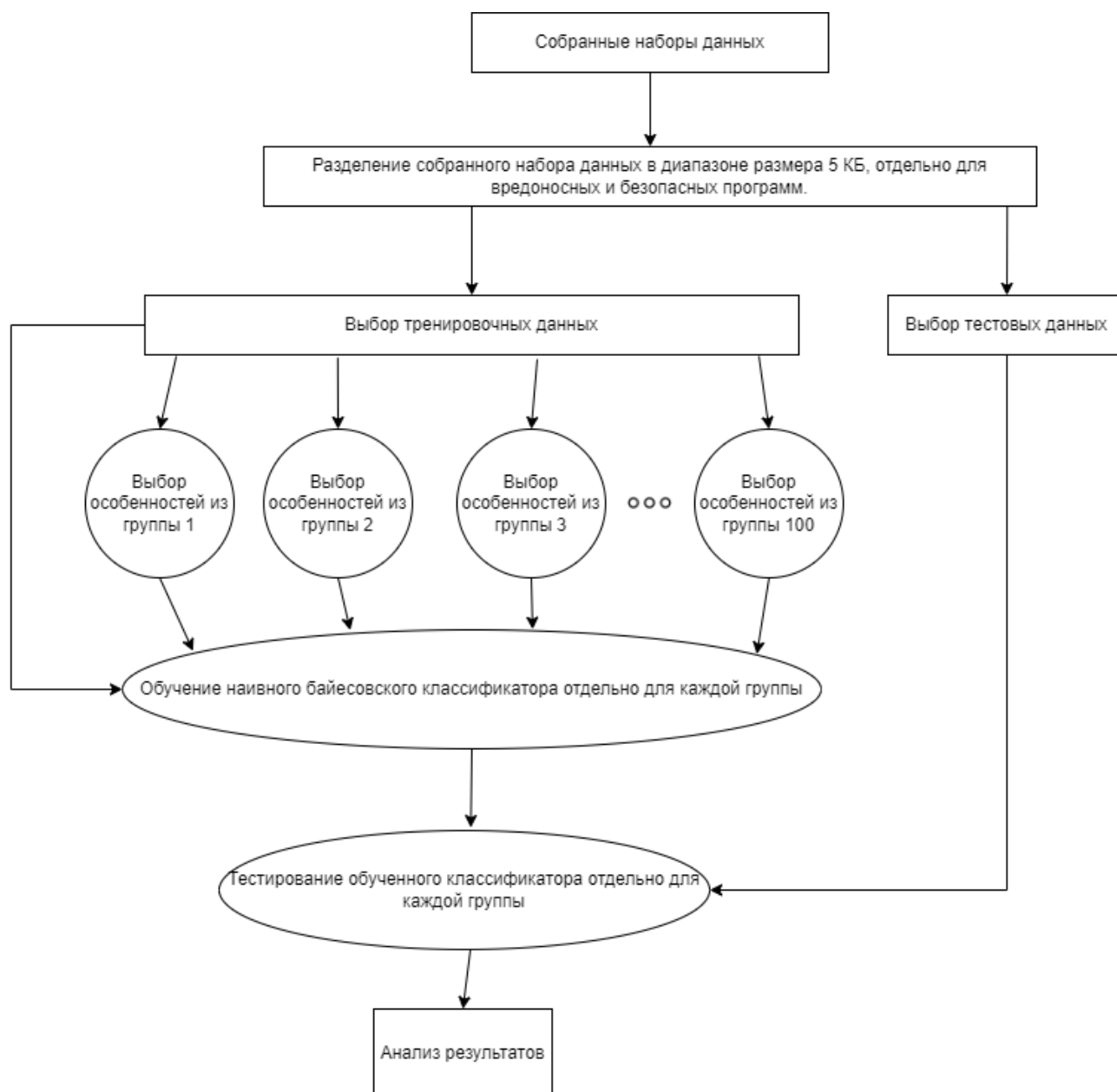


Рис. 2. Блок-схема обнаружения неизвестных вредоносных программ путем разделения исполняемых файлов на группы

Следующий подход оперирует анализом появления кодов операций (функций) путем группировки исполняемых файлов [3]. Авторы утверждают, что данный метод основан на том, что разница между размерами любых двух вредоносных программ находится в пределах 5 КБ. Суть алгоритма заключается в том, что на вход подаются предварительно обработанные данные всех групп, а именно: количество безвредных исполняемых файлов, количество исполняемых файлов вредоносных программ, список функций. Для всех доброкачественных исполняемых файлов вычисляется нормализованная частота каждого кода операции, для всех исполняемых файлов вредоносных программ также вычисляется нормализованная частота каждого кода операции. Далее находится разность значений частот файлов. Затем происходит сортировка полученных значений и устанавливается пороговое значение таким образом, чтобы из каждой группы было выбрано не менее 10 кодов операций. Блок-схема обнаружения вредоносных программ представлена на рисунке 3.

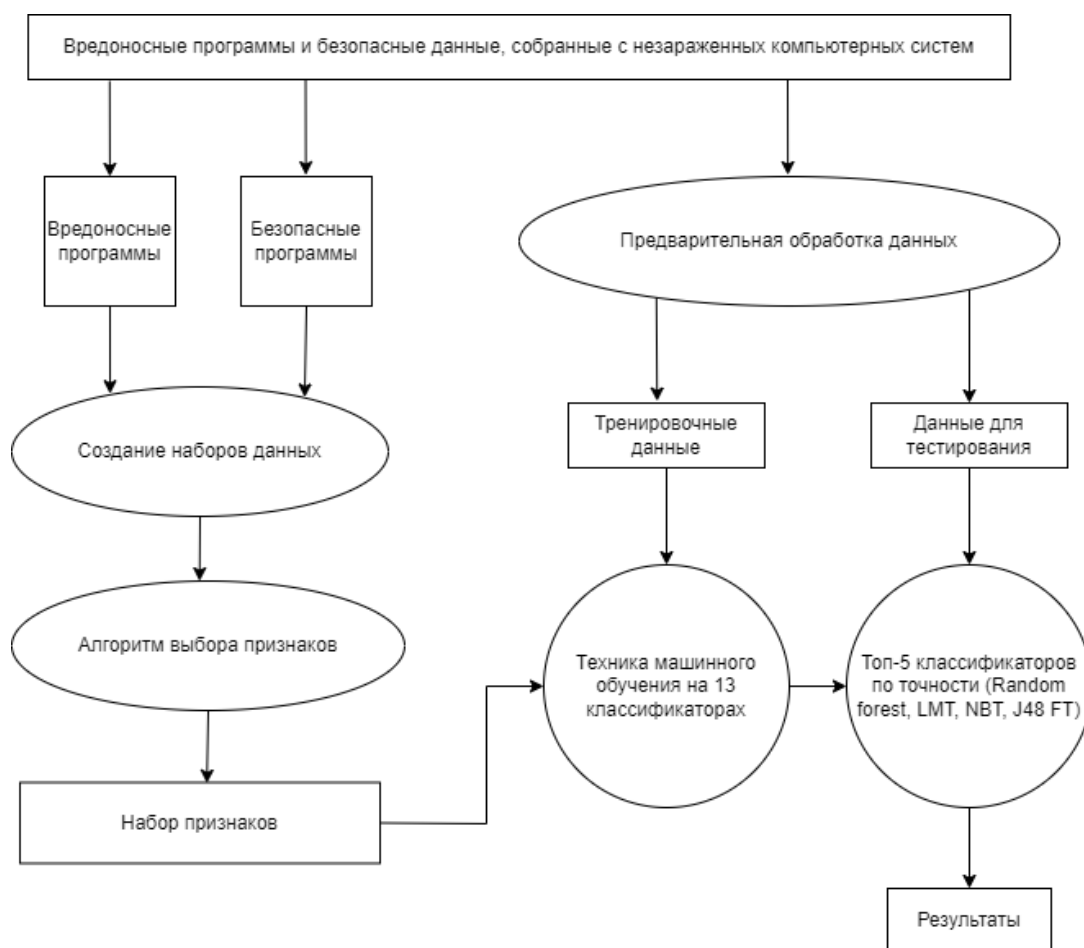


Рис. 3. Блок-схема обнаружения неизвестных вредоносных программ

Полученные результаты анализа будут использованы в дальнейшей работе по разработке модели обеспечения кибербезопасности с целью увеличения эффективности работы систем по выявлению и предотвращению атак на системы цифровой экономики.

Литература

1. Voronin E.A., Kozlov S.V., Kubankov A.N. Integrated Control of Economic Security by Machine Learning in the Information Space of the Digital Economy // Systems of Signals Generating and Processing in the Field of on Board Communications. – 2021. – Pp. 1–5.
2. Ashu Sharma, Sanjay K. Sahay, Abhishek Kumar. Improving the Detection Accuracy of Unknown Malware by Partitioning the Executables in Groups // Advanced Computing and Communication Technologies. – 2016. – Pp. 421–431.
3. Ashu Sharma, Sanjay K. Sahay. An effective approach for classification of advanced malware with high accuracy // International Journal of Security and Its Applications. – 2016. – Pp. 249–266.

УДК 004.03

ПОДХОДЫ К РАЗРАБОТКЕ ПРИЛОЖЕНИЙ С NOSQL И SQL СУБД

Сергеев А.И.¹ (студент)

Научный руководитель – канд. техн. наук, доцент Зудилова Т.В.¹

¹Университет ИТМО

e-mail: andrey0503@gmail.com

В процессе проектирования и разработки приложений всегда возникает вопрос, связанный с тем, какая система управления базами данных (СУБД) лучше подходит для хранения данных реализуемого приложения. Для решения данного вопроса предлагаются критерии оценки характеристик SQL и NoSQL СУБД, влияющих на качество и удобство разработки программного обеспечения.

Ключевые слова: SQL СУБД, NoSQL СУБД, разработка приложений, критерии выбора СУБД, качество ПО.

В современном мире практически каждое приложение хранит всю необходимую информацию в базе данных. Однако, в процессе разработки приложения у архитекторов и разработчиков всегда возникает вопрос, связанный с тем, какая система управления базами данных лучше подходит для реализуемого приложения. Этот вопрос является одним из самых важных и сложных в процессе разработки приложения, так как занимает много времени и требует большого анализа как SQL и NoSQL СУБД, так и требований к самому приложению. Выбранная система управления базами данных должна удовлетворять всем текущим потребностям, а в идеале и будущим. Более того, неправильный выбор может привести к неудобствам в процессе разработки приложения, к некорректной работе приложения, а в худшем случае и к дальнейшим ресурсоемким миграциям на другую СУБД.

Оптимальным решением данного вопроса является определение критериев оценки характеристик SQL и NoSQL СУБД, влияющих на качество и удобство разработки программного обеспечения (ПО). Конечно же список критериев может изменяться в зависимости от требований к приложению и нужд разработчиков, однако, среди них можно выделить наиболее важные критерии, которые необходимо всегда учитывать при выборе системы управления базами данных.

Качество ПО – это комплекс характеристик программного продукта, определяющих способность выполнять возложенные на него функции. Основные характеристики качества программного обеспечения согласно стандарту ГОСТ Р ИСО/МЭК 25010-2015: функциональная пригодность, уровень производительности, совместимость, удобство использования, надежность, защищенность, сопровождаемость, переносимость [1].

В данной работе под качеством ПО будут рассматриваться только следующие характеристики:

1. Функциональная пригодность, так как на функциональность ПО может повлиять целостность данных в выбранной СУБД. Например, если выбрать для банковской системы СУБД, не соответствующую требованиям ACID, то при сбоях во время записи данных возможна ситуация, когда деньги были списаны с одного счета, но не были зачислены на другой.
 2. Уровень производительности. На эффективность ПО может повлиять как производительность выбранной СУБД, так и возможность масштабирования СУБД. Даже если сама система будет легко масштабироваться, то неправильно выбранная СУБД может стать узким местом и не дать ПО обеспечить необходимую производительность.
-

3. Надежность. Под надежностью ПО понимают бесперебойное выполнение возлагаемых на него задач на заданных условиях в течение установленного времени [2]. На надежность программного обеспечения очень сильно влияет отказоустойчивость СУБД, ведь большинство приложений не смогут корректно функционировать без работоспособной базы данных.

Другие характеристики, определяющие качество ПО, согласно стандарту ГОСТ Р ИСО/МЭК 25010-2015, не будут рассматриваться, так как выбор SQL и NoSQL СУБД не оказывает существенного влияния на эти характеристики.

Под удобством разработки ПО подразумевается возможность быстрого изменения схемы данных приложения, что напрямую зависит от модели данных выбранной СУБД. Однако, данная характеристика сильно зависит от целей приложения и проработанности требований к нему. Например, если разработчикам необходимо за минимальный срок написать прототип приложения, то схема данных может сильно изменяться в процессе реализации данного прототипа. Более того, если планируется разработка полноценного приложения, но для него не сформулированы окончательные требования, то схема данных также может сильно изменяться в процессе реализации системы.

Таким образом, для выбора SQL и NoSQL СУБД при разработке приложений предлагаются следующие критерии оценки характеристик:

1. Модель данных – это совокупность структур данных и операций их обработки. Существуют такие модели данных, как реляционная модель, пары «ключ-значение», документная, колоночная, графовая и др. Выбор в сторону конкретной модели нужно делать в начале проектирования и отталкиваться от данных, которые будут храниться в базе, а также от операций над этими данными [3]. От выбранной модели данных будет зависеть, какой язык запросов будет использоваться, а также как быстро мы сможем менять схему данных в процессе разработки.
2. Целостность данных. Под данным критерием подразумевается соответствие СУБД требованиям ACID, наличие целостности сущностей (наличие уникального и неизменного первичного ключа для каждой записи в рамках одного отношения), целостности ссылок, а также наличие возможности создать другие правила ограничения целостности данных [4]. Необходимая целостность данных выбирается в зависимости от цели создания приложения и предъявленных к нему требований.
3. Масштабируемость. В зависимости от требований и предполагаемой нагрузки на приложение и СУБД следует рассмотреть наличие возможности не только вертикального масштабирования (увеличение вычислительных мощностей сервера), но и горизонтального масштабирования (увеличение количества серверов в кластере).
4. Отказоустойчивость. Под данным критерием подразумевается наличие возможности быстрого переключения роли основной базы данных на другой сервер в случае отказа первого. Данный критерий сильно связан с масштабируемостью СУБД и также зависит от предъявленных к приложению требований.

После определения критериев необходимо оценить характеристики SQL и NoSQL СУБД. Далее представлены характеристики для реляционных СУБД:

1. Модель данных. В реляционной модели данные представлены в виде таблиц. Перед созданием таблицы задается строгая схема хранимых данных, связи между таблицами также строго определены. В реляционных СУБД для работы с данными используется язык SQL. Также стоит обратить внимание на тот факт, что все изменения в структуре объектов приложения необходимо применять и к структуре таблиц, так как физическая структура данных не соответствует объектной модели приложения. Таким образом, реляционные СУБД идеальны для работы со структурированными данными, структура которых не подвержена частым изменениям, так как любое изменение структуры данных влечет за собой необходимость изменения всех данных, которые уже находятся в таблицах, что часто является сложной и трудоемкой задачей.

2. Целостность данных. В реляционных СУБД реализуются свойства ACID (атомарность, непротиворечивость, изолированность, надежность) [5] необходимые для обеспечения консистентности транзакций и данных. В таких системах поддерживается целостность сущностей, целостность ссылок, а также в них можно указывать другие правила ограничения целостности (например, ограничение атрибута, ограничение типа). Благодаря всем перечисленным характеристикам реляционные СУБД минимизируют риски потери или искажения данных и считаются очень надежными.
3. Масштабируемость. Данные СУБД поддерживают вертикальную масштабируемость (увеличение мощности сервера), однако, полноценное горизонтальное масштабирование не предусмотрено. Реляционные СУБД смогут обеспечить партиционирование (разделение) таблиц внутри одного сервера и репликацию баз данных, но не смогут обеспечить шардирование данных для параллельной обработки группой серверов. Если к приложению предъявлены требования о реализации шардирования данных, стоит рассмотреть возможность использования нереляционных СУБД.

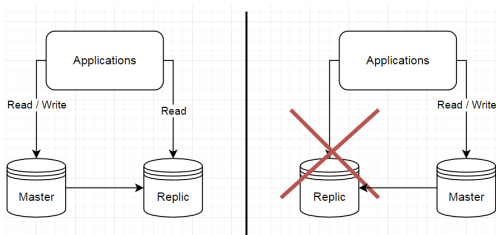


Рис. 1. Пример схемы репликации реляционной БД

4. Отказоустойчивость. Благодаря возможности репликации базы данных реляционные СУБД могут обеспечивать достаточно высокую отказоустойчивость. Так как при выходе из строя главного сервера его место займет сервер-реплика, пример работы данной схемы представлен на рисунке 1. Таким образом, после правильной конфигурации кластера серверов, можно добиться необходимой отказоустойчивости СУБД.

Далее определим характеристики нереляционных СУБД на основании полученных критериев:

1. Модель данных. В отличие от реляционных, в нереляционных базах данных схема данных является динамической и может меняться в любой момент времени. Физические объекты в NoSQL обычно можно хранить прямо в том виде, в котором с ними потом работает приложение. Базы данных NoSQL подходят для хранения больших объемов неструктурированной информации, а также хороши для быстрой разработки и тестирования гипотез. В них можно хранить данные любого типа и добавлять новые в процессе работы. Базы данных NoSQL предоставляют разнообразные модели данных, такие как пары «ключ-значение», документы и графы, оптимизированные под конкретные задачи для высокой производительности и масштабируемости.
2. Целостность данных. Свойства ACID не реализуются в полной мере в нереляционных СУБД, так как в них используется более гибкая модель, обеспечивающая шардирование данных. В NoSQL СУБД поддерживается целостность сущностей, однако, не поддерживается целостность ссылок. Таким образом, NoSQL СУБД не могут полностью обеспечить целостность данных.
3. Масштабируемость. Нереляционные СУБД имеют распределенную архитектуру, поэтому хорошо масштабируются не только вертикально, но и горизонтально, обеспечивая при этом высокую пропускную способность. Данные СУБД поддерживают репликацию и шардирование – разбиение базы данных на части, которые хранятся на разных серверах. На рисунке 2 приведен пример шардирования и репликации на примере нереляционной СУБД MongoDB.
4. Отказоустойчивость. Благодаря возможностям горизонтального масштабирования данные СУБД могут обеспечивать высокую отказоустойчивость. Правильно сконфигурированный кластер, состоящий из нескольких серверов, считается отказоустойчивым, так как в случае отказа одного из серверов, его данные не будут утеряны, а будут читаться и изменяться в соседних серверах. Это становится возможным благодаря шардированию и репликации данных.

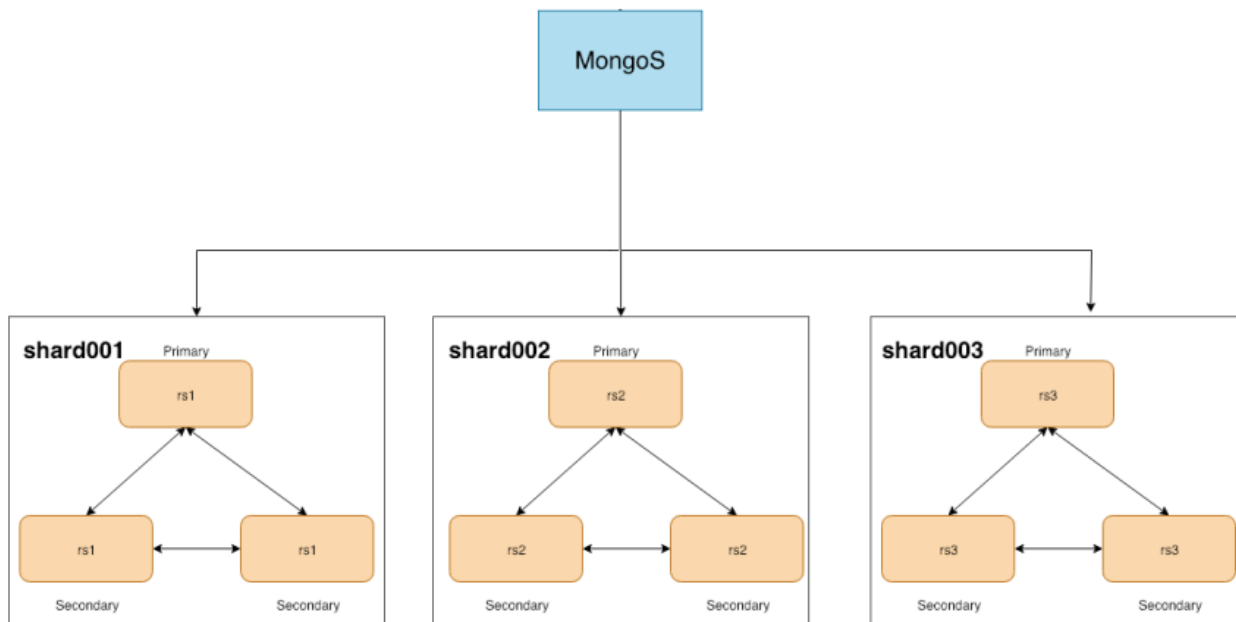


Рис. 2. Пример шардирования и репликации в NoSQL СУБД

Универсальных систем управления базами данных, которые подойдут под любые задачи, не существует. Для выбора SQL и NoSQL СУБД необходимо провести подробный анализ разрабатываемого приложения, сформулировать к нему требования и на основе данных требований и критериев, полученных в процессе написания работы, выбрать наиболее подходящую СУБД. Сейчас очень популярна практика, когда и SQL, и NoSQL СУБД используются в приложении для решения различных задач. Хотя NoSQL СУБД и стали популярными благодаря быстрдействию и хорошей масштабируемости, в определенных ситуациях стоит сделать выбор в стороне реляционных СУБД. В дальнейшем планируется спроектировать приложение и определить для него наиболее подходящую СУБД с помощью требований к приложению и критериев, полученных в ходе выполнения работы.

Литература

1. ГОСТ Р ИСО/МЭК 25010-2015. Требования и оценка качества систем и программного обеспечения (SQuaRE). Модели качества систем и программных продуктов. [Электронный ресурс]. – Режим доступа: <https://docs.cntd.ru/document/1200121069> (дата обращения: 10.12.2021).
2. Обеспечение качества программного обеспечения. [Электронный ресурс]. – Режим доступа: <https://www.a1qa.ru/blog/obespechenie-kachestva-programmnogo-obespecheniya/> (дата обращения: 10.12.2021).
3. Критерии выбора СУБД при создании информационных систем. [Электронный ресурс]. – Режим доступа: <http://www.interface.ru/home.asp?artId=2147> (дата обращения: 15.12.2021).
4. Целостность данных (Data integrity). [Электронный ресурс]. – Режим доступа: <https://wiki.loginom.ru/articles/data-integrity.html> (дата обращения: 16.12.2021).
5. Что такое NoSQL? [Электронный ресурс]. – Режим доступа: <https://aws.amazon.com/ru/nosql/> (дата обращения: 18.12.2021).

УДК 608.2

ИССЛЕДОВАНИЕ ИНФОРМАЦИОННЫХ ПРОЦЕССОВ ПОДДЕРЖКИ ВРАЧЕБНОЙ ПРАКТИКИ ВРАЧА СТАЦИОНАРА

Фокина С.Г.¹ (студент)

Научный руководитель – канд. физ.-мат. наук Орлов Г.М.¹

¹Университет ИТМО

e-mail: sun-vetochka@yandex.ru

Материал представляет собой краткий обзор существующих мировых практик по внедрению цифровых информационных процессов во врачебную практику, а также обозначение ключевых проблем, связанных с внедрением этих процессов. В статье рассмотрена роль врача, как источника первичной медицинской информации, на основе которой формируется база медицинских данных. Представлена гипотеза о причинах низких показателей полноты и качества, приведены факты, подтверждающие данную гипотезу. Материал, изложенный в статье, может быть использован медицинскими учреждениями, а также органами управления здравоохранением.

Ключевые слова: информационные процессы, электронная медицинская карта, телемедицина, форматно-логический контроль.

Одним из ключевых стратегических направлений в области цифровой трансформации является развитие цифровых информационных процессов в сфере здравоохранения. В 2022 г. из федерального бюджета на это планируется выделить 1,2 трлн руб. [1].

В Санкт–Петербурге благодаря системе федерального и регионального финансирования количество государственных медицинских организаций, которые приобрели медицинскую информационную систему (МИС), также провели модернизацию уже внедренной МИС, с целью ее приведения в соответствие с требованиями приказа Министерства здравоохранения РФ N 911н по передаче данных в региональную и федеральную информационные системы [2] достигло 100 процентов.

Данные, передаваемые МИС, стали основой для создания различных сервисов, таких как портал «Здоровье петербуржца», используемых жителями Санкт–Петербурга. Применяемые правительством РФ меры по борьбе с новой коронавирусной инфекцией явились толчком к резкому увеличению количества ежедневных посетителей личного кабинета на портале «Здоровье петербуржца» до 300 тысяч посетителей в день. Наиболее востребованным сервисом стал сервис, предоставляющий возможность записи на вакцинацию против данного вида инфекции [3]. В свою очередь, сервисы и требования к подсистемам формируют необходимость более структурировано вести данные в МИС.

При стремительном развитии цифровых процессов в сфере здравоохранения остаются нерешенными следующие проблемы:

- проблемы нормативно–правового регулирования, не определен порядок предоставления доступа к персональным данным пациента, до конца не определена ответственность всех участников цифровых информационных процессов;
 - проблема неравенства в доступности получения качественной телемедицинской консультации «врач–врач» – врачи отдаленных регионов нашей страны не имеют технической возможности использования таких технологий как телемедицина;
 - в ряде регионов не определен порядок оплаты услуг, которые оказываются с применением цифровых технологий. К таким услугам относится телемедицина [4];
-

- отсутствие стандартизации и унификации данных, передаваемых между различными сетями и системами [5];
- проблема, связанная с обеспечением кадрами, путем подготовки специалистов на уровне образовательных учреждений.

В решении данных проблем определяющую роль играет государство, как основной двигатель внедрения цифровых технологий, содействие которого необходимо для решения вопросов нормативно–правового регулирования, решения вопросов финансирования и более широкого применения цифровых технологий в сфере здравоохранения.

В развитии цифровизации здравоохранения одну из ключевых ролей играет врач, как один из первичных источников внесения медицинских данных в региональную информационную (РЕГИЗ) и федеральную информационную (ЕГИСЗ) системы, используемых как на уровне региона комитетом по здравоохранению (КЗ), так и на федеральном уровне министерством здравоохранения Российской Федерации (МЗ) (рис. 1).

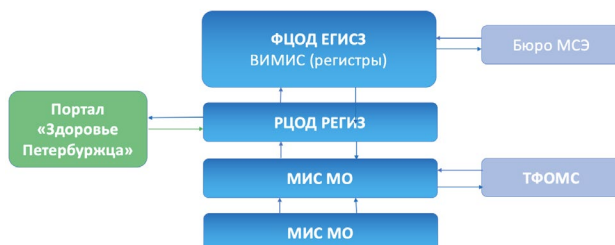


Рис. 1. Схема информационного взаимодействия между источниками и потребителями информации

На основе данных принимаются управленческие решения о пути развития, и, в случае некачественных данных, будут приняты неверные решения. Также от полноты и качества данных зависят такие показатели как:

- степень доверия со стороны ключевых потребителей – пациента и врача;
- получение услуг, оказываемых посредством цифровых технологий, пациентом (инвалидность через бюро медико-социальной экспертизы (МСЭ));
- получение оплаты учреждением за оказанные цифровые услуги.

С цифровизацией информационных процессов врачу стали доступны такие инструменты как: электронная медицинская карта (далее ЭМК) и телемедицинская консультация (ТМК).

ЭМК открывает перед врачом возможности:

- перехода от бумажного ведения истории болезни к электронному,
- получения оперативного доступа к клиническим рекомендациям,
- получения доступа к электронным медицинским данным пациента.

Также ЭМК является отличным инструментом для сбора, обработки, хранения и передачи данных о пациенте в региональную и федеральную информационные системы, которые в дальнейшем используются на горизонтальном уровне – врачами других организаций и на вертикальном уровне – медицинским информационным аналитическим центром (МИАЦ), КЗ, территориальным фондом обязательного медицинского страхования (ТФОМС), МЗ.

ТМК позволяет врачу перевести процессы из офлайн режима в онлайн. В режиме онлайн врачу стали доступны консультации в режимах «врач–врач», «врач–пациент», также появилась возможность прохождения обучения дистанционно. ТМК имеет хорошие перспективы развития в условиях эпидемиологической обстановки.



Рис. 2. Схема проверки медицинских данных

На сегодняшний день каждая система, участвующая в обмене электронной медицинской документации, имеет форматно–логический контроль, оценивающий правильность заполнения полей электронного документа (рис. 2). Несмотря на программные проверки, все же лучшим контроллером своих данных остается пациент, имеющий доступ к личному кабинету на портале «Здоровье Петербуржца», куда выгружаются результаты обследований, протоколы осмотра и иная электронная медицинская документация.

Для оценки полноты и качества медицинских данных, которые врач как оператор предоставляет в региональную и федеральную систему существуют методики контроля. В Санкт–Петербурге для оценки полноты и качества используется система рейтингов, доступная руководителям МО для контроля выполнения требований по передаче данных вышестоящих организаций. Методика расчета показателей передачи данных, на примере стационарных случаев медицинского обслуживания (СМО), происходит следующим образом:

– индикатор полноты ведения – отношение переданных стационарных случаев медицинского обслуживания (СМО), завершенных в отчетный период, к среднему количеству СМО за период предыдущего года;

– индикатор качества ведения – доля завершенных стационарных СМО с эпикризами (в случае расчета индикатора доли закрытых стационарных СМО с эпикризами) в общем количестве случаев, завершенных в отчётный период [6].

На графике (рис. 3) на примере одного из стационаров Санкт–Петербурга приведены показатели полноты и качества передаваемых данных, по этим данным можно проследить изменение показателей полноты и качества данных, передаваемых из стационара в региональную систему, а также сделать оценку эффективности использования информационных процессов в практике врача стационара. Как можно заметить на графике, интегральный рейтинг стационара, зависящий от индикаторов полноты и качества ведения СМО, за выбранный период не поднимался выше 60 процентов. И более того с августа 2021 года потерпел резкое падение с 60 до 40 процентов, после чего уже не поднимался выше 50 процентов. Эти данные дали почву для размышления и поиска как причины резкого снижения показателя в августе 2021 года, так и причины недостижения линии интегрального рейтинга отметки в 100 процентов. В результате проведенного исследования была выявлена причина резкого спада, связанная с изменением требований МЗ, а именно включением в методику расчета требования по передаче электронных документов за электронной подписью врача. Падение показателя показало недостаточную готовность учреждения к новым требованиям. Более глубокого анализа требует поиск причин того, что линия интегрального рейтинга МО по полноте и качеству ведения СМО колеблется от 60 до 40 процентов, не достигая показателя в 100 процентов. В связи с тем, что данный показатель зависит от индикаторов полноты и качества ведения СМО, следует обратить внимание на ведение этих данных в МИС, а также на требования к передаче данных из МИС в РЕГИЗ. Для проведения исследования использованы следующие инструменты – изучение нормативной базы, используемой врачом стационара в своей практике, для понимания общих требований к заводимым в систему медицинским данным, а также глубинные интервью с врачами одного из отделений стационара, для выявления возможных трудностей на местах, связанных с отсутствием оборудования, некорректной настройкой системы, также для оценки эффективности использования системы врачом.

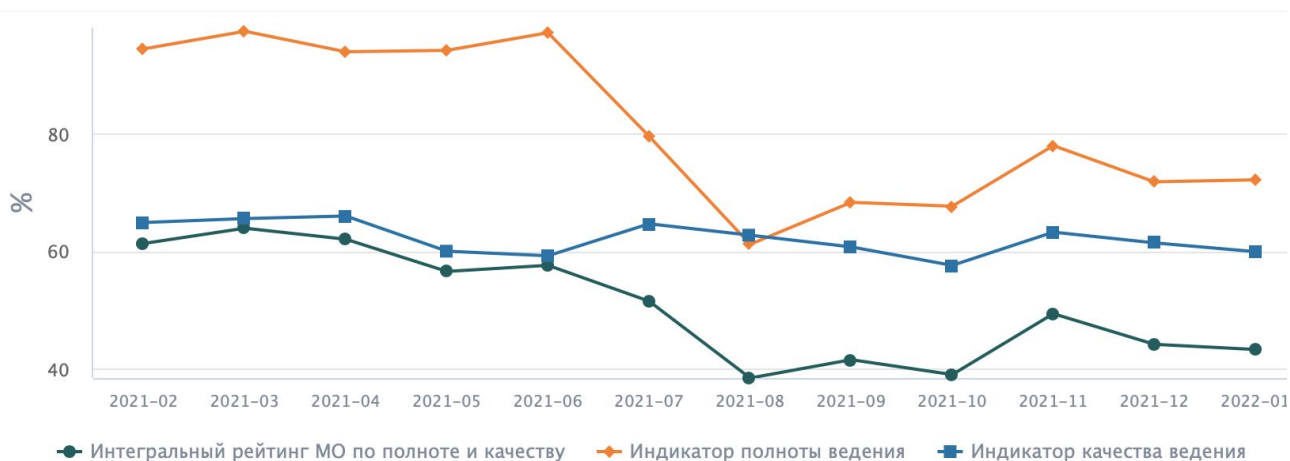


Рис. 3. Динамика показателей полноты и качества передаваемых данных

Автором статьи была выдвинута следующая гипотеза – интегральный рейтинг МО по полноте и качеству ведения СМО не достигает 100 процентов, поскольку врач не использует возможности МИС в полном объеме.

При изучении нормативной документации были выявлены следующие трудности, с которыми сталкивается врач в своей практике:

- использование большого количества нормативных документов, которые часто изменяются и дополняются правками, за которыми врач не может уследить [7];
- необходимость дублирования информации, в связи с тем, что новые сервисы не имеют интеграции с МИС.

В результате проведенных интервью с врачами неврологического отделения одного из стационаров г. Санкт–Петербурга, с использованием анкеты, содержащей вопросы, позволяющие определить эффективность использования МИС в своей работе, была получена следующая информация:

- у врача отсутствует полная информации о возможностях МИС учреждения;
- отсутствует форматно–логический контроль ошибок, учитывающий актуальные требования к сбору и передаче медицинских данных;
- в МИС отсутствуют поля протокола, необходимые врачу для автоматического формирования отчетов;
- врач ведет две истории болезни бумажную и электронную.

По работе этого отделения был проведен анализ данных по отказам страховых компаний, из-за несоответствия историй болезни стандартам лечения. Результат показал, что основные замечания были связаны с несоблюдением сроков предоставления услуг в рамках ведения стационарного случая.

В рамках работы была оценена динамика роста количества пациентов на этом отделении, так как эта информация также является причиной увеличения нагрузки на врача и, как следствие, является причиной низкий показателей полноты и качества передаваемых данных. В 2021 году количество пациентов увеличилось почти на 10 процентов относительно 2020 года.

Суммируя вышесказанное, можно сделать следующий вывод:

- врач не видит в МИС средство по снижению нагрузки, не использует ее возможности в полном объеме. Это в свою очередь, не только не ведет к снижению нагрузки на врача, но, наоборот, приводит к ее увеличению. Этот фактор оказывает влияние на полноту и качество ведения СМО врачом и их дальнейшей передачи в региональную и федеральную системы.

Дальнейший план исследования подразумевает следующие действия:

1. Оценить мотивацию и вовлеченность врачей в использовании информационных процессов, с учетом отсутствия юридической силы у ряда документов, передаваемых из МИС в РЕГИЗ, а также отсутствия оплаты определенных видов услуг в регионах РФ, для оценки степени их влияния на колебание показателей полноты и качества передаваемых данных.
2. Провести реинжиниринг существующих информационных процессов, настроить работу врачей отделений стационара по следующему алгоритму:
 - провести анализ существующих в системе информационных процессов «как есть»;
 - совместно с врачом отделения разработать правила проверки заведения данных в системе с учетом нормативных документов, которыми должен руководствоваться врач в своей врачебной практике;
 - перевести бумажные формы документов на уровне учреждения в электронные, обеспечить возможность их подписания электронной подписью для придания им юридической силы;
 - с использованием разработчиков настроить в системе «генератор отчетов», с возможностью оперативного сбора данных, внесенных в систему.
3. Разработать практические рекомендации, основанные на анализе внешних факторов, влияющих на полноту и качество передаваемых данных.

Литература

1. Калиновская Е. Расходы федерального бюджета на здравоохранение снизятся в 2022 году на 9%. [Электронный ресурс]. – Режим доступа: <https://medvestnik.ru/content/news/Rashody-federalnogo-budjeta-na-zdravooohranenie-snizyatsya-v-2022-godu-na-9.html> (дата обращения: 18.01.2022).
2. Приказ Министерства здравоохранения РФ от 24 декабря 2018 г. N 911н “Об утверждении Требований к государственным информационным системам в сфере здравоохранения субъектов Российской Федерации, медицинским информационным системам медицинских организаций и информационным системам фармацевтических организаций”. [Электронный ресурс]. – Режим доступа: <https://www.garant.ru/products/ipo/prime/doc/72117630/> (дата обращения: 18.02.2022).
3. 2021: Рост количества посетителей до 300 тыс. в день. [Электронный ресурс]. – Режим доступа: https://www.tadviser.ru/index.php/Продукт:Здоровье_петербуржца_%28gorzdrav.spb.ru%29 (дата обращения: 28.02.2022).
4. Третьякова Е.П. Правовые аспекты регулирования телемедицины // Цифровое право. – 2020. – Т.1. – №2. – С. 53–66.
5. Прилуков М.Д. Проблемы правового регулирования телемедицины. Российский и Международный опыт // Вестник Нижегородского университета им. НИ Лобачевского. – 2018. – №. 6. - С. 136–141.
6. Методика расчета рейтинга полноты и качества ведения ЭМК петербуржца. [Электронный ресурс]. – Режим доступа: <https://spbmiac.ru/wp-content/uploads/2019/06/Методика-расчета-рейтинга-полноты-и-качества.pdf> (дата обращения: 18.02.2022).. Информационно-аналитический модуль. [Электронный ресурс]. – Режим доступа: <https://iam.gorzdrav.spb.ru/ui/> (дата обращения: 09.03.2022).
7. Распоряжение Правительства РФ от 29 декабря 2021 г. № 3980-р об утверждении стратегического направления в области цифровой трансформации здравоохранения. [Электронный ресурс]. – Режим доступа: <https://www.garant.ru/products/ipo/prime/doc/403236631/> (дата обращения: 28.02.2022).

УДК 608.2

СРАВНИТЕЛЬНЫЙ АНАЛИЗ ФУНКЦИОНАЛЬНОСТИ ОТКРЫТОЙ ЧАСТИ ГОСУДАРСТВЕННЫХ ИНФОРМАЦИОННЫХ СИСТЕМ В ЧАСТИ ПУБЛИКАЦИИ ДАННЫХ ОБ ОБЪЕКТАХ ГОРОДСКОЙ СРЕДЫ

Хабирова А.Р.¹ (студент)

Научный руководитель – канд. экон. наук Кононова О.В.¹

¹*Университет ИТМО*

e-mail: lina.khabirova@gmail.com

В работе рассмотрена концепция открытых данных, а также проблемы и решения для работы с открытыми данными. Раскрыто понятие открытых данных с точки зрения Российского законодательства, отмечена актуальность и преимущества работы с данными. Проведен сравнительный анализ функциональных возможностей порталов открытых данных. Отмечено, что распространению открытых данных способствует комплекс существующих технических решений.

Ключевые слова: общедоступная информация, открытые данные, машиночитаемые данные, портал открытых данных, наборы данных, открытое правительство.

Внедрение механизмов «электронного правительства» и трансформация способов организации государственного управления, способствовали развитию идеологии открытых данных. Концепция открытых данных является одним из направлений реализации обеспечения доступа к информации. Первоначально идея получила распространение в академических кругах как движение, направленное на развитие открытых научных сообществ путем обеспечения свободного доступа к академическим данным, опубликованным в специальных цифровых хранилищах. Позже данная идея стала приобретать и политический смысл.

В основе концепция открытых данных лежит идея о свободном доступе государственных данных для групп заинтересованных пользователей с возможностью распространения в машиночитаемом формате, без ограничений авторского права и иных механизмов контроля. Так информационная открытость и расширение форм представления и использования данных о деятельности государственной власти становится приоритетными направлениями развития информационной политики многих государств. Открытые данные все больше признаются в качестве новой формы инфраструктуры, которая определяет то, как организуется взаимодействие правительства, предприятий и граждан в сетевом пространстве [1].

В соответствии с Федеральным законом «Об информации, информационных технологиях и о защите информации» государственные органы и органы местного самоуправления обязаны представлять заинтересованным лицам информацию о своей деятельности. Одной из допустимых форм является возможность размещения общедоступной информации, созданной в пределах полномочий органов государственной власти, путем публикации в сети Интернет. Согласно Российскому законодательству, под общедоступной информацией, размещаемой в форме открытых данных, является информация, «публикуемая обладателями в сети «Интернет» в формате, допускающем автоматизированную обработку без предварительных изменений человеком в целях повторного ее использования».

Возможность публикации данных в сети Интернет во многом должна обеспечить реализацию потребностей в информации для максимального числа конечных пользователей. Формирующаяся в текущий период база открытых данных становится базой для сервисов и

услуг как для коммерческой, так и некоммерческой сфер. Однако, отсутствие четких требований к размещению открытых данных и их публикация на разных информационных порталах сформировало ряд существенных недостатков: данные размещаются на различных ресурсах, что усложняет поиск данных, имеют разную форму публикации, требуют дополнительной обработки, обновляются не регулярно, часть данных дублируются на разных порталах, существует проблема разрозненности информации [2].

Согласно положениям Федерального закона от 27 июля 2006 г. № 149-ФЗ «Об информации, информационных технологиях и о защите информации», государственные данные, возникающие в рамках исполнения государственными ведомствами своих полномочий и предназначенные в том числе для межведомственного взаимодействия, размещаются в государственных информационных системах. Общедоступная информация, аккумулируемая данными системами, также подлежит размещению в форме открытых данных. Государственные информационные системы становятся ключевыми источниками государственных данных в России [3].

С целью решения проблем унификации и доступности наборов данных, способов формирования и публикации ведется системная работа по созданию единых порталов открытых данных. Применение унифицированных форм размещения информации в форме открытых данных обеспечит возможность объединять данные, получаемые и используемые на различных условиях, что позволяет расширять возможности автоматизированных систем, за счёт исключения ручных обработок. Отдельная информация (справочная), получаемая за счет совместимой семантики и согласованной структуры данных, позволяет расширить функциональность средств автоматизации. Формируемая инфраструктура открытых государственных данных позволит снизить запросы и обращения в соответствующие государственные, муниципальные органы.

Однако, перечень информации, требуемый к размещению на городских или национальных порталах открытых данных, ограничивается соответствующими нормативно правовыми актами. Так в Санкт-Петербурге перечни информации утверждены Постановлением правительства Санкт – Петербурга «О размещении информации в государственной информационной системе Санкт-Петербурга «Открытые данные Санкт-Петербурга». С целью представления общедоступной информации, не попадающей в данный перечень, но интересной конечным пользователям, был разработан открытый модуль государственной информационной системы «Объекты городской среды Санкт-Петербурга» (далее – ГИС ОГС). Отдельным преимуществом, по сравнению с порталом «Открытые данные», является скорость обновления данных. Поскольку данные в открытый модуль поступают из автоматизированной системы, входящий в состав ГИС ОГС, данные обновляются автоматически. Кроме того, на стадии разработки находятся сервисы, с помощью которых заинтересованные пользователи смогут подавать заявки на предоставление государственных услуг.

В рамках исследования возможностей модернизации функционала и интерфейса открытой части ГИС ОГС Санкт-Петербурга для повышения эффективности размещаемых и потенциально возможной к размещению данных проведен анализ порталов открытых данных: Портал открытых данных г. Москва, Портал открытых данных г. Санкт-Петербург, Портал открытых данных г. Казани.

Условия использования открытых данных, технологическая инфраструктура, требования к публикации реестра наборов данных и самих наборов регламентированы «Методическими рекомендациями по публикации открытых данных государственными органами и органами местного самоуправления, а также техническими требованиями к публикации открытых данных». В задачи портала открытых данных, как web-ресурса, входят хранение, включая предыдущие версии наборов данных при обновлении, сбор и предоставление доступа к размещенным данным. Сравнительный анализ позволяет сделать вывод о том, что в качестве основы для структуры и реализованного функционала порталы использовали методические рекомендации, поскольку порталы между собой схожи и соответствуют (хоть и не в полной мере) требованиям изложенных в рекомендациях.

Архитектура ресурсов позволяет отметить, что основной упор сделан на простоту взаимодействия с порталом. При сравнении порталов можно отметить, что наборы данных представлены в виде таблиц с понятной архитектурой. Доступ к открытым данным реализован при помощи двух типов интерфейсов. На порталах обеспечено представление данных в человекочитаемом представлении с использованием средств визуализации: для удобства пользователей реализованы возможности просмотра наборов данных на самом портале. Представлены возможности работы с документами в формате применим для обмена данными внутри сетей между машинами с разными архитектурами – машиночитаемое представление. Форматы, основанные на текстовых файлах, являются наиболее распространенными: CSV, XML. Интерфейс в машиночитаемом представлении делится на 2 типа: файловый и REST API [4].

Порталы обеспечивают доступ к метаданным открытых данных таких как: реестр наборов открытых данных, паспорт, структура, статистическая информация. Работа в реестрах должна обеспечивать навигацию по наборам и возможность извлекать ключевые параметры этих наборов. Сами наборы данных представлены в определенной иерархии с сохранением предыдущих версий, при наличии [5]. Что касается структуры, в большинстве случаев, это документ, описывающий как тот или иной набор данных необходимо отражать пользователю, либо данные, требуемые для машинной обработки. В функции паспорта набора открытых данных входит идентификация наборов данных в реестре. Сведения, содержащиеся в паспорте, включают в себя информацию о наборе данных, публикаторе, ссылки на набор данных и соответствующие версии, а также информацию о периодах публикации данных в наборе. Под статистической информацией подразумевают информацию, оставляемую пользователями: количество скачиваний или просмотров набора данных, оценки пользователей.

Кроме того, каждый портал представляет возможности работы с картографическим слоем, в случае если данные содержат геолокационную информацию.

Что касается открытой части ГИС ОГС, первичный обзор и тестирование работы с данными позволяет сделать вывод, что за основу взята схожая схема представления открытых данных, однако, функционал на текущий момент является ограниченным (таблица), а интерфейс слабо адаптирован для работы внешних пользователей.

Таблица. Функциональные возможности открытой части ГИС ОГС

Функциональные возможности ГИС ОГС
Работа с объектами на карте
Возможность работы с реестровыми записями
Обеспечен поиск информации по реестрам, контактам, разделам
Возможность обращения в службу поддержки
Карта
Выбор слоя
Переход к доступной информации о выбранном объекте
Скачивание связанных документов
Обеспечен поиск по адресу
Обеспечен поиск по доступным слоям
Возможность определения текущей геопозиции
Работа в полноэкранном режиме
Реестры
Возможность выбора соответствующего реестра через выпадающий список
Возможность отображение всех объектов реестра на карте
Возможность скачивания набора данных (только отображенных на странице). <i>Доступные форматы CSV, XLSX</i>
Возможность перехода с каждой реестровой записи на карту
Возможность настройки колонок (отображение только необходимых)
Переключение на формат расширенного (поиска по каждому столбцу)
Возможность быстрого поиска через строку «поиск»
Выбор количества отображаемых записей на странице
Переход по страницам

Проведенный обзор предметной области позволяет сделать вывод о том, что в настоящее время наблюдается устойчивая тенденция к повышению открытости государства, однако, сам процесс раскрытия данных значительно затягивается.

Стоит отметить, что достаточно высокий уровень развития вычислительной техники позволяет расширять список машиночитаемых форматов до мультимедийной информации, что может рассматриваться как потенциал к развитию порталов, в том числе открытого модуля ГИС ОГС. Развитие порталов и систем, размещающих открытые данные требуется для реализации потенциала открытых данных. Для этого необходимо, чтобы порталы обеспечивали возможность легкого поиска данных и работы с ними, а также обладали понятной структурой. Доступность и удобство работы с данными в свою очередь стимулируют развитие и реализацию идей по созданию социально значимых сервисов, тем самым повышая гибкость и эффективность экономики, улучшая благосостояние граждан и качество жизни.

В рамках дальнейшего исследования предполагается подробное рассмотрение архитектур порталов открытых данных. Основой для проведения оценки функционала должна стать разработка предложений и обоснование системы показателей оценки. С целью определения возможностей для развития интерфейса открытой части ГИС ОГС планируется проведение обзора особенностей и технологий реализации пользовательского интерфейса. На основе полученных результатов планируется осуществить проектирование открытой части ГИС ОГС на примере выбранных объектов городской среды, а также разработка ряда рекомендаций по модернизации функционала и интерфейса открытой части ГИС ОГС Санкт-Петербурга.

Литература

1. Околеснова О.А. Открытые данные как инновационный механизм раскрытия информации // Государственный аудит. Право. Экономика. – 2017. – № 2 – С. 64–67.
2. Verhulst S., Young A. Open data in developing economies. [Электронный ресурс]. – Режим доступа <http://odimpart.org/developingeconomies.html> (дата обращения: 25.02.2022)
3. Оценка открытости государственных информационных систем: аналитический отчет. – М.: Изд-во ЦПУР. – 2020. – С. 8.
4. Староверов В.В. Открытые данные // International Journal of Open Information Technologies. – 2016. – № 4. – С. 62–67.
5. Ирхин Ю.В. Электронное правительство и общество: Мировые реалии и Россия // Социологические исследования. – 2020. – №1. – С. 73–82.

УДК 004.048

АНАЛИЗ ВОЗМОЖНОСТИ ПРИМЕНЕНИЯ МАШИННОГО ОБУЧЕНИЯ И НЕЙРОННЫХ СЕТЕЙ ДЛЯ ПРОЕКТИРОВАНИЯ И ОПТИМИЗАЦИИ ТОРГОВЫХ РОБОТОВ

Чагина П.А.¹ (студент)

Научный руководитель – канд. техн. наук, доцент Ананченко И.В.^{1,2}

¹Университет ИТМО

¹Санкт-Петербургский государственный технологический институт (технический университет)

e-mail: chagina.pa@mail.ru, anantchenko@yandex.ru

В работе оцениваются возможности применения нейронных сетей разных типов в алгоритмах торговых роботов для финансового валютного рынка Форекс, анализируются возможности нейронных сетей разных типов, используемых для прогнозирования, их достоинства и недостатки, оцениваются перспективы использования методов машинного обучения в обучении нейронных сетей для синтеза нейросетевых моделей, применяемых для разработки и оптимизации торговых роботов.

Ключевые слова: нейронная сеть, нейрон, торговый робот, Форекс, машинное обучение, торговые стратегии.

Невозможно представить развитие мировой экономики без финансовых рынков. Они являются важным элементом для развития также и России, где популярен рынок Forex (Форекс), который занимает 90% мирового рынка капитала. Профессионалы, торгующие на рынке – трейдеры, должны обладать умением концентрироваться, быть психологически устойчивыми и быть способными справляться со стрессом, усталостью и личными эмоциями. Отсутствие данных качеств негативно сказывается на качестве их работы. Данную проблему можно решить использованием автоматических торговых систем [1, 2]. Нейронные сети эффективны при решении сложнейших задач. К их достоинствам можно отнести умение самообучаться, что выделяет их. В настоящее время отсутствуют готовые решения (проекты или многофункциональные торговые платформы) с применением нейронных сетей для разработки нейросетевых торговых роботов.

Проведенный анализ представленной тематической информации на платформе российской научной электронной библиотеки (<https://www.elibrary.ru/>) позволил выявить, что треть литературных источников связаны именно с нейронными сетями и относятся к 2017 – 2021 годам, что видно из таблицы и говорит об активном интересе к данной теме и ее актуальности. Компьютерную нейронную сеть можно определить как массив искусственных нейронов, имеющих связь между собой.

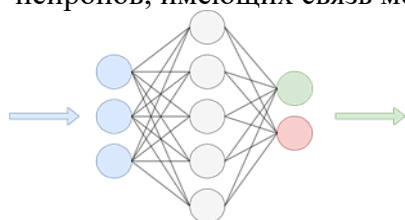


Рис. 1. Модель нейронной сети

На рисунке 1 представлена схема простой нейронной сети. Нейроны изображены в виде кружков: синим цветом обозначены нейроны, на которые подаются входные данные, серым цветом обозначен скрытый слой, а красным и зеленым нейроны, дающие выходные данные. Стрелками обозначен поток данных. В скрытом слое нейронов выполняется аналитическая деятельность. Количество нейронов для скрытого слоя определяется в соответствии со сложностью решаемой задачи.

Таблица 1. Результат обзора актуальных источников

Ключевые слова	Общее количество публикаций	Количество публикаций за период 2017–2021 гг.
нейронные сети, торговые роботы, валютные рынки	483	159
торговые роботы, валютный рынок, трейдинг	210	138
машинное обучение, торговые роботы, Форекс	95	12
торговые роботы, торговый бот, машинное обучение, Форекс	7	4
торговые роботы, нейронные сети, машинное обучение, Форекс	18	14
торговые роботы, трейдинг, Форекс, валютная биржа	32	66

Нейрон – несложный процесс преобразования входного сигнала в выходной сигнал. Выбранная модель искусственного нейрона относится к модели Мак – Каллока – Питса, в модели x_n – множество данных, w_i – множество весов.

Для построения нейронной сети первоначально необходимо спроектировать нейрон. Из набора таких нейронов будет состоять нейронная сеть (рис. 2). На вход нейрон получает множество данных p_i и множество коэффициентов w_i , а на выходе происходит проверка значения активационной функции, в результате нейрон может быть активирован или нет.

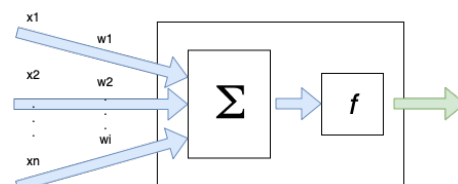


Рис. 2. Математическая модель нейрона

Помимо наличия модели нейрона при проектировании нейронной сети необходимо спроектировать скрытые слои и обозначить количество нейронов во входном и выходном слое [3]. Нейросеть требует проведения обучения перед введением ее в работу. Обучение нейросети проводится при помощи алгоритмов и методов машинного обучения: обучение с учителем, обучение без учителя, обучение с подкреплением.

В рамках проводимой работы нами были рассмотрены следующие нейросети:

- Perceptron – сеть, состоящая из сенсоров, ассоциативных и реагирующих элементов, основана на математической модели, имитирующей восприятие информации мозгом.
- Feed Forward Neural Networks – сеть, работающая на основе метода обратного распространения ошибки.
- Auto Encoder Network – сеть, предназначенная для снижения размерности и фильтрации шумов.
- Convolutional Neural Network – сеть, обладающая устойчивостью при работе с меняющимися изображениями, данное свойство нейронной сети позволяет фиксировать изменения на графиках валют.
- Recurrent Neural Network – сеть может работать с нефиксированным объемом данных, выступающим в роли краткосрочной памяти, способна решать проблему минимизации затраченных вычислительных ресурсов

Торговый робот – компьютерная программа, в задачи которой входит совершать финансовые операции купли-продажи финансовых инструментов, основываясь на решениях, которые торговый робот принял в рамках работы алгоритма самостоятельно. Основной целью торгового робота является автоматизация процесса торговли с целью получения прибыли [4].

Бессистемная торговля на рынке Форекс приводит к потерям капитала. Если продолжительное время наблюдать за движением цены, можно выявить некоторые последовательности, закономерности, зависимости, которые могут быть положены в основу разрабатываемых торговых стратегий. Торговый робот разрабатывается на основе торговой стратегии, с помощью которой формируется решение во время торговли. Торговая стратегия – это определенные правила для систематизации торгового процесса. Данный свод правил определяет момент входа и выхода из сделки, а также ситуации, когда процесс совершения

торговых операций нужно отложить [5]. Торговая стратегия помогает исключить влияние человеческого фактора (эмоции, усталость и т.д.). Торговая стратегия, своеобразный план заключения сделок, ведения торговли на рынке Форекс.

На выбор торговой стратегии влияет:

- логическое обоснование ее применения;
- конкретная валютная пара рубль/доллар или евро/доллар;
- правила входа (сигналы на открытие позиции);
- правила выхода;
- объем торгового лота.

Разработка торгового робота – сложный процесс из-за необходимости учета влияния хаотической составляющей, влияющей на изменения курсов финансовых инструментов, используемых для торговли на валютных рынках. Использование нейронных сетей и алгоритмов машинного обучения для проектирования и разработки торговых роботов позволит автоматизировать процесс подстройки параметров торговых роботов с учетом динамики изменения рынка.

Литература

1. Пестерева Л.И. Что такое биржевой робот: принцип работы, виды, плюсы и минусы, топ 5 лучших. [Электронный ресурс]. – Режим доступа: <https://greedisgood.one/robot-dlya-torgovli-na-birzhe> (дата обращения: 18.02.2022).
2. Ананченко И.В., Чагина П.А. Применение нейронных сетей в алгоритмах торговых роботов на рынке Форекс // Приоритетные направления инновационной деятельности в промышленности. Сб. научных статей III международной научной конференции. Казань. – 2021. – С. 133–134.
3. Юрков Д.А. Синтез нейронных сетей с фрактальной структурой // Восточно-европейский журнал передовых технологий. – 2009. – Т.4– № 3(40). – С.39– 44.
4. Буккунов С.В., Климин П.Ю. Автоматизированная торговая система для работы на финансовых рынках // Инженерный вестник Дона. – 2019. – №1(6). – С. 87–92.
5. Ананченко И.В., Чагина П.А. Практическое применение нейронных сетей: проблемы и решения // European research. Сб. статей XXIX Международной научно-практической конференции. Пенза. – 2020. – С. 27–30.

УДК 314.04; 159.9

ПРИЛОЖЕНИЯ МОБИЛЬНОГО ЗДРАВООХРАНЕНИЯ: ИССЛЕДОВАНИЕ МОТИВАЦИИ РУССКОГОВОРЯЩИХ ПОЛЬЗОВАТЕЛЕЙ

Чижи́к А.В.¹, Мельникова С.А.¹ (студент), Ковальчук С.В.¹

¹*Университет ИТМО*

e-mail: melnikova-rostovskaya@yandex.ru

Статья посвящена описанию исследования уровня готовности населения Санкт-Петербурга к использованию телемедицинских технологий. Авторы приводят данные опроса, проведенного в декабре 2021 года, а также результаты автоматического анализа отзывов на медицинские мобильные приложения. В качестве результата исследования авторы приводят построенную мотивационную модель, которая учитывает особенности поведения пользователя на этапе установки приложения и в фазе первых недель использования.

Ключевые слова: мотивационная модель, mHealth, телемедицина, UX, тональность отзыва.

Электронное (eHealth) и мобильное (mHealth) здравоохранение становятся важными компонентами общей системы здравоохранения. Это происходит под влиянием многих факторов, среди которых в том числе и потребность в ускорении процессов формального взаимодействия с пациентами (запись на прием, обмен документами и т.п.). Однако, несмотря на неоспоримые преимущества интеграции подобных сервисов, для того чтобы электронные медицинские услуги были действительно эффективными для системы здравоохранения, они должны входить в систему мировоззрений обывателя, иначе при необходимости решения проблем лечения и профилактики здоровья человек не может полностью использовать потенциал ИКТ в этой сфере [1]. В рамках данного исследования мы стремились проанализировать уровень готовности россиян к интеграции философии телемедицины в свои повседневные практики. Так как многие практики eHealth в той или иной степени вошли в повседневные практики населения, то было решено сфокусироваться на анализе ситуации с мотивированностью граждан использовать мобильное здравоохранение, которое на данный момент развито намного меньше. В связи с этим было выделено три основных цели исследования:

1. Выяснить доступность интернета.
2. Проанализировать, какими мобильными устройствами пользуются люди чаще всего (и с какими целями).
3. Понять текущее мировоззрение (в том числе проблемы, которые есть у людей, и цели, которые они готовы решать с помощью mHealth).

Дизайн исследования предполагал использование двух взаимодополняющих методов: классический социологический опрос (для выявления основных трендов) и анализ публичных данных из интернета (для апробации возможности их использования при построении мотивационных моделей и, соответственно, анализа проблем, с которыми встречаются пользователи).

Опрос был проведен в декабре 2021 года среди жителей города Санкт-Петербург и проходил 14 дней, было опрошено 95 респондентов. Отметим, что на вопрос «Является ли для вас доступным интернет?» 90% респондентов ответили «да», оставшиеся 10% (ответ «нет») уточнили, что используют для выхода в сеть смартфон, и интернет-соединение,

предоставляемое операторами сотовой связи, имеет недостатки: дорогой тариф или неполная зона покрытия.

Было стратегически важно для дальнейшего хода исследования получить благодаря опросу репрезентативные данные, что на наш взгляд, обеспечивается за счет формирования равнозначных по количеству возрастных выборок (рис. 1) и охватом профессиональных сфер (рис. 2).

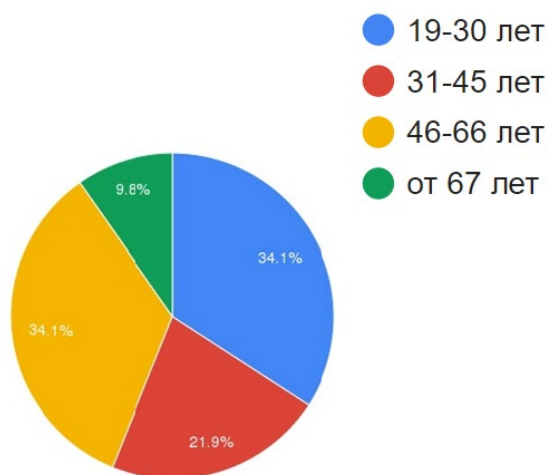


Рис. 1. Возрастной состав респондентов

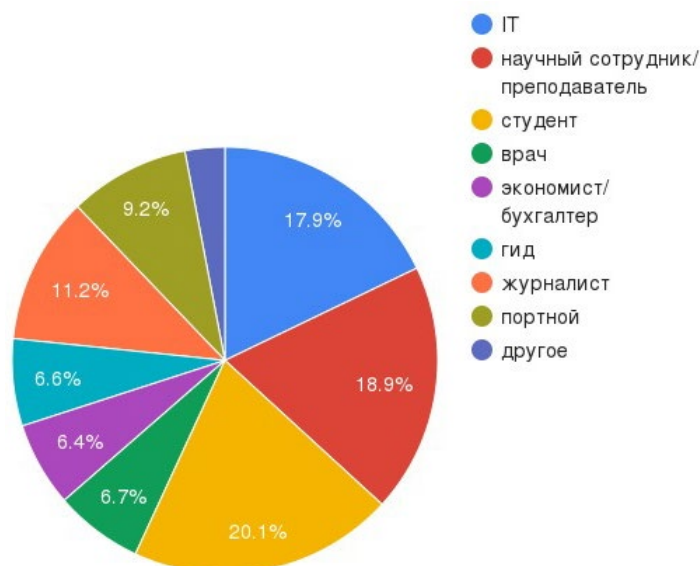


Рис. 2. Сфера профессиональной деятельности респондентов

Использование mHealth-приложений предполагает по умолчанию наличие смартфона с достаточным объемом памяти (также, например, при разработке необходимо учитывать, какая операционная система чаще всего используется пользователями и т.п.) [2]. В базе схема взаимодействия пользователя и приложения (и, при наличии, устройства, собирающего данные) может быть изображена следующем образом (рис. 3):

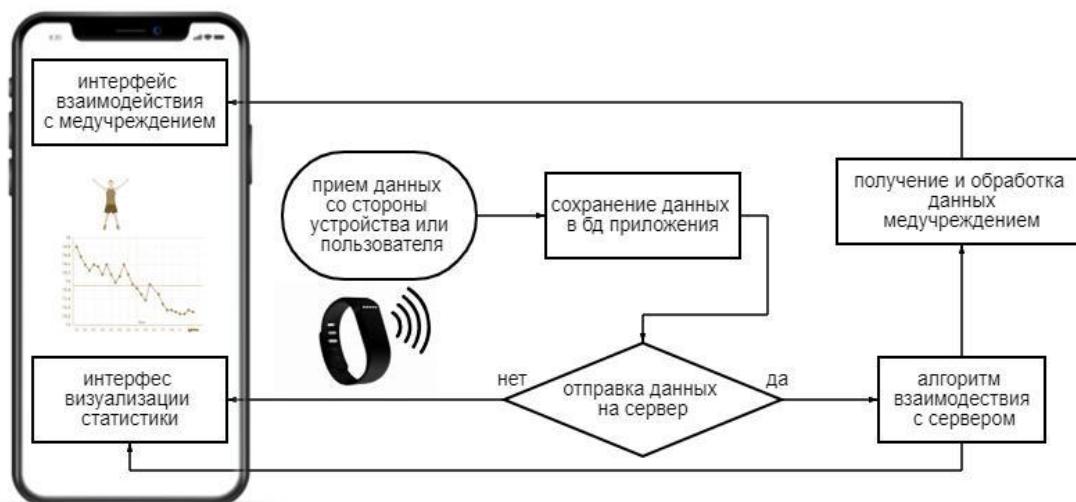
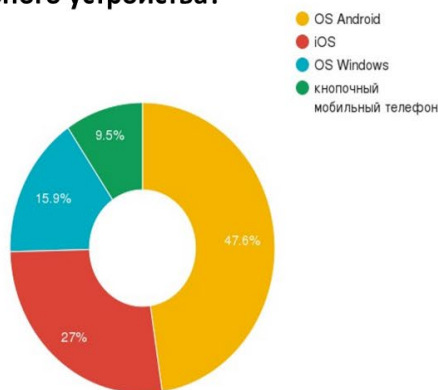


Рис. 3. Усредненный алгоритм взаимодействия пользователя и мобильного медицинского приложения

Нами было выяснено, что среди респондентов 47% использует OS Android, 27% – iOS, 75% опрошенных не собираются менять мобильное устройство, 17% на момент опроса планировали покупку нового мобильного устройства с той же операционной системой, что и используемое в текущем. Соответственно, при разработке стоит начинать с охвата целевой аудитории, которая использует Android и расширяться до iOS на следующем шаге. Стоит также отметить, что 9,5% из числа респондентов используют кнопочный телефон (представители четвертой возрастной группы – 67–84 года). В своих повседневных практиках

большинство респондентов использует следующий функционал мобильного устройства: сообщения, звонки, мессенджеры социальных сетей, приложения развлекательного характера, бизнес-приложения, обучающие приложения (в порядке убывания востребованности) (рис. 4).

Какая операционная система у вашего мобильного устройства?



Собираетесь ли вы менять в ближайшее время ваше мобильное устройство?



Какие функции вы чаще всего используете в своем мобильном устройстве?

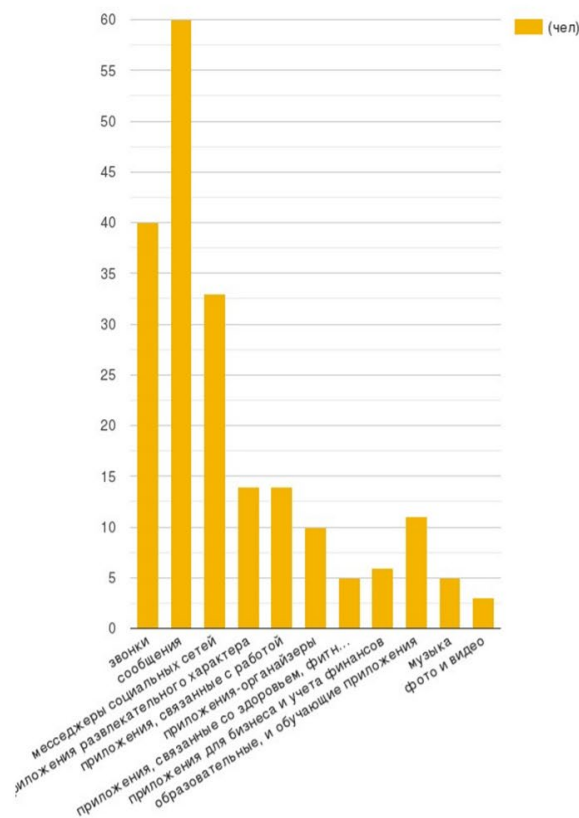


Рис. 4. Информация о предпочтениях при выборе операционной системы, наиболее востребованных функциях, а также о планах на смену мобильного устройства

Исходя из присутствующего интереса к функционалу коммуникации, можно выдвинуть гипотезу, что мобильный сервис в сфере медицины может быть выполнен или как чат-бот на платформе одной из социальных сетей (включая сервисы обмена мгновенными сообщениями), или в виде отдельного приложения, которое в качестве внутренних и внешних стимулов будет иметь триггеры похожие на те, что присутствуют в выше названных приложениях из списка активно используемых респондентами.

Важным пунктом нашего исследования был анализ степени посвященности жителей Санкт-Петербурга с концепцией телемедицины. Ниже представлен набор диаграмм, демонстрирующих владение респондентов термином «телемедицина» (рис. 5). Первые три диаграммы демонстрируют ситуацию для отдельно взятых возрастных групп, четвертая показывает общий тренд вовлеченности общества в это понятие.

Если рассматривать усредненный результат (четвертая диаграмма на рис. 5), то более половины опрошенных утверждает, что знает, о чем идет речь, однако, обращаясь к сегментации по возрастным группам, видна тенденция отсутствия владения данным термином среди людей старше 65 лет (третья диаграмма), также можно констатировать, что наибольшая осведомленность характерна для молодого сегмента респондентов (люди до 45 лет). Мы попробовали переформулировать вопрос (рис. 6) – первые четыре диаграммы демонстрируют тенденции в каждой из возрастных групп, пятая диаграмма – общая ситуация по итогу сложения всех ответов респондентов.



Рис. 5. Ответ респондентов на вопрос «Знаете ли вы, что такое телемедицина?»

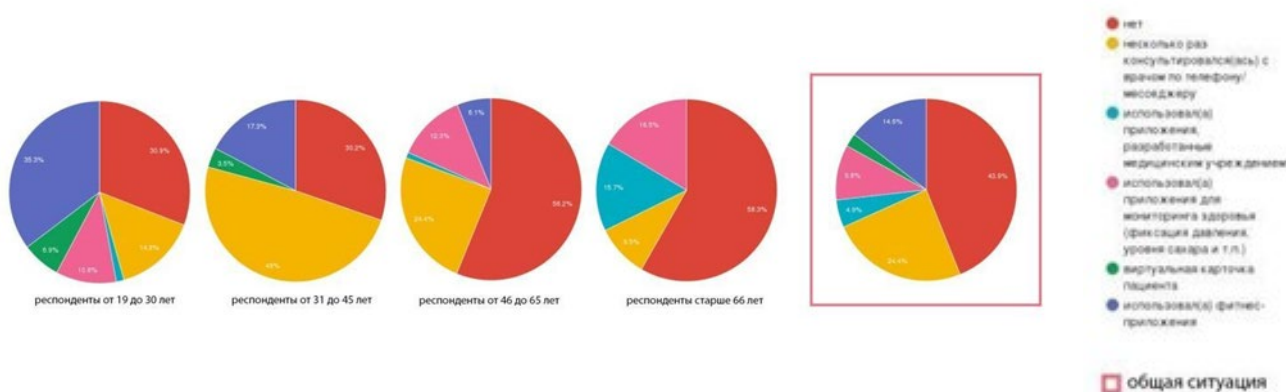


Рис. 6. Ответ респондентов на вопрос «Приходилось ли Вам пользоваться телемедицинскими технологиями?»

Исходя из продемонстрированных диаграмм, можно сделать вывод, что действительно около 43% из общего числа опрошенных не пользовались на данный момент сервисами, относящимися к телемедицине. В целом видны основные тренды поведения индивидов. Использование фитнес-приложений актуально только для самой молодой возрастной группы (19 – 30 лет). Если обратиться к людям среднего возраста (31 – 45 лет) видна тенденция к коммуникации с врачом в дистанционном формате. Для возрастной группы 46 – 65 лет помимо общей апатии по отношению к использованию телемедицинских технологий заметен, тем не менее, тренд на использование приложений для мониторинга отдельных показателей физического состояния. Можно сделать вывод, что на данный момент общество не находится в состоянии принятия телемедицины как привычной практики, во многом, на наш взгляд, это связано с отсутствием понимания, что такие сервисы существуют и функционируют. Из этой части исследования вытекают основные мотивационные стимулы к использованию мобильных медицинских приложений с разделением на возрастные сегменты:

- 19 – 30 лет: поддержание здоровья (фитнес) и дистанционный контакт с врачом;
- 31 – 45 лет: дистанционный контакт с врачом и поддержание здоровья (фитнес);
- 46 – 65 лет: дистанционный контакт с врачом и поддержание здоровья (мониторинг отдельных показателей);
- старше 66 лет: поддержание здоровья (мониторинг отдельных показателей) и приложения, разработанные медицинским учреждением.

Опрос показал, что возможности дистанционной связи с врачом интересуют значительное число респондентов, однако, понимают люди под этим очень разные сущности: от возможности получать дистанционно направления на обследования до регулярной связи с лечащим врачом с целью сообщения сведений о своем самочувствии в процессе лечения. Очевидно, что в зависимости от масштабности ожиданий оценка приложений с функционалом дистанционной коммуникации с медицинским учреждением будет разной.

В ходе исследования было выяснено, что в текущем есть сформированная привычка записываться к врачам с desktop-устройств. Причины для этого были выявлены разные (удобство заполнения форм с компьютера, недостаточно удобный интерфейс мобильных приложений и т.п.), но, на наш взгляд, важно, что сама проблема записи на прием решена

(паттерн поведения присутствует) и, таким образом, функционал мобильных приложений на данный момент имеет смысл видеть в продумывании стратегий коммуникации с врачом и в разработке алгоритмов мониторинга здоровья.

Далее мы провели пилотное исследование отзывов пользователей мобильных приложений (Google Store). Выбрали 4 категории приложений: мониторинг давления и пульса (288 шт.); трекеры сахара (373 шт.); приложения для напоминания о приеме лекарств (382 шт.); запись к врачу (292 шт.). Судя по оценкам и текстовым отзывам хуже всего ситуация обстоит с записью и коммуникацией с врачом. Отметим, что это логично, так как для нормального функционирования подобных приложений необходимо централизованно модернизировать медицинскую систему, а среди приложений, попавших в анализ, очень мало разработок вообще имеет статус государственных (то есть напрямую относящихся в текущие реалии клиник).

Менее 25% из общего числа приложений публикуют в своих профайлах информацию о сохранении персональных данных пользователя (то есть поясняют до момента установки, какими данными приложение будет пользоваться, куда и в каком виде они будут сохраняться). При этом среди приложений типа «запись к врачу» 83,5% разработок не снабжены таким описанием, что является достаточно странной тенденцией, так как наибольший оборот конфиденциальной информации может быть предусмотрен именно в них. При анализе полного описания приложений было выявлено, что для органайзеров по приему лекарств характерно наличие функции мониторинга отдельных физических показателей пользователя (чаще всего – давление и пульс). Была найдена корреляция высокого рейтинга и данного факта. При этом зеркальной ситуации не наблюдается при анализе приложений для мониторинга здоровья: дополнительных функций в этой категории разработок практически нигде не анонсировано, и влияния на рейтинг это не оказывает. То есть для людей оказывается удобным иметь полифункциональное приложение в случае, когда основной мотивацией к установке является органайзер по приему лекарств, а обратной зависимости нет.

Мы проанализировали корреляцию оценки в баллах и тональности отзыва, оказалось, что чем ниже оценка, тем нейтральнее отзыв, рисунок 7. Очевидно, что именно эти отзывы могут содержать наиболее важную информацию о потребностях в доработки UX-дизайна и функционала.

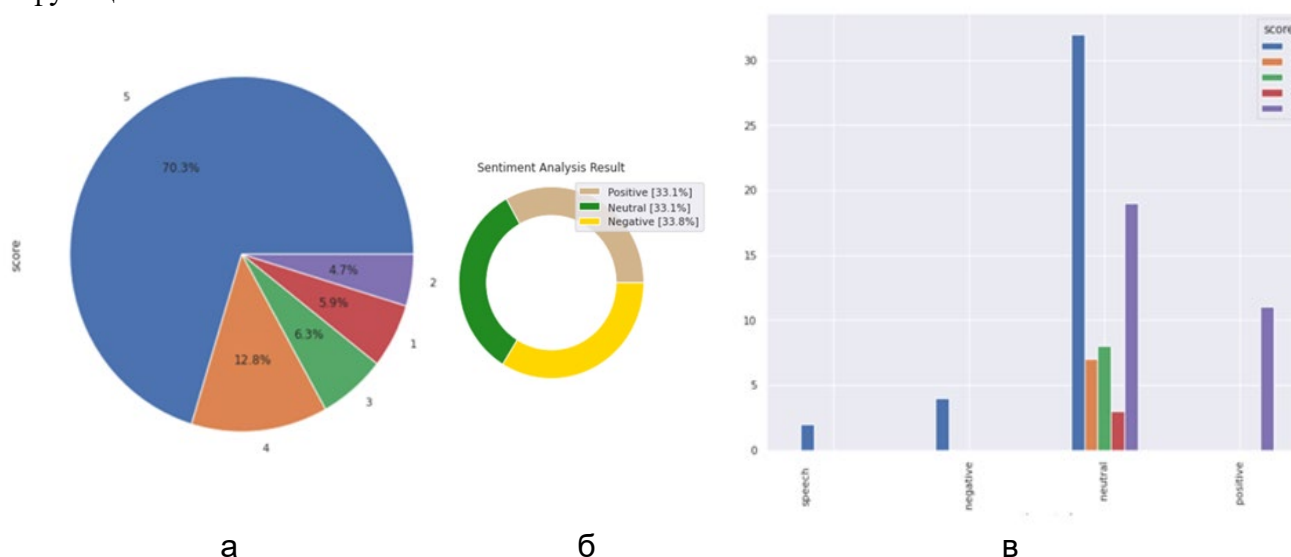


Рис. 7. Диаграммы: а) соотношение оценок пользователей для приложения Medisafe; б) оценка тональности текстовых отзывов; в) соотнесение числовой оценки и тональности отзывов

В соответствии с гипотезой далее мы проанализировали внутреннее содержание нейтральных отзывов с низкой оценкой. В частности, для приложения Medisafe было выявлено, что пользователей беспокоит концепция авторизации, а также методы выставления дозировок и времени начала приема препарата.

Из проведенного исследования стало понятно, что проблема мотивации пользователей к использованию мобильного приложения должна разделяться на две подпроблемы [3]:

1. Мотивация индивида к установке приложения.
2. Его удержание (длительное использование приложения).

Системы мониторинга показателей и органайзеры по приему лекарств с точки зрения побуждения пользователя к действию (установки приложения) схожи, так как представляют собой виртуального помощника (основная функция фиксация данных и/или напоминание о действии).

В качестве агрегированного результата исследования нами была разработана следующая модель мотивации (рис. 8).

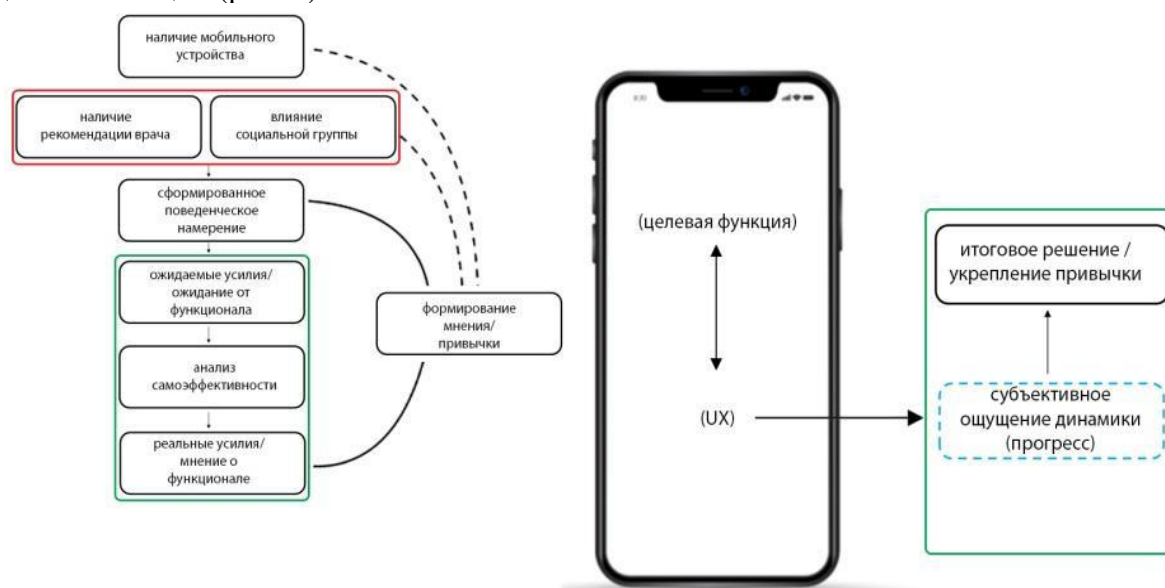


Рис. 8. Модель мотивации

Такая модель предполагает интеграцию атрибута автономии и учет стремления индивида к взаимодействию с социальной группой из теории самодетерминации в модель формирования привычки, что позволяет заменить фазы вознаграждения и инвестиций на актуальные для сферы здоровья категории. Модель включает в себя полный цикл взаимодействия пользователя приложением: мотивация к установке мобильного медицинского приложения, эмоциональные фазы первых взаимодействий, формирование решения (об отказе или о продолжении использования).

Литература

1. Данилов А.Б., Данилов Ал.Б. Биопсихосоциокультурная модель и хроническая боль // Современная терапия в психиатрии и неврологии. – 2013. – №1. – С. 30–36.
2. Terry M. Medical apps for smartphones. *Telemed J E Health* 2010. – 16. – Pp. 17–22.
3. Manning M. The effects of subjective norms on behaviour in the theory of planned behavior: a meta-analysis // *British Journal of Social Psychology*. – 2009. – V.48. – №4. – Pp. 649–705.

УДК 379.8

РАЗРАБОТКА ПРИЛОЖЕНИЯ ДЛЯ ИНТЕРАКТИВНЫХ ГОЛОГРАФИЧЕСКИХ ВЫСТАВОК

Шаров Д.Д.¹ (аспирант), Бурак П.В.¹ (студент)

Научный руководитель – д-р техн. наук, профессор Стафеев С.К.¹

¹Университет ИТМО

e-mail: danyelsharov@gmail.com

Статья представляет подход к формированию цифровой образовательной среды, способствующей повышению эффективности функционирования интерактивных образовательных выставок, посвященных голографии. Развитие подобных просветительных проектов в начале 2010-х годов способствовало росту интереса к голографическим технологиям, а представленная в докладе тематика исследования, по мнению авторов будет способствовать росту интереса и в новом десятилетии.

Ключевые слова: голография, интерактивные голографические выставки, онтология, принятие решений, DSS.

Разрабатываемые центром выставочного проектирования Университета ИТМО интерактивные голографические выставки сочетают в себе голографические и интерактивные экспонаты. Такой подход позволяет визуально познакомить посетителя с такой технологией как голография, а также дать ему представление о способах создания голограмм и физических явлениях, лежащих в их основе. В последние годы, согласно статистике посещения Музея оптики и данным о поисковых запросах пользователей Google, интерес к подобным выставкам начал нарастать.

N, шт/мес.

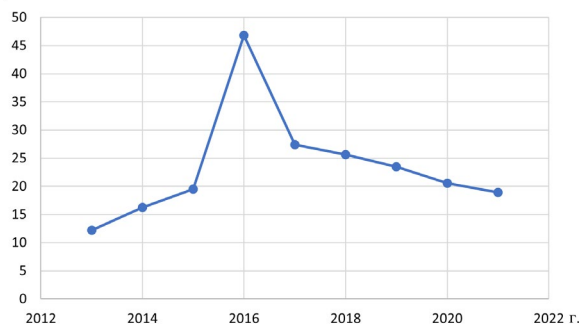


Рис. 4. Статистика поисковых запросов по слову «голограмма» с 2013 года по н.в.

На приведенном графике (рис. 1) приведена наиболее стабильная часть всей истории поисковых запросов по слову голограмма. Начиная с 2013 года, в интернете, интерес к голографическим технологиям стал стабильно ненулевым, но, тем не менее, невысоким. По этой причине начиная с 2015 года Музеем оптики Университета ИТМО стали регулярно проводиться выставочные мероприятия, целью которых является развитие интереса молодого поколения к данной технологии, привлечение потенциальных абитуриентов в Университет ИТМО, а также популяризация науки.

По мере проведения выставок стало заметно, что выставочный формат способствует развитию знаний и интереса к данной области. Более того, начиная с 2015 года стало возрастать число интерактивных голографических выставок, однако, на тот момент всё ещё не удавалось достоверно оценить эффективность таких мероприятий. Первый анализ был проведён в 2019 году [1] и позволил определить аудиторию посетителей музея и их интересы. Эти данные были учтены при формировании выставки в Шанхайском музее науки и технологий (SSTM), которая оказалась самой масштабной голографической выставкой, которую посетило порядка 400 тысяч посетителей.

Выставка в SSTM подтвердила, что интерес к голографии зависит от охвата посетителей, а также, что для формирования у посетителей более полного представления о технологии, требуется интерактивная цифровая среда, которая была бы для посетителя интереснее

типичных описаний у экспоната, а также давала бы релевантную для пользователя информацию, учитывая социокультурные особенности посетителя. Также, стало понятным и то, что типичный метод тестирования, применяемый при оценке эффективности выставок является устаревшим и требует его цифровизации и интеграции в интерактивную среду. Перечисленные выше события привели организаторов к необходимости разработки такой среды взаимодействия посетителя (пользователя) и организатора выставки (исследователя).

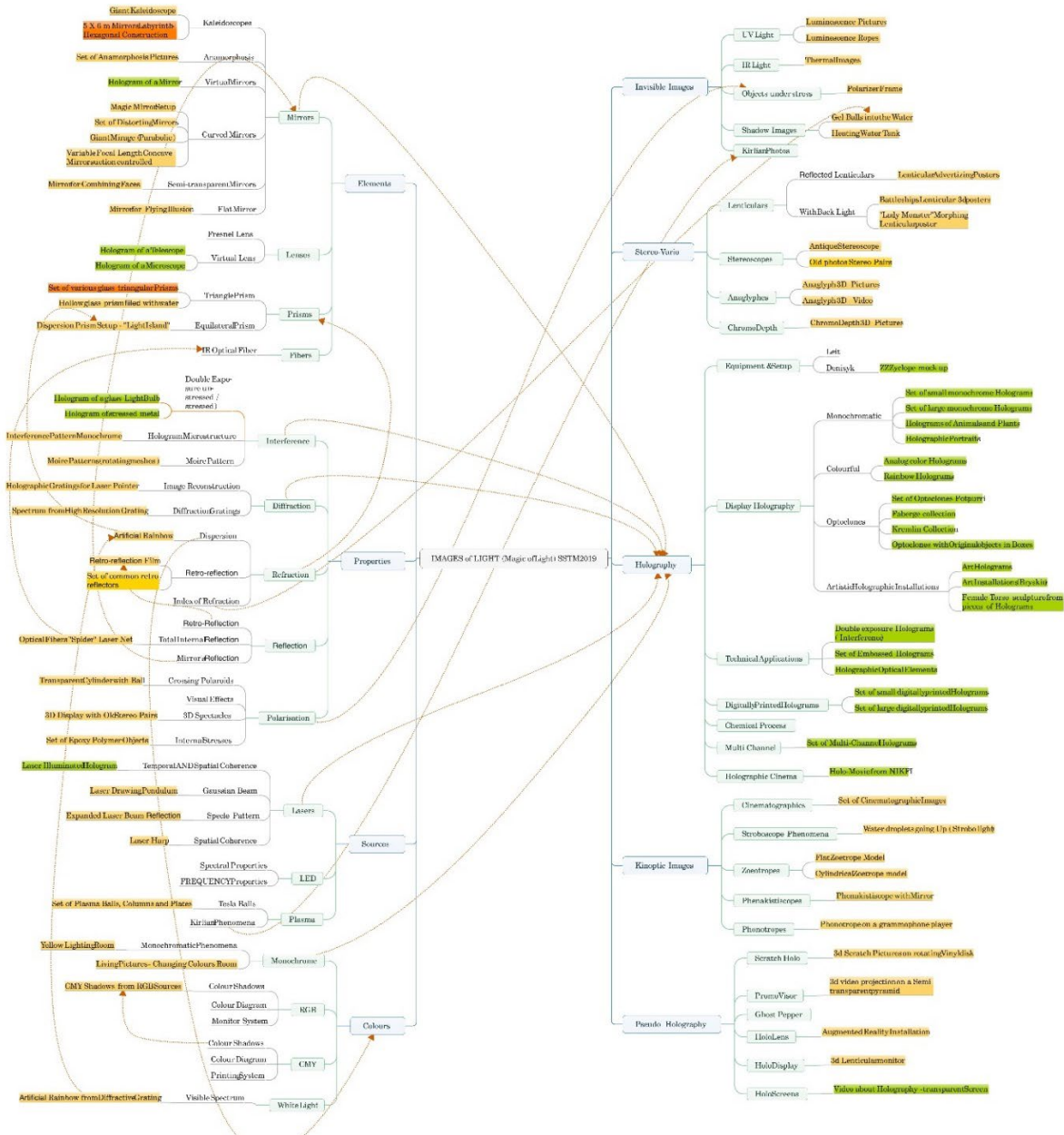


Рис. 2. Структура онтологии интерактивной голографической выставки

Такая среда для посетителя должна хранить в себе полную информацию о выставочном контенте, организованную в базу данных. Данная база данных хранит несколько описаний о каждом экспонате для выдачи посетителю конкретной целевой аудитории, а также способна анализировать активность посетителя: вести учет времени прочтения описаний и пребывания у экспоната, составлять пользовательский маршрут по экспозиции, а также оценивать вклад выставки в понимание посетителем данной области знаний.

Рассмотрим детальнее вышеупомянутое краткое описание системы. Во-первых, в основе системы лежит онтология, содержащая формализованные описания ко всем экспонатам выставки и связанные с ними понятия. На рисунке 2 представлена одна из версий модели такой онтологии.

Описания экспонатов и понятия в приведенной онтологии имеют ряд версий, отличающихся по уровню сложности – от простого и понятного даже ребенку, до продвинутого,

дающего исчерпывающую для эксперта информацию. Решение о предоставлении того или иного описания посетителю будет выдано на основании выбора пользователем своего статуса (школьник, взрослый, пенсионер). В дальнейшем система может поменять сложность описаний, если ответы пользователя на предлагаемые системой вопросы будут неверными. Во-вторых, система будет предлагать посетителю участвовать в опросах, тематически связанных с посещаемыми экспонатами. В системе продумано несколько типов опросов: входной, промежуточный, финальный и остаточный. Решение о частичном или полном прохождении опросов принимает сам пользователь, соглашаясь, или отказываясь от участия. Подобная интеграция опросов делает их более привлекательными для посетителя и позволяет привлечь к ним большее число людей, что положительно сказывается на качестве выборки. В-третьих, собранные анонимизированные данные могут быть изучены для формирования персонализированных предложений посетителю. К таким относятся советы по поиску экспоната и персонализированный маршрут по выставке. Собранные данные также полезны и организаторам, поскольку позволяют оценить множество параметров – среднее время пребывания на выставке, доступность описаний для понимания посетителем, удобство расположения экспоната, поиск целевой аудитории и так далее.

Оптимальным способом реализации данной системы кажется геймификация. Этот подход сегодня активно используется исследователями во всем мире [3] как средство привлечения аудитории, так и эффективный инструмент сбора данных в руках исследователя.

Взаимодействие с пользователем реализуется через графический веб-интерфейс системы UI, как показано на рисунке 3. UI связывается с серверными данными по технологии Rest API и представляет возможность как предоставлять информацию пользователю системы, так и проводить геймифицированные тесты. Серверные данные содержат в себе вышеописанную онтологическую базу данных, модель которой представлена на рисунке 3, а также блок тестовых заданий. Помимо этого, система имеет внешнее средство Analytics DB, связанное с UI и регистрирующее пользовательскую активность.

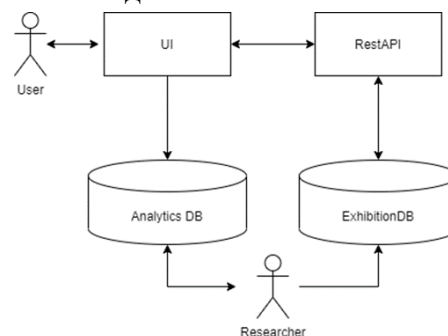


Рис. 3. Планируемая архитектура системы

На основании полученных данных исследователь проводит анализ и делает выводы об эффективности выставки.

Приведенная в данном докладе система окажется полезным инструментом и будет нести в себе двойную функцию: с одной стороны это исчерпывающая и доступная информационная база, способствующая полноценному освоению образовательной компоненты выставки, а с другой стороны это эффективное средство анализа качества проводимых выставок. Результатом работы данной системы является понятная и доступная цифровая среда, скорректированное и дополненное выставочное пространство. В случае реализации данной системы организаторы ожидают рост интереса посетителей к подобным интерактивно-голографическим выставкам и как возможное следствие, ещё больший рост интереса к образовательным программам в области фотоники, спектроскопии и голографии.

Литература

1. Stafeev S., Sarakinos A., Bobritskaya E., Lembessis A., Sharov D., Ipatova T. "Images of Light" with magic, art & science: Re-educating the public in holography, light and vision // Science Education and Museums. – 2019. – Vol.5. – №2. – Pp. 132–145.
2. Молотков И.И., Новиков Ф.А. Онтология дискретной математики в образовании. Компьютерные инструменты в образовании // Компьютерные инструменты в образовании. – 2021. – №1. – С. 69–85.
3. Hamari J., Koivisto J., Sarsa H. Does Gamification Work? – A Literature Review of Empirical Studies on Gamification // Proceedings of the 47th Hawaii International Conference on System Sciences, Hawaii, USA, January 6-9. – 2014. DOI:10.1109/HICSS.2014.377.

УДК 656.18

УСТОЙЧИВАЯ МОБИЛЬНОСТЬ: ИСПОЛЬЗОВАНИЕ ЭЛЕКТРО-САМОКАТОВ В ЕВРОПЕЙСКИХ ГОРОДАХ: ПРОБЛЕМЫ И ПЕРСПЕКТИВЫ

Шнейдер Л.¹ (аспирант)

Научный руководитель – канд. псих. наук, доцент Шмелева И.А.¹

¹*Университет ИТМО*

e-mail: schneider551204@gmx.de

Данная статья посвящена относительно новому виду микромобильности - совместному использованию электро-самокатов. С помощью углубленного обзора литературы и небольшого исследования о развитии совместного использования электронных самокатов в Брюсселе, Бельгия, в статье рассматривается вопрос, оказывают ли электронные самокаты значительное влияние на мобильность и загрязнение окружающей среды в городских районах. Выяснилось, что не все преимущества, заявленные операторами совместного использования электронных самокатов, подтверждены научными исследованиями, и что IT-технологии могут быть как полезными, так и затрудняющими использование совместных электронных самокатов.

Ключевые слова: электро-самокаты, городская мобильность, микромобильность, устойчивая мобильность, док-станции.

В городских районах по всему миру можно заметить многочисленные негативные последствия нынешних транспортных систем, ориентированных на использование автомобилей, такие как пробки, загрязнение воздуха и шум. Умные решения в области мобильности и особенно системы совместного использования, такие как прокат самокатов, могут помочь в решении этих проблем. С 2017 года совместное использование парка электрических самокатов стало привычным явлением, в настоящее время услуги совместного использования электро-самокатов доступны примерно в 177 городах мира. Технология, лежащая в основе электро-самокатов, не нова, поскольку эти самокаты являются моторизованными версиями хорошо известных кик-самокатов, которые питаются от аккумуляторной батареи.

Идея совместного использования электро-самокатов сравнима с принципами других систем проката. Люди пользуются самокатами по мере необходимости без затрат и ответственности за владение самокатами. В большинстве систем электро-самокаты являются "без доковыми", что означает отсутствие стационарных станций, и самокаты можно оставлять и забирать в любом месте в пределах обслуживаемой территории. Приложение для мобильных телефонов используется для аренды и сдачи самокатов. Это делает систему проката самокатов более дешевой в эксплуатации, чем система проката велосипедов на базе станций, и более удобной для пользователя, поскольку ему не нужно искать станцию в конце поездки.

В данной статье исследуется вопрос: оказывает ли совместное использование электронных самокатов значительное влияние на мобильность и загрязнение воздуха в городских районах, и какую роль играют здесь существующие модели совместного использования электро-самокатов и IT-инфраструктура.

История системы проката электро-самокатов

В последние годы прокат самокатов получил стремительное развитие, сначала в США, а затем и по всему миру [1]. Развитие совместного использования электро-самокатов началось

в 2017 году, когда первый прокат появился в крупных городах США [8]. В то же время было много жалоб на то, что самокаты паркуются в неудобных местах - например, у подъездов домов или посреди тротуара, их бросают в озера, крадут и ломают [8]. В конце 2018 года лидер американского рынка Lime расширил свою деятельность на города Европы. Поскольку в Бельгии и России не было никаких оговорок в отношении электро-самокатов, они появились в этих странах уже в конце 2018 года, а другие страны последовали за ними позже.

Однако, в большинстве стран, где разрешен прокат электро-самокатов, возникли проблемы, схожие с теми, что существуют в США. Помимо того, что самокаты паркуются в неудобных местах, еще одной проблемой, с которой столкнулись многие европейские города, была очень быстрая езда на самокатах по тротуарам или в пешеходных зонах [21]. В связи с этим многие города расширили существующие ограничения для электро-самокатов без док-станций и ввели ограничения скорости для них в определенных районах города.

К концу 2020 года из-за глобальной пандемии COVID-19 люди стали избегать общественного транспорта и переходить на индивидуальные виды транспорта [7]. В связи с этим рост использования электро-самокатов наблюдался как в США, так и в Европе [16]. Тем не менее, [2] наблюдалось снижение использования общественного транспорта и MaaS в связи с мерами по сдерживанию, которые требовали от многих людей оставаться дома. В связи с этим компании, предоставляющие самокаты, временно сократили количество услуг и уволили персонал как в Европе, так и в Северной Америке. Тем не менее, аналитики ожидают, что индустрия совместного использования самокатов выйдет из кризиса [3].

Традиционные преимущества и недостатки проката электро-самокатов

Если рассматривать преимущества совместного использования электро-самокатов, то, прежде всего, они удобны для пользователей благодаря концепции электро-самокатов без док-станций. Кроме того, самокаты разгоняются без особых физических усилий и ездят гораздо быстрее обычных велосипедов, не требуя при этом почти никаких физических усилий [10]. Тем не менее, в целом, электро-самокаты не составляют серьезной конкуренции велосипедам общего пользования, поскольку они используются совершенно по-разному [20]. Наконец, как и велосипедисты, водители самокатов экономят время, поскольку им не приходится сталкиваться с пробками в городских центрах [22].

Операторы часто утверждают, что прокат электро-самокатов может помочь сократить использование автомобилей и тем самым уменьшить пробки. Более того, поскольку они электрические и, следовательно, имеют нулевые выбросы углекислого газа, утверждается, что они помогают снизить загрязнение и улучшить качество воздуха [4]. Однако, последовательных доказательств этого предположения нет. Хотя оператор совместного использования самокатов Lime обнаружил, что каждая четвертая поездка на электро-самокате заменяет поездку на моторизованном транспортном средстве, независимые исследования не подтвердили эти выводы [17]. Исследования в Бельгии, Франции и США показали, что электро-самокаты в основном заменяют поездки, которые ранее совершались на общественном транспорте или проходили пешком [5, 12, 15]. Это указывает на то, что использование электро-самокатов может, в том числе, стать серьезной проблемой для здоровья человека из-за снижения физической активности. Более того, хотя при езде на электро-самокате не образуется никаких выхлопных газов, выбросы CO₂ есть, если рассматривать их полный жизненный цикл. В [11] было обнаружено, что производство и распространение электро-самокатов общего пользования приводит к огромным выбросам, что делает их менее экологичными, чем езда на велосипеде, ходьба пешком и использование общественного транспорта, поэтому сомнительно, что самокаты общего пользования приводят к снижению загрязнения в настоящее время [18].

Еще одним часто заявляемым преимуществом проката самокатов является то, что они могут стимулировать использование общественного транспорта [14]. Общественный транспорт может быть непривлекательным, если следующая остановка общественного транспорта находится далеко или обслуживается не часто. Однако, люди могут использовать самокаты, чтобы добраться до остановки общественного транспорта или до места назначения [24]. Это привело бы к пересадке на общественный транспорт, что в свою очередь привело бы

к снижению пробок и загрязнения воздуха. Однако, исследования, проведенные в Германии и Франции, показали, что электро-самокаты общего пользования используются преимущественно приезжими или туристами [25, 15]. Это может свидетельствовать о том, самокаты используются в меньшей степени для поездок по последнему километру пути («первой/последней мили») и в большей степени для поездок на отдых.

Проблемой внедрения электро-самокатов было и большое количество жалоб на то, что самокаты паркуются в неудобных местах и мешают движению пешеходов. Однако, исследования, проведенные в США, показали, что лишь меньшинство совместно используемых электронных самокатов припарковано неправильно и еще меньшее таким образом, что это мешает движению пешеходов [12, 8]. Тем не менее, во многих городах были введены специальные парковочные зоны для совместного использования электро-самокатов, чтобы гарантировать, что они не будут мешать другим участникам дорожного движения. Однако, такие парковочные зоны противоречат принципу без-доковых станций. Более того, создание парковочных зон может привести к необходимости перераспределения самокатов между различными парковочными зонами. Это может стать проблемой, поскольку транспортные средства, используемые для перераспределения в системе, очевидно, выбрасывают CO₂, что подрывает цель снижения выбросов CO₂. Наконец, хотя это и не является прямой проблемой для использования электро-самокатов, наблюдаются неравные модели участия среди пользователей электро-самокатов – в США было установлено, что люди в возрасте от 25 до 34 лет и мужчины гораздо чаще ездят на электро-самокатах, чем другие категории граждан [22].

Проблемы и недостатки проката электро-самокатов, связанные с информационными технологиями

Наряду с традиционными преимуществами и недостатками использования электро-самокатов существуют также связи с инфраструктурой информационных технологий. Во-первых, данные GPS, собранные операторами, могут помочь решить некоторые из вышеупомянутых традиционных проблем. Эффективно используя эти данные, можно получить представление о местоположении, количестве и продолжительности поездок на электро-самокатах в городе. Это может помочь в управлении распределением парка совместно используемых электро-самокатов, выявлении пробелов в обслуживании и управлении парковкой и загрузкой электро-самокатов [6]. Однако, при использовании этих данных, компаниям необходимо учитывать соображения конфиденциальности.

Кроме того, виртуальные геозоны могут помочь в решении проблем, связанных с неправильной парковкой. Геозонирование — это программная функция, которая с помощью технологии GPS или RFID устанавливает географические границы, которые могут быть использованы для отображения в мобильном приложении зон парковки или зон, где запрещено парковаться или где существуют специальные правила использования электро-самокатов [23].

В следующем разделе будет приведено краткое исследование развития проката электро-самокатов в Брюсселе, Бельгия.

Пример Брюсселя, Бельгия

Брюссельский столичный регион (БСР) состоит из 19 муниципалитетов, включая город Брюссель. Площадь БСР составляет 161 км², население – 1,2 миллиона человек. Это делает его самым густонаселенным регионом Бельгии. Брюссель имеет холмистый рельеф, что создает проблемы для существующей системы совместного использования велосипедов, поскольку подъем на велосипеде по холмам требует больших физических усилий. Поэтому электро-самокаты могут иметь довольно большой потенциал в этом городе.

Поэтому неудивительно, что одним из первых городов, куда Lime распространился в 2018 году в Европе, стал Брюссель. Как и в США, система оказалась весьма успешной. В период с ноября 2018 года по январь 2019 года система зарегистрировала 150 000 поездок, которые составили 220 000 километров. Такая популярность привела к тому, что дополнительные операторы проката вывели свои самокаты в Брюссель, а продажи электро-самокатов выросли [19]. Однако, как и в США, стали поступать жалобы на электро-самокаты, припаркованные в неудобных местах, где они мешали другим пользователям дорог и

пешеходных дорожек. В связи с этим с 1 февраля 2019 года все компании, предлагающие бездоковые электро-самокаты, должны были получить лицензию от министерства транспорта БСР. Согласно этой лицензии, компании обязаны эксплуатировать свои транспортные средства используя экологически чистую энергию, следить за тем, чтобы пользователи соблюдали Правила дорожного движения, и передавать информацию о пользователях в БСР. Кроме того, было ограничено количество транспортных средств, которые могут быть размещены в определенных местах в регионе. Для обеспечения соблюдения этих правил операторы должны платить штраф за каждый неправильно припаркованный электро-самокат. В дополнение к обязательной лицензии были введены добровольные зоны парковки для электро-самокатов и без доковых велосипедов. Кроме того, инструкции по использованию систем должны быть доступны на английском, французском и голландском языках. В связи с появлением в Брюсселе дополнительных компаний по продаже самокатов к июню 2019 года на улицах БСР было разбросано около 4 000 самокатов с ценой 1 евро за поездку + 0,15 евро за минуту [9]. Это огромное количество самокатов усугубило существующую проблему с парковкой. В связи с этим городские власти начали рассматривать возможность введения зон, где парковка электро-самокатов будет запрещена, или будут созданы специальные парковочные места для электро-самокатов.

В октябре 2019 года в БСР были введены первые специальные парковочные места для электро-самокатов. В дополнение к этому министр транспорта БСР объявил о создании так называемых парковочных «дропзон» рядом с пешеходными переходами для электро-самокатов, которые заменили парковочные места для автомобилей рядом с переходами. Уже в течение 2019 года произошла консолидация брюссельского рынка проката электро-самокатов. В августе 2019 года из-за ухода двух операторов из города было вывезено 1 300 самокатов [19].

В связи с глобальной пандемией COVID - 19 и соответствующими мерами по ее локализации компании, предоставляющие самокаты в совместное пользование, сократили свои услуги к середине марта 2020 года, но в мае 2020 года обе компании возобновили свою работу [13]. О восстановлении индустрии самокатов в Брюсселе говорит и тот факт, что в июне 2021 года новая компания (Bolt) запустила в город 1 100 электронных самокатов [26]. В Брюсселе электро-самокаты в основном используются для коротких поездок, которые ранее совершались на общественном транспорте или пешком, и лишь иногда заменяют поездки на автомобиле или мотоцикле. Типичный водитель электро-самоката в Брюсселе – молодой мужчина с высшим образованием от 25 до 34 лет. Средняя поездка длится от пяти до десяти минут [5].

Появление электрических самокатов по всему миру - одно из последних достижений в постоянно развивающейся индустрии интеллектуальной мобильности и услуг совместного пользования. Хотя технология, лежащая в основе электро-самокатов, не нова, внедрение информационных технологий и аккумуляторов привело к созданию нового вида городского транспорта, который может оказать значительное влияние на мобильность и выбросы загрязняющих веществ и выбросы CO₂ в городских районах. Однако, на данный момент доказательства такого влияния не являются убедительными, поскольку в основном заменяются поездки на общественном транспорте, велосипеде и пешком. Более того, поскольку при производстве и распространении электронных самокатов образуются огромные выбросы CO₂, они менее экологичны, чем велосипеды, пешие прогулки или общественный транспорт. Кроме того, из-за жалоб на то, что самокаты паркуются в неудобных местах, многие города ввели специальные парковочные места для совместного использования электронных самокатов, что может помешать концепции без доковых совместных электро-самокатов. В связи с этими неоднозначными результатами сомнительно, что в настоящее время прокат электро-самокатов действительно способен внести реальный вклад в городскую мобильность.

Необходимо проследить, какое влияние на развитие данной инфраструктуры в городах оказывают различные действия, предпринимаемые операторами проката электро-самокатов, а также муниципалитетами для решения вышеупомянутых проблем. Для определения этих эффектов и того, какое влияние IT-технологии могут оказать на прокат электро-самокатов в Брюсселе, необходимы дополнительные исследования.

Литература

1. Baek K., Lee H., Chung J.H., Kim J. Electric scooter sharing: How do people value it as a last-mile transportation mode? // *Transp Res D Transp Environ.* – 2021. – V.90. – Pp. 102642.
2. Beck M.J., Hensher D.A. Insights into the impact of COVID-19 on household travel and activities in Australia – The early days under restrictions // *Transp Policy.* – 2020. – V.96. – Pp. 76–93.
3. Berg Insight. *The Bike and Scootersharing Telematics Market (2nd Edition)*. Berg Insight. – 2020. – 4 p.
4. Boglietti S., Barabino, B. & Maternini, G. Survey on e-Powered Micro Personal Mobility Vehicles: Exploring Current Issues towards Future Developments // *Sustainability.* – 2021. – V.13. – Pp. 3692.
5. Bruxelles Mobilité. *Enquête sur l'usage des Trotinettes Électriques à Bruxelles*. Bruxelles Mobilité. – 2019. – 60 p.
6. Cardell M., Holm Møller, T. How micro mobility is moving cities into a sustainable future. [Электронный ресурс]. – Режим доступа: https://www.ey.com/en_gl/automotive-transportation/how-micromobility-is-moving-cities-into-a-sustainable-future (дата обращения: 10.03.2022).
7. De Vos J. The effect of COVID-19 and subsequent social distancing on travel behavior // *Transp. Res. Interdiscip. Perspect.* – 2020. – V.5. – Pp. 100121.
8. Fang K., Weinstein Agrawal A., Steele J., John J. H., Hooper A.M. Where Do Riders Park Dockless, Shared Electric Scooters? Findings from San Jose, California // *Mineta Transportation Institute Publications.* – 2018. – 5 p.
9. Galindo G. Unruly e-scooter parking to be regulated in Brussels. [Электронный ресурс]. – Режим доступа: <https://www.brusselstimes.com/brussels/58266/unruly-e-scooter-parking-to-be-regulated-in-brussels/> (дата обращения: 10.03.2022).
10. Gössling S. Integrating E-scooters in Urban transportation: Problems, Policies, and the Prospect of System Change // *Transp Res D Transp Environ.* – 2020. – V.79. – Pp. 102230.
11. Hollingsworth J., Copeland B., Johnson J.X. Are e-scooters polluters? The environmental impacts of shared dockless electric scooters // *Environ. Res. Lett.* – 2019. – V.14. – №8. DOI:10.1088/1748-9326/ab2da8.
12. James O., Swiderski J. I., Hicks J., Teoman D., Buehler R. Pedestrians and e-scooters: An initial look at e-scooter parking and perceptions by riders and non-riders // *Sustainability.* – 2019. – V. 11. – № 20. DOI:10.3390/su11205591.
13. Johnston J. Lime scooters gradually return to Brussels from today. [Электронный ресурс]. – Режим доступа: <https://www.brusselstimes.com/brussels/111758/lime-scooters-gradually-return-to-brussels-from-today/> (дата обращения: 10.03.2022).
14. Kong H., Jin S.T., Sui D.Z. Deciphering the relationship between bike sharing and public transit: Modal substitution, integration, and complementation // *Transport. Res. Part D: Trans. Environ.* – 2020. – V.85. – Pp. 102392.
15. Krier C., Chrétien J., Louvet N. Uses and users of free-floating e-scooters in France. 6t-bureau de recherche, 2019. [Электронный ресурс]. – Режим доступа: <https://6-t.co/en/free-floating-escooters-france/> (дата обращения: 10.03.2022).
16. Lime. *End of Year Report 2020 Official One-Year Report*. [Электронный ресурс]. – Режим доступа: <https://www.li.me/pdf/eoy-report.pdf> (дата обращения: 10.03.2022).
17. Lime. *Research Finds Most E-Scooter Riders Are Local Commuters, Not Tourists*. [Электронный ресурс]. – Режим доступа: <https://www.li.me/second-street/research-finds-e-scooter-riders-local-commuters-not-tourists> (дата обращения: 10.03.2022).
18. Luo H., Kou Z., Zhao F., Cai H. Comparative life cycle assessment of station-based and dock-less bike sharing systems // *Resour. Conserv. Recycl.* – 2019. – V.146. – Pp. 180–189.
19. McCullough E. Sales of e-scooters in Brussels are on the rise. [Электронный ресурс]. – Режим доступа: <https://www.brusselstimes.com/brussels/65148/sales-of-e-scooters-in-brussels-are-on-the-rise-mobility/> (дата обращения: 10.03.2022).

20. McKenzie G. Spatiotemporal comparative analysis of scooter-share and bike-share usage patterns in Washington, D.C. // *J. Transp. Geogr.* – 2019. – V.78. – Pp. 19–28.
21. Pfeiffer-Goldmann D. Frankfurt und die Elektroroller: Mit festen Parkzonen gegen das Chaos. [Электронный ресурс]. – Режим доступа: <https://www.fnp.de/frankfurt/frankfurt-feste-parkzonen-sollen-das-chaos-beseitigen-90658538.html> (дата обращения: 10.03.2022).
22. Sanders R.L., Branion-Calles M., Nelson T.A. To scoot or not to scoot: Findings from a recent survey about the benefits and barriers of using E-scooters for riders and non-riders // *Transp Res Part A Policy Pract.* – 2020. – V.139. – Pp. 217–227.
23. Shadel Fischer P. Understanding and Tackling Micromobility: Transportation’s New Disruptor. [Электронный ресурс]. – Режим доступа: https://www.ghsa.org/sites/default/files/2020-08/GHSA_MicromobilityReport_Aug31Update.pdf (дата обращения: 10.03.2022).
24. Shaheen S., Chan N. Mobility and the sharing economy: Potential to facilitate the first-and last-mile public transit connections // *Built Environ.* – 2016. – V.42. – Pp. 573–588.
25. Tack A., Klein A., Bock B. E-Scooters in Germany. A data-driven contribution to the ongoing debate. [Электронный ресурс]. – Режим доступа: <https://scooters.civity.de/en> (дата обращения: 10.03.2022).
26. The Brussels Times. 1,100 e-scooters: Belgium’s newest rental company lands in Brussels. [Электронный ресурс]. – Режим доступа: <https://www.brusselstimes.com/brussels/174209/another-e-scooter-rental-service-enters-the-brussels-market-bolt-green-recycle/> (дата обращения: 10.03.2022).

УДК 004.891.2

МЕТАСHERCHANT: ИСПОЛЬЗОВАНИЕ Hi-C СВЯЗЕЙ ПРИ ВЫДЕЛЕНИИ ГЕНОМНОГО КОНТЕКСТА ИЗ МЕТАГЕНОМОВ

Шостина А.Д.¹ (студент), Иванов А.Б.¹ (аспирант)

Научный руководитель – канд. техн. наук Ульянов В.И.¹

¹*Университет ИТМО*

e-mail: shostina77@gmail.com

Работа выполнена в рамках темы НИР №621290 «Алгоритмы извлечения признаков для сравнительного анализа групп метагеномных образцов полногеномного секвенирования с использованием графов де Брейна и применением мультимодальных омиксных данных».

В работе предложен метод построения геномного контекста с учетом Hi-C связей, который был реализован на основе приложения MetaCherchant. Расширенный геномный контекст позволит анализировать Hi-C связи между геномами бактерий, плазмид и вирусов.

Ключевые слова: геномный контекст, Hi-C секвенирование, полногеномное секвенирование, граф де Брейна, антибиотико-резистентный ген.

Метагеномика занимается изучением генома не отдельного организма, а совокупности обитателей микробных сообществ, живущих в разных природных условиях. Выяснение глобального разнообразия микробиоты важно для понимания роли микробиоты в здоровье хозяина и для обнаружения способов модулирования микробного сообщества для профилактики и лечения заболеваний. Например, анализ микробиоты кишечника позволяет выделить бактерии, которые имеют свойство устойчивости к антибиотикам, так как в микробиоте кишечника было обнаружено большое количество антибиотико-резистентных генов.

В данной работе рассматривается два способа получения метагеномных данных: метод полногеномного секвенирования (WGS) и метод определения конформации хромосом (Hi-C). Результатом работы обоих методов является набор пар ридов (прочтений). Ридом называется небольшой фрагмент нуклеотидной последовательности генома. В ходе полногеномного секвенирования риды считываются со случайных позиций. На основе полученных пар ридов происходит сборка генома, то есть восстановление его последовательности целиком.

Для сборки генома используются графы де Брейна. Граф де Брейна – это ориентированный граф с петлями и кратными ребрами, вершинами которого являются различные k-меры, а ребрами k+1-меры. K-мер – это нуклеотидная последовательность длины k. Ребро соединяет два k-мера, если один k-мер является его префиксом, а другой k-мер его суффиксом. Сжатый граф де Брейна – это граф де Брейна, в котором все пути без ответвлений были сжаты в одну вершину. Таким образом, был получен граф, вершинами которого являются контиги. Контиг – это однозначно определенная нуклеотидная последовательность максимальной длины.

Hi-C секвенирование используется для определения пространственной организации генома. Пары Hi-C ридов являются фрагменты нуклеотидной последовательности, которые в пространстве расположены близко друг к другу. Это могут быть как риды из одного генома, так и риды из двух различных геномов. Однако, пары Hi-C ридов всегда лежат внутри одной клетки.

В области здравоохранения существуют проблемы, связанные с анализом геномного контекста, например, проблема устойчивости бактерий к некоторым видам антибиотиков. Антибиотико-резистентные гены (АРГ) часто располагаются в плаزمиде – небольших молекулах ДНК, физически обособленных от основного генома. В природе плазмиды обычно содержат гены, повышающие приспособленность бактерий к окружающей среде. Например, бактерия приобретает свойство устойчивости к антибиотику, если плазида с АРГ находится в клетке бактерии. Такие бактерии можно определить только при помощи метода секвенирования Hi-C, так как между геномом бактерии и геномом плазмиды есть Hi-C связи.

Геномный контекст – это подграф графа де Брейна, в котором содержится исследуемый ген. Таким образом, анализ геномного контекста, построенного с учетом Hi-C связей, поможет решить проблему поиска бактерий, которые приобрели свойство устойчивости к антибиотикам за счет наличия плазмиды с АРГ.

Для возможности проведения анализа расширенного геномного контекста необходимо разработать метод построения и визуализации геномного контекста с учетом Hi-C связей.

Построение геномного контекста с учетом Hi-C связей

Построение геномного контекста будем проводить в четыре этапа. На первом этапе осуществляется построение геномного контекста вокруг анализируемого гена. Данный этап выполняется при помощи приложения MetaCherchant [1].

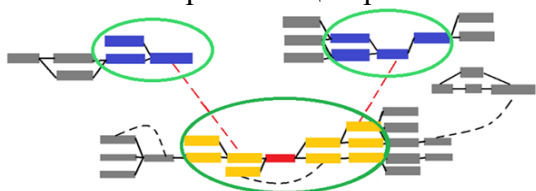


Рис. 1. Геномный контекст, построенный с учетом Hi-C связей

На рисунке 1 приведен пример исходного геномного контекста. Исследуемый ген был изображен в виде красного прямоугольника. Контиги, вошедшие в геномный контекст, покрашены в желтый цвет, а не вошедшие в исходный контекст контиги – в серый и синий цвет.

На втором этапе происходит поиск Hi-C ридов, которые лежат вне построенного геномного контекста, однако, имеют Hi-C связи с ним. На рисунке 1 Hi-C связи, которые будут использоваться при построении расширенного геномного контекста, изображены красной пунктирной линией, а Hi-C связи, не используемые при построении геномного контекста, черной пунктирной линией.

Для поиска интересующих нас пар Hi-C ридов сделаем картирование пар Hi-C ридов на контиги исходного геномного контекста при помощи утилиты bwa [2]. В результате картирования каждому Hi-C риду будет сопоставлен контиг, в котором он находится. Если такого контига нет, то рид будет отмечен как неизвестный. Затем найдем такие неизвестные Hi-C риды, парный рид которых был сопоставлен с контигом из исходного контекста. Тем самым, мы найдем все пары Hi-C ридов, расширяющие исходный геномный контекст.

На третьем этапе осуществляется построение объединенного геномного контекста вокруг анализируемого гена и всех Hi-C ридов, выделенных на этапе два. Приложение MetaCherchant позволяет выделять контекст вокруг нескольких нуклеотидных последовательностей, следовательно, его можно использовать для построения контекста на этом этапе. На рисунке 1 синим цветом отмечены те контиги, которые входят в расширенный геномный контекст, но не присутствуют в исходном геномном контексте.

На четвертом этапе геномный контекст визуализируется. В качестве средства визуализации графа де Брейна в данной работе используется приложение Bandage [3]. Однако, приложение Bandage при текущей реализации не позволяет отображать Hi-C связи, поэтому необходимо разработать способ их визуализации. Предлагается изображать Hi-C связи пунктирной линией, соединяющей середины контигов.

Пример геномного контекста для бактерии Сальмонелла

Тестирование разработанных методов построения и визуализации геномного контекста производилось на синтетическом примере, то есть используемые пары WGS и Hi-C ридов были сгенерированы при помощи утилит InSilicoSeq [4] и sim3C [5].

Для генерации использовался фрагмент генома бактерии Сальмонелла и геном ее

плазмиды. Длина нуклеотидной последовательности генома фрагмента бактерии и плазмиды составила примерно 100000 нуклеотид. При секвенировании было сгенерировано 30000 пар WGS и Hi-C ридов. При выделении геномного контекста использовалось значение покрытия 5. Это значит, что каждый k-мер, используемый при построении графа де Брейна, содержится как минимум в пяти различных ридах. Геномный контекст был ограничен максимальным радиусом, то есть максимальной дистанцией между любым k-мером и анализируемым геном. Использовался максимальный радиус равный 10000. Вследствие этого размер исходного контекста также составил 10000 нуклеотид, а размер расширенного геномного контекста составил 200000, то есть используемые геномы были восстановлены полностью.

На рисунке 2 изображен расширенный геномный контекст, построенный для фрагмента бактерии *Сальмонелла* и ее плазмиды. Исследуемый ген был выбран из генома плазмиды, поэтому весь исходный контекст расположен в плазмиде. Исследуемый ген выделен красным цветом, а исходный контекст – оранжевым. Следует отметить, что фрагмент генома бактерии *Сальмонелла* попал в расширенный контекст из-за того, что между геномом плазмиды и геномом бактерии есть Hi-C связи.

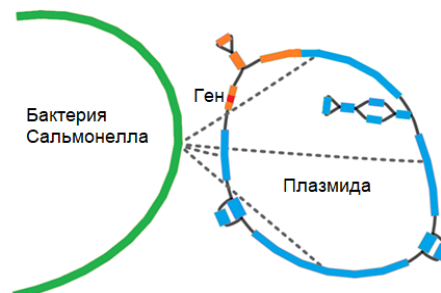


Рис. 2. Геномный контекст для бактерии *сальмонелла* и ее плазмиды

Таким образом, в данной работе был разработан метод построения геномного контекста с учетом Hi-C связей. Данный метод был реализован на основе приложения MetaCherchant. В дальнейшем планируется протестировать разработанный метод на большом количестве реальных и сгенерированных данных. Также планируется улучшить предложенный способ визуализации графа де Брейна с Hi-C связями в приложение Bandage.

Литература

1. Olekhnovich E.I., Vasilyev A.T., Ulyantsev V.I., Kostryukova E.S., Tyakht A.V. MetaCherchant: analyzing genomic context of antibiotic resistance genes in gut microbiota // *Bioinformatics*. – 2018 – Т.34. – Вып.3. – С. 434–444.
2. Li Heng. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM // *ArXiv* :1303.3997. – 2013. – Т.1303.
3. Ryan R.W., Mark B.S., Justin Z., Kathryn E.H. Bandage: interactive visualization of de novo genome assemblies // *Notes Bioinformatics*. – 2015. – Т.31. – С. 3350–3352.
4. Gourelé H., Karlsson-Lindsjö O., Hayer J. and Bongcam-Rudloff E. Simulating Illumina data with InSilicoSeq. // *Bioinformatics*. – 2019. – Т.35. – Вып.3. – С. 521–522.
5. DeMaere M.Z., Darling A.E. Sim3C: simulation of Hi-C and Meta3C proximity ligation sequencing technologies // *GigaScience*. – 2018. – Т.7. – Вып.2. – С. 1–12.

УДК 004.925; 004.942

СТАНОВЛЕНИЕ ЕСТЕСТВЕННОЙ КОНВЕРГЕНЦИИ ЧЕЛОВЕЧЕСКОГО ЗРЕНИЯ В ВИДЕОПРОЗРАЧНЫХ СИСТЕМАХ СМЕШАННОЙ РЕАЛЬНОСТИ

Эседов Н.Э.¹ (студент), Жданов А.Д.¹

Научный руководитель – канд. физ.–мат. наук Жданов Д.Д.¹

¹*Университет ИТМО*

e-mail: narimanesedov@bk.ru, ddzhdanov@mail.ru, andrew.gtx@gmail.com

В статье описан подход, позволяющий восстанавливать изображение, сформированное видео-прозрачной системой смешанной реальности, в соответствии с естественной конвергенцией глаз пользователя устройства, полученной из системы отслеживания зрачков устройства смешанной реальности. Способ восстановления изображения основан на использовании расширенного (2,5-мерного) представления изображения, полученного, например, с использованием 3D-сканера, который строит карту глубин сцены, и его дальнейшего восстановления путем проецирования точек исходного изображения на точки изображения, соответствующие требуемой конвергенции, методом «перефокусировки» для каждой точки, при условии, что оптические системы, которые формируют изображения реального мира на ЖК-экранах, и окуляры, которые проецируют эти изображения в глаза пользователя, не изменяют свои характеристики и положение. Продемонстрированы преимущества и недостатки данного метода.

Ключевые слова: визуальное восприятие, смешанная реальность, виртуальная реальность, отслеживание зрачка, фотореалистичный рендеринг, конфликт вергенции.

Быстрая интеграция систем виртуальной, дополненной и смешанной реальности в человеческую жизнь требует длительного времени пребывания человека в виртуальной среде [1]. Такие устройства могут использоваться практически во всех сферах человеческой деятельности, от использования в промышленности до личного домашнего использования, от проектирования самолетов до простой помощи по дому, как например, для демонстрации рецептов при приготовлении пищи. Это приводит к более строгим требованиям к разработчикам оптических систем и разработчикам программного обеспечения для исследования и разработки новых способов как сделать эти системы максимально возможно удобными для конечного пользователя [2]. Это включает в себя не только требования к улучшению качества ЖК-матриц и оптики, но и к более комфортному восприятию формируемого изображения смешанной реальности для глаз и мозга пользователя [3].

Имеется два основных типа систем смешанной реальности: оптически-прозрачная и видео-прозрачная. Главное различие между этими системами заключается в том, как пользователь наблюдает за реальным миром. В ситуации оптически-прозрачной системы пользователь устройства свободно наблюдает за реальным миром [4, 5], а в случае видео-прозрачной системы смешанной он наблюдает реальный мир, спроецированный с ЖК-экранов на сетчатку глаза [6, 7]. Как итог, в первом случае, когда пользователь наблюдает реальный мир с помощью оптически-прозрачной системы, конвергенция зрения идентично направлению взгляда, однако, в случае видео-прозрачной системы иначе, так как направление взгляда возможно не соответствует конкретизированному направлению параллельных камер устройства.

В рамках данной статьи авторы сосредоточились на конфликте вергенции зрения, возникающей при использовании видео-прозрачных систем смешанной реальности. При наблюдении за реальным миром с использованием видео-прозрачной системы смешанной реальности требуется дополнительная обработка изображений [8, 9], чтобы снизить расхождение вергенции, возникающее между камерами устройства и глазами пользователя в соответствии с данными, полученными от системы отслеживания направления взгляда глаз пользователя, как показано на рисунке 1.

На рисунке 2 показано, как пользователь видео-прозрачной системы наблюдает реальность, с помощью такого типа устройств.

Конфликт вергенции вызван фиксированным направлением камер устройства, которые получают изображение реального мира, а конфликт аккомодации вызван постоянством фокусного расстояния оптической системы, которая формирует изображение реального мира на сетчатке глаза человека, проецируемого с ЖК матрицы.

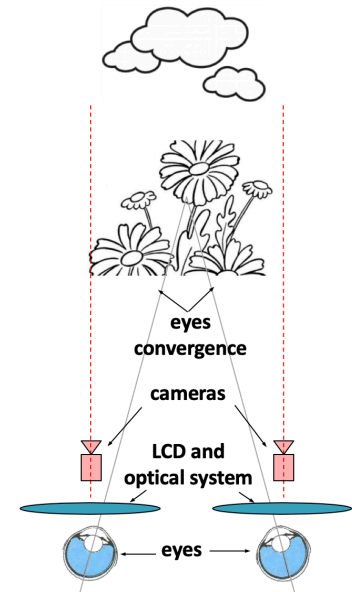


Рис. 1. Конфликт вергенции в видео-прозрачной системе смешанной реальности

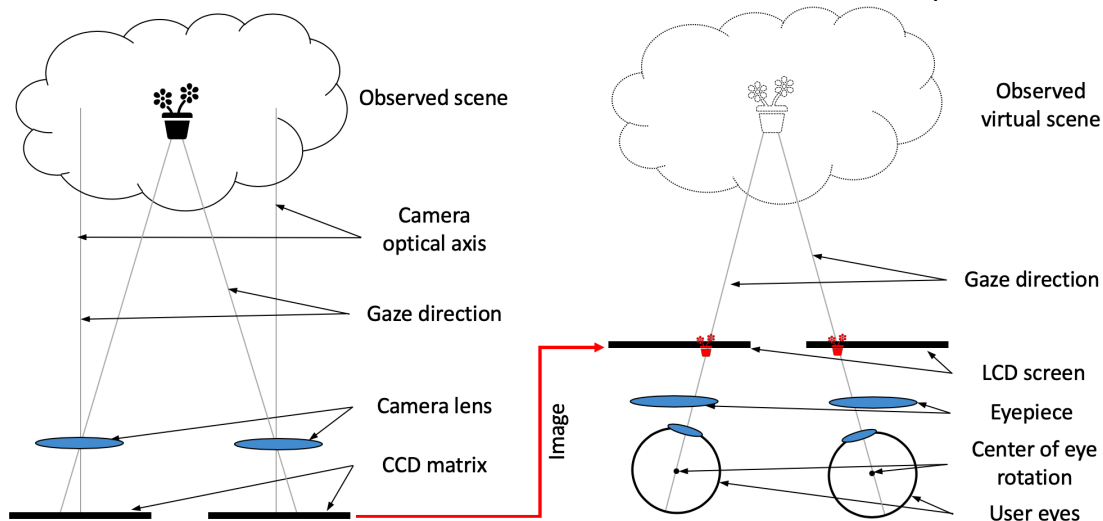


Рис. 2. Наблюдение реального мира при использовании видео-прозрачной системы смешанной реальности

С другой стороны, при наблюдении за объектами виртуального мира оба типа систем смешанной реальности в равной мере подвержены конфликтам вергенции [10], и метод исправления таких конфликтов одинаков для обоих типов систем. Для преодоления конфликта зрительного восприятия, вызванного несоответствием ориентации фронтальных камер устройства смешанной реальности и параметров глаз пользователя устройства, требуется специальная обработка изображений. Эта обработка предназначена для обновления изображения, отображаемого на ЖК-экранах устройства смешанной реальности, чтобы объект, просматриваемый пользователем устройства, находился в фокусе и смещался в соответствии с положением глаз пользователя. Обязательным условием устранения дискомфорта вергенции является отслеживание направления взгляда. Для этого используется система слежения за зрачком, которая может определять положение зрачка и вычислять соответствующее направление взгляда в режиме реального времени с достаточной точностью. Зная направление взгляда, система синтеза изображения виртуального мира может формировать изображение для левого и правого глаза пользователя устройства с учетом правильного схождения направления взгляда глаз пользователя.

В видео-прозрачной системе смешанной реальности этот метод коррекции изображения виртуального мира можно использовать для коррекции изображений реального мира, снятых фиксированными параллельно-направленными передними камерами устройства. Следует отметить, что данная система смешанной реальности формирует двумерное изображение реального мира, в котором теряется расстояние до точки, где расположен наблюдаемый объект. Однако, при вычислении нового положения точки изображения при вращении оптической системы вокруг зрачка камеры расстояние до объекта принципиальной роли не играет. На рисунке 3 показано, как меняется видимое изображение при вращении камеры вокруг входного зрачка оптической системы камеры.

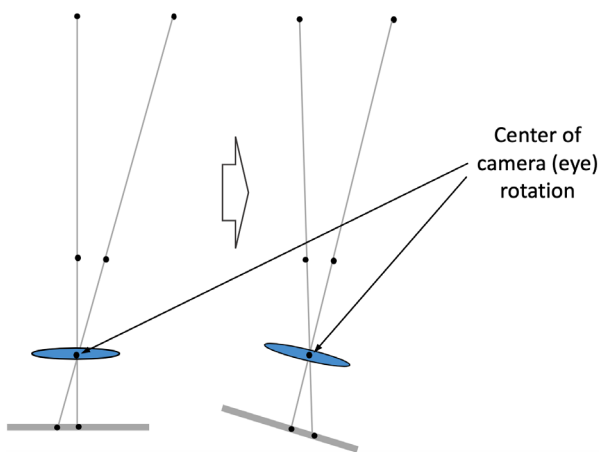


Рис. 3. Поворот камеры вокруг зрачка оптической системы камеры

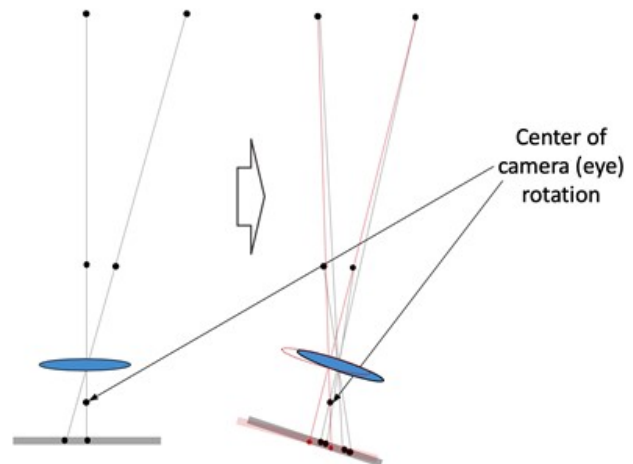


Рис. 4. Поворот камеры вокруг средней точки между входным зрачком оптической системы камеры и матрицей

Координаты точек изменились, но точки, которые изначально находились на том же луче, проходящем через входной зрачок камеры, после поворота камеры останутся на том же луче, и в результате ей нужно только пересчитать положение пикселей при новом угле камер, чтобы исправить изображение в соответствии с новыми настройками камеры. Таким образом, это преобразование может быть выполнено на основе 2D-изображения реального мира для левого и правого глаза, исходя из правильного направления взгляда, полученного от системы отслеживания зрачка. Если предположить, что виртуальные камеры системы смешанной реальности имитируют реальные глаза пользователя устройства, то их следует повернуть относительно некоторой средней точки между входным зрачком оптической системы камеры и матрицей, как показано на рисунке 4.

Для воссоздания изображения с параллельных камер с необходимой конвергенцией, полученной от системы слежения за глазами, используются соответствующие исходные изображения высокого разрешения и карты глубины точек сцены (облако точек, полученное из систем сканирования сцены). Простые геометрические преобразования восстановят изображения для левого и правого глаза пользователя с исправленной конвергенцией. Поскольку изображение сцены, создаваемое камерами смешанной реальности, имеет дискретную структуру, в процессе рендеринга с компенсацией конвергенции могут возникать перекрывающиеся или отсутствующие точки изображения. В случае наложения возникает излишняя яркость, а при отсутствии точек на финальном изображении появятся черные пиксели.

В итоге, предложенный метод реконструкции конвергенции состоит из следующих шагов:

- сформировать карту диспаратета двух изображений, снятых параллельными фронтальными камерами, и восстановить облако точек сцены или 2,5-мерное представление реального мира;
- сформировать две виртуальные камеры на основе информации от системы отслеживания зрачка для левого и правого глаз в системе координат восстановленного на предыдущем шаге облака точек реального мира;

- используя полученное 2,5-мерное представление реального мира и две виртуальные камеры выполнить рендеринг изображений, которые должны быть видны левым и правым глазами пользователя при соответствующей конвергенции;
- интерполировать «черные» точки, которые могут появиться в процессе рендеринга изображения;
- применить дефокусировку полученного изображения, чтобы размыть полученные изображения в соответствии с ожидаемым фокусным расстоянием глаз пользователя, полученным из условия естественного соответствия вергенции и аккомодации зрения.

Исследование показало, что конвергенцию можно правильно реконструировать по изображениям, снятым статическими камерами с параллельными осями, формирующими изображения для левого и правого глаза. В результате восстановленные изображения практически не содержат артефактов, а оставшиеся артефакты изображения сосредоточены в основном на границах объектов. Кроме того, обнаружена ошибка реконструкции карты глубин сцены, которая очень мало влияет на качество конечного изображения, а локализация ошибок реконструкции ограничена высокочастотными областями исходных изображений. Более серьезной проблемой является наличие в сцене зеркальных и прозрачных объектов, которые вносят неоднозначность в определение карты глубины и могут внести заметные искажения в сформированное изображение.

Еще одна проблема зрительного дискомфорта, выходящая за рамки данного исследования, – это коррекция изображения в соответствии с аккомодацией глаз пользователя. Когда глаз сфокусирован на объекте, все другие объекты сцены, не находящиеся в плоскости фокуса, расфокусируются. Для этого на этапе синтеза изображения добавляется расфокусировка, соответствующая реальному восприятию глазом дистанции фокусировки. Функция рассеяния точки вне плоскости представляет собой цилиндр, радиус цилиндра которого зависит от фокусного расстояния глаза пользователя, диаметра зрачка и смещения точки обзора от плоскости объекта. Учитывая, что диаметр зрачка и фокусное расстояние глаза пользователя постоянны для всего изображения, радиус функции масштабируется только в соответствии с расстоянием до точки. В результате, если у нас есть «идеальное» изображение камеры-обскуры, и мы знаем расстояние от зрачка до точек сцены, соответствующих точкам изображения, мы можем просто свернуть изображение, используя цилиндрическую функцию. Единственной технической сложностью операции свертки является переменный радиус функции фильтра, который необходимо вычислять для каждой точки изображения. Другая проблема лежит в области оптики и связана с адаптацией фокусировки окуляра к текущей конвергенции глаз.

Литература

1. Potemin I.S., Zhdanov D.D., Zhdanov A.D., Bogdanov N.N., Voloboy A.G. Hybrid ray tracing method for photorealistic image synthesis in head-up displays // Proc SPIE 10690, 5 June 2018, Optical Design and Engineering VII. – 2018. – Pp. 106900I.
2. Potemin I.S., Livshits I., Zhdanov D., Zhdanov A., Bogdanov N. An application of the virtual prototyping approach to design of VR, AR, and MR devices free from the vergence-accommodation conflict // Proc SPIE 10694, 28 May 2018, Computational Optics II. – 2018. – Pp. 1069404.
3. Potemin I.S., Zhdanov A., Bogdanov N., Zhdanov D., Livshits I., Wang Y. Analysis of the visual perception conflicts in designing mixed reality systems // Proc SPIE 10815, 5 November 2018, Optical Design and Testing VIII. – 2018. – Pp. 108150U.
4. Microsoft Corporation Microsoft HoloLens. [Электронный ресурс]. – Режим доступа: <https://www.microsoft.com/en-us/hololens> (дата обращения: 05.03.2022).
5. Magic Leap Inc Home. [Электронный ресурс]. – Режим доступа: <https://www.magicleap.com> (дата обращения: 05.03.2022).
6. Hewlett-Packard Company HP Windows Mixed Reality Headset. [Электронный ресурс]. – Режим доступа: <https://www8.hp.com/us/en/campaigns/mixedrealityheadset/overview.html> (дата обращения: 05.03.2022).

7. Microsoft Corporation Windows Mixed Reality. [Электронный ресурс]. – Режим доступа: <https://www.microsoft.com/en-us/windows/windows-mixed-reality> (дата обращения: 05.03.2022).
8. Benko H., Ofek E., Zheng F., Wilson A.D. FoveAR: Combining an Optically See-Through Near-Eye display with Spatial Augmented Reality Projection // Proc UIST'15, November 11 - 15 – Charlotte, NC, USA. – 2015. – Pp. 129–135.
9. Xiao R., Benko H. Augmenting the Field-of-View of Head-Mounted Displays with Sparse Peripheral Displays // Proc CHI'16, May 07 - 12 – San Jose, CA, USA. – 2016. – Pp. 1221–1232.
10. Kramida G. Resolving the Vergence-Accommodation Conflict in Head-Mounted Displays // IEEE Transactions on Visualization and Computer Graphics, July 1. – 2016. – Vol.22. – Issue 7. – Pp. 1912–1931.

УДК 004.048

РАСПОЗНАВАНИЕ ИМЕНОВАННЫХ СУЩНОСТЕЙ С ПОМОЩЬЮ DOCCANO

Яшихина Е.В.¹ (студент)

Научный руководитель – канд. экон. наук. Кононова О.В.¹

¹Университет ИТМО

e-mail: e.yashikhina@yandex.ru

Работа выполнена в рамках НИР №621304 «Разработка сервиса тематической кластеризации корпуса текстов «Развитие цифрового государственного управления в Российской Федерации» на основе машинного обучения».

В работе приводятся виды аннотации и основные функциональные возможности инструмента Доссано. Рассмотрены подходы к решению задачи распознавания именованных сущностей и сложность ее реализации. Выделены метрики для оценки моделей распознавания именованных сущностей.

Ключевые слова: распознавание именованных сущностей, классификация текстов, аннотация, обработка данных, нейронная сеть.

Модели машинного обучения нуждаются в размеченных данных, но большинство данных, как правило, собираются в необработанном формате. Поэтому первым шагом перед построением модели будет подготовка данных – разметка корпусов текстов экспертами предметной области. Для этого стоит использовать инструмент аннотации текста с открытым исходным кодом Доссано [1]. Он позволяет создать набор данных за несколько часов, создав проект, загрузив данные и начав аннотацию.

В Доссано различают следующие типы аннотаций:

- ручная аннотация (пользователь может начать определять разметку, после назначения сущностей, нажав на кнопку метки);
- полуавтоматическая аннотация (модель машинного обучения загружается для автоматического аннотирования загруженных данных; пользователь может внести некоторые исправления в неправильные прогнозы);
- автоматическая аннотация (загруженный документ автоматически аннотируется обученным документом модели машинного обучения) [2].

Существует три основных функции, которые можно выполнить с помощью Доссано. Каждая функция имеет свой собственный эффективный формат вывода:

- анализ настроений – это задача классификации текстов и тем по различным категориям. Поскольку текст может относиться к нескольким категориям, аннотация может быть многозначной;
- машинный перевод – одна из последовательных задач, которая позволяет выполнять несколько ответов, если может быть предоставлено более одного ответа;
- распознавание именованных сущностей (Named Entity Recognition, NER) – одна из задач маркировки последовательностей.

В то время как нейролингвистическое программирование помогает в понимании человеческого языка, методы машинного обучения помогают в анализе, категоризации и повышении точности анализируемых данных с течением времени. Чтобы понять, что такое

сущность, модель извлечения сущности должна сначала быть способна идентифицировать слова или строку слов, которые составляют сущность. Тогда он должен быть в состоянии классифицировать их соответствующим образом.

Текст выбирается и аннотируется в соответствии с определенными сущностями. Необходимо обнаружить, какая из последовательности слов – это именованная сущность и понять, к какому классу эта именованная сущность относится [3]. Чтобы узнать, что такое сущность, модель NER должна уметь обнаруживать слово или строку слов, которые образуют сущность, и знать, к какой категории сущностей она принадлежит. В общей постановке задачи необходимо разместить входной корпус текста и выделить самую сущность, ее расположение, людей, связанных с этой сущностью.

Где используется NER? Распознавание именованных сущностей используется везде, где обрабатываются большие объемы контента. Например, СМИ, журналисты ежедневно генерируют большое количество онлайн-контента.

Алгоритм извлечения сущностей может автоматически сканировать целые статьи и определять, какие важные люди, организации, продукты, места или общие ключевые слова появляются в них. После извлечения этой информации она помогает автоматически классифицировать статьи в определенных иерархиях. На основе этой информации результаты поиска могут быть составлены более точно, контент может быть разделен на тематические кластеры, сообщения, связанные с контентом, могут отображаться пользователю или разыгрываться целевые рекламные объявления. В дополнение к использованию на новостных порталах рекомендательные функции медиа – сервисов также основаны на распознавании именованных сущностей.

Стандартного набора классов нет, все зависит от задачи исследования. Наибольшее распространение для широкого спектра задач получило выделение таких сущностей, как персона (Per) – имена, фамилии, отчества людей; локация (Loc) – местоположение; организация (Org) – названия организаций, компаний, объединений; разное (Misc) [3]. В эту группу входят все прочие типы сущностей, если их более тщательное разделение не требуется для целей исследования.

На первый взгляд кажется, что с именованными сущностями не должно возникнуть особых сложностей. Ведь большинство из них являются именами собственными, которые пишутся с большой буквой. Но во многих языках, таких как китайский или арабский, нет больших букв, а в некоторых — не только имена собственные пишутся с большой буквы. Например, в немецком языке с большой буквы пишутся вообще все существительные.

Также вызывает трудности обстоятельство, что именованные сущности редко состоят из одного слова и не всегда очевидны границы сущности. Поэтому различают начало, середину и конец именованной сущности. Для обозначения первого слова используют префикс «B» (beginning), для последнего слова — «E», а для всех слов между — «I» (intermediate) [4]. Таким образом, задача сводится к пословной классификации.

Обычно для построения классификатора на большом количестве текстов с разметкой именованных сущностей тренируют нейросетевую модель. Процесс обучения инструмента NER состоит из сравнения функций с аннотированным вручную текстом; так называемый учебный корпус. Этими функциями могут выступать списки слов, которые включают все имена персонажей, местоположения и организации. Кроме того, части речи могут быть включены не только для самой именованной сущности, но и для предыдущих и следующих слов. Только сочетание различных функций приводит к хорошим результатам с автоматическим обнаружением именованных сущностей может иметь разные значения в разных контекстах.

Хорошие результаты дают и классические классификаторы, работающие с предзаданным множеством признаков. Для решения задач NER также используют именованные сущности, которые уже собраны в списки, газетеры [4]. Они содержат всевозможные названия стран, континентов, городов, списки имен, организаций. В газетир имен помимо полных имен добавлены распространенные сокращенные имена.

Еще применяются системы, основанные на правилах, в них прописываются шаблонные схемы именованных сущностей. Но проблема такого подхода заключается в том, что подготовка правил требует очень много времени, а малейшее отступление от них приведет к ошибке.

Для оценки моделей NER используют метрики — полнота и точность. Полнота дает отношение правильно помеченных сущностей к общему количеству, которое должно было быть помечено. Точность — это отношение правильно помеченных сущностей к общему количеству помеченных. F-мера — это единая метрика, которая сочетает в себе как полноту, так и точность. Одним из предостережений при использовании этих метрик является область действия. В реальном приложении они применяются к реальным именованным сущностям.

Так аннотирование данных вручную происходит медленно. Полуавтоматический подход заключается в предоставлении файла, содержащего набор шаблонов или правил. Другой подход заключается в том, чтобы начать с существующей модели и вручную исправить ошибки [5].

Рекомендуется подход, который заключается в том, что модель помечает однозначные сущности на первом проходе. Последующие проходы используют уже помеченные сущности для устранения неоднозначностей в отношении других сущностей.

По сути, итеративные подходы, которые начинаются с шаблонов, словарей или однозначных сущностей, начинаются с высокой точности, но с низкой полнотой. Полнота улучшается с помощью уже помеченных сущностей и обратной связи. Аннотированные вручную данные обычно резервируются для оценки модели.

Аннотация важна не только для создания наборов данных, но и для понимания области исследования. Обработка естественного языка наблюдала сдвиг парадигмы в точности за последние несколько лет. Эти большие успехи во многом можно отнести к появлению глубокого обучения. Данная работа дает понимание данных, возможность экспертам обсудить разногласия и лучше понять трудности в области исследования.

Литература

1. Доссано. [Электронный ресурс]. – Режим доступа: <https://github.com/doccano/doccano> (дата обращения: 02.02.2022).
2. Musabeyezu F. Comparative study of annotation tools and techniques // the African University of Science and Technology. – 2019. – С. 35–39.
3. Kauffmann A., Rey F., Atanassova I., Gaudinat A., Greenfield P., Madinier H., Cardey S. Indirectly named entity // Journal of Computer-Assisted Linguistic Research 5. – 2021. – С. 27–46.
4. Маслова М.А., Дмитриев А.С., Холкин Д.О. Методы распознавания именованных сущностей в русском языке // Инженерный вестник Дона. – 2021. – №7. – С. 64–77.
5. Ivanitskiy R., Shipilo A., Kovriguina L. Russian named entities recognition and classification using distributed word and phrase representations // SIMBig. – 2016. – С. 150–156.

СОДЕРЖАНИЕ

КОМПЬЮТЕРНЫЕ ТЕХНОЛОГИИ И УПРАВЛЕНИЕ.....	10
Антонов Р.М. СЕМАНТИЧЕСКАЯ СЕГМЕНТАЦИЯ КАК ЧАСТНЫЙ СЛУЧАЙ ЗАДАЧИ КЛАССИФИКАЦИИ	11
Барина Я.В., Менщиков А.А. ИССЛЕДОВАНИЕ ЭФФЕКТИВНОСТИ ВНЕДРЕНИЯ HONEУРОТ-РЕШЕНИЙ В ОБЛАСТИ ИОТ	14
Басова Т.В., Андреев Ю.С. РАЗРАБОТКА МЕТОДИКИ АКТИВНОГО КОНТРОЛЯ РЕЖУЩЕГО ИНСТРУМЕНТА НА СТАНКАХ С ЧПУ В УСЛОВИЯХ БЕЗЛЮДНОГО ПРОИЗВОДСТВА	17
Власова М.А., Шевченко А.Н., Дементьев М.М., Безмен Г.В. РАЗРАБОТКА СИСТЕМЫ ТЕРМОСТАБИЛИЗАЦИИ ДЛЯ УСТАНОВКИ ПО КОНТРОЛЮ ПАРАМЕТРОВ ГАЗОВЫХ ЯЧЕЕК КВАНТОВЫХ ДАТЧИКОВ	20
Голованов А.А., Беззатеев С.В. КРИПТОГРАФИЧЕСКИЙ МЕТОД ОБЪЕДИНЕНИЯ СООБЩЕНИЙ НЕСКОЛЬКИХ ПОЛЬЗОВАТЕЛЕЙ В ОДНОМ ШИФРТЕКСТЕ.....	24
Гонтарь Д.А., Егоров Д.А. ОБЗОР МЕТОДОВ ПОСТРОЕНИЯ ТЕМПЕРАТУРНОЙ МОДЕЛИ ВОЛОКОННО- ОПТИЧЕСКОГО ГИРОСКОПА	28
Григорьев А.М., Менщиков А.А. АНАЛИЗ ЭФФЕКТИВНОСТИ ПРИМЕНЕНИЯ НЕЙРОННЫХ СЕТЕЙ В ЗАДАЧАХ КЛАССИФИКАЦИИ ОБЪЕКТОВ КИБЕРАТАКИ НА ОСНОВЕ ОТКРЫТЫХ ИСТОЧНИКОВ ДАННЫХ.....	31
Губин М.С. АВТОМАТИЗАЦИЯ ТЕХНОЛОГИИ КОМПЛЕКСНОГО НЕРАЗРУШАЮЩЕГО КОНТРОЛЯ КАЧЕСТВА ПОКРЫТИЙ ИЗДЕЛИЙ СЛОЖНОЙ ГЕОМЕТРИЧЕСКОЙ ФОРМЫ.....	33
Дакуо Ж.-М.Н., Иогансон И.Д., Давыдов В.В. СРАВНИТЕЛЬНЫЙ АНАЛИЗ ID-BASED СИСТЕМ, ОСНОВАННЫХ НА ФУНКЦИЯХ СПАРИВАНИЯ НА ЭЛЛИПТИЧЕСКИХ КРИВЫХ.....	36
Дербенёва Е.С., Авксентьева Е.Ю. СРЕДСТВО ДЛЯ АВТОМАТИЗАЦИИ КОНТЕЙНЕРИЗАЦИИ ПРИЛОЖЕНИЙ С ВОЗМОЖНОСТЬЮ РАСШИРЕНИЯ ИСПОЛЬЗУЕМОГО СТЕКА ТЕХНОЛОГИЙ	40
Завитаев А.С. АНАЛИЗ ИСТОЧНИКОВ ПОГРЕШНОСТЕЙ ЯДЕРНОГО МАГНИТНОГО ГИРОСКОПА ПРИ МЕХАНИЧЕСКИХ ВОЗДЕЙСТВИЯХ.....	45
Иогансон И.Д., Дакуо Ж.-М.Н., Беззатеев С.В. СХЕМА РАЗДЕЛЕНИЯ СЕКРЕТА НА МНОГОМЕРНЫХ УРАВНЕНИЯХ	47
Зименко К.В., Афанасьев М.Я. ПРОБЛЕМЫ ПРИМЕНЕНИЯ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА В АДДИТИВНОМ ПРОИЗВОДСТВЕ	50
Климова Е.Н., Шевченко А.Н., Драницына Е.В. ИССЛЕДОВАНИЕ ВОЗМОЖНОСТИ ИСПОЛЬЗОВАНИЯ АМОРФНОГО ЖЕЛЕЗА В КАЧЕСТВЕ МАГНИТНОГО ЭКРАНА ВОГ.....	54

Колганов О.А., Федоров А.В., Лукьянов Г.Н. ОБОСНОВАНИЕ ВЫБОРА МЕТОДА РЕГИСТРАЦИИ ПАРАМЕТРОВ ДВИЖЕНИЯ УДАРНИКА ПРИ ДИНАМИЧЕСКОМ ИНДЕНТИРОВАНИИ	58
Колесникова Д.В., Юрьева Р.А. ГИПЕРАВТОМАТИЗАЦИЯ КАК СПОСОБ ПОВЫШЕНИЯ РЕНТАБЕЛЬНОСТИ ПРОИЗВОДСТВА	62
Кочарян А.М., Мирошник А.А., Хмелев А.П., Москальчук В.Р., Еремина В.В. ОБЕСПЕЧЕНИЕ ИНФОРМИРОВАНИЯ ЖИТЕЛЕЙ ОБЛАСТИ О ХОДЕ РЕАЛИЗАЦИИ ПРОГРАММЫ БЛАГОУСТРОЙСТВА ДВОРОВЫХ ТЕРРИТОРИЙ ПРИ ПОМОЩИ ТЕХНОЛОГИЙ МАШИННОГО ОБУЧЕНИЯ	67
Кузнецова В.Е., Юрьева Р.А. ИССЛЕДОВАНИЕ И РАЗРАБОТКА ОБЛАЧНОЙ АРХИТЕКТУРЫ РАСПРЕДЕЛЕННОГО ПРЕДПРИЯТИЯ.....	70
Кустов Е.Ф., Леевик А.Г. ОБЗОР И АНАЛИЗ СХЕМ РАЗДЕЛЕНИЯ СЕКРЕТА, ПОСТРОЕННЫХ НА РЕШЕТКАХ	76
Леевик А.Г., Кустов Е.Ф. АНАЛИЗ СОВРЕМЕННЫХ СХЕМ ДОКАЗАТЕЛЬСТВА С НУЛЕВЫМ РАЗГЛАШЕНИЕМ, ОСНОВАННЫХ НА ТЕОРИИ РЕШЕТОК	80
Лихтенберг А.М., Беззатеев С.В. КВАНТОВОЕ РАСПРЕДЕЛЕНИЕ КЛЮЧЕЙ: ОБЗОР РЕШАЕМЫХ ЗАДАЧ	84
Ляховенко Ю.А. ИССЛЕДОВАНИЕ ВЛИЯНИЯ HUMAN FACTORS НА СКОРОСТЬ ДВИЖЕНИЯ РЕЛЬСОВОГО ГОРОДСКОГО ТРАНСПОРТА	88
Малый В.В., Костюхин А.С., Кинжагулов И.Ю. РАЗРАБОТКА ТЕХНОЛОГИИ КОНТРОЛЯ ПАЯНЫХ СОЕДИНЕНИЙ ТЕПЛООБМЕННЫХ АППАРАТОВ.....	93
Меренков Д.Н. ОБЕСПЕЧЕНИЕ ИНФОРМАЦИОННОЙ БЕЗОПАСНОСТИ В МЕДИЦИНСКИХ ВЕБ- ПРИЛОЖЕНИЯХ	96
Милантьев С.А., Литвинов П.А., Бессмертный И.А. ПОИСК АНОМАЛЬНЫХ ПОСЛЕДОВАТЕЛЬНОСТЕЙ ПИКСЕЛЕЙ В РАМКАХ АНАЛИЗА СНИМКОВ КОЖНЫХ ПОРАЖЕНИЙ	100
Мыськив И.И., Заколдаев Д.А. ОЦЕНКА УСТОЙЧИВОСТИ ТЕХНИЧЕСКОГО ЗРЕНИЯ В БЕСПИЛОТНЫХ ТРАНСПОРТНЫХ СРЕДСТВАХ.....	104
Никитина А.О., Киприянов К.В. ИСПОЛЬЗОВАНИЕ СИСТЕМНОЙ ИНЖЕНЕРИИ И ИМИТАЦИОННОГО МОДЕЛИРОВАНИЯ ДЛЯ СКЛАДСКИХ СИСТЕМ	108
Павлов В.А., Мордвинцева Ю.А., Кофонов Р.В. ПРИМЕНЕНИЕ АДДИТИВНЫХ ТЕХНОЛОГИЙ В ПРОТОТИПИРОВАНИИ И ТОПОЛОГИЧЕСКОЙ ОПТИМИЗАЦИИ В РОБОТОТЕХНИКЕ	112
Павлов В.А., Полицинский А.С., Макаренко З.В., Сунцова Д.И., Бахолдин П.П., Маргун А.А., Кремлев А.С. РАЗРАБОТКА TELEGRAM-БОТА «ОПРЕДЕЛЕНИЕ УРОВНЯ ГОТОВНОСТИ ТЕХНОЛОГИИ».....	115

Сергеев А.С., Алифанова И.Е. МОДЕЛЬ РАЗДЕЛЬНО-СОВМЕЩЕННОГО ОПТИКО-АКУСТИЧЕСКОГО ПРЕОБРАЗОВАТЕЛЯ	121
Сергеев А.С. РАЗРАБОТКА ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ ДЛЯ ОБРАБОТКИ, ВИЗУАЛИЗАЦИИ И ХРАНЕНИЯ РЕЗУЛЬТАТОВ УЛЬТРАЗВУКОВОГО КОНТРОЛЯ.....	126
Татаров Д.А., Менщиков А.А. МЕТОД ТЕСТИРОВАНИЯ БЕЗОПАСНОСТИ КОМПЬЮТЕРНЫХ СИСТЕМ НА ОСНОВЕ КОНЦЕПЦИИ ЦИФРОВЫХ ДВОЙНИКОВ	128
Титов Р.У. , Моторин А.В. АЛГОРИТМЫ СОВМЕСТНОЙ НАВИГАЦИИ И КАРТОГРАФИРОВАНИЯ, ПОСТРОЕННЫЕ НА ОСНОВЕ ПРОЦЕДУР ЧАСТИЧНОГО АНАЛИТИЧЕСКОГО ИНТЕГРИРОВАНИЯ.....	131
Хуцаева А.Ф., Давыдов В.В. АНАЛИЗ СТОЙКОСТИ КРИПТОГРАФИЧЕСКИХ СХЕМ, ОСНОВАННЫХ НА ИЗОГЕНИЯХ НЕСУПЕРСИНГУЛЯРНЫХ ЭЛЛИПТИЧЕСКИХ КРИВЫХ.....	135
Чернов Р.И., Тихомиров А.В., Еремук В.В., Гришенцев А.Ю. РАЗРАБОТКА МЕТОДА ПОВЫШЕНИЯ ЭФФЕКТИВНОСТИ МАГНИТООПТИЧЕСКОГО ДАТЧИКА ГРАДИЕНТА МАГНИТНОГО ПОЛЯ С ТОРСИОННЫМ ПОДВЕСОМ ЧУВСТВИТЕЛЬНОГО ЭЛЕМЕНТА ЗА СЧЕТ УМЕНЬШЕНИЯ ВРЕМЕНИ УСПОКОЕНИЯ	139
Шамраев А.Д., Колюбин С.А. ИССЛЕДОВАНИЕ АЛГОРИТМОВ УПРАВЛЕНИЯ ДВИЖЕНИЕМ ЧЕТВЕРОНОГИМИ РОБОТАМИ.....	143
Якимов Я.Д. ПОСТРОЕНИЕ СИСТЕМЫ ДЛЯ ОБНАРУЖЕНИЯ ПРЕПЯТСТВИЙ НА ЖЕЛЕЗНОДОРОЖНОМ ТРАНСПОРТЕ.....	149
ФОТОНИКА.....	152
Азина Л.В., Панов Д.Ю., Спиридонов В.А., Волковицкий В.Е., Романов А.Е. РАЗРАБОТКА АКТИВНОГО ЭЛЕМЕНТА ОПТИЧЕСКОГО ДАТЧИКА ИСКРОВОГО ПРОБОЯ ДЛЯ РАСПРЕДЕЛИТЕЛЬНЫХ ЩИТОВ ЭЛЕКТРОСЕТЕЙ	153
Астахов Ю.И. ВЛИЯНИЕ ИСКУССТВЕННОГО СВЕТА НА АТМОСФЕРУ ОБЩЕСТВЕННЫХ ПРОСТРАНСТВ	157
Безуглый А.М. ИССЛЕДОВАНИЕ ВОЗМОЖНОСТИ УВЕЛИЧЕНИЯ РАБОЧЕГО ПОЛЯ ПРИ ИСПОЛЬЗОВАНИИ НЕПЛОСКИХ ПРИЕМНЫХ ПЛОЩАДОК ЗЕРКАЛЬНЫХ КАМЕР.....	160
Бердюгина Д.А., Пастухов А.С., Смирнов Ю.Ю. АНАЛИЗ ХАРАКТЕРИСТИК СИСТЕМ ВИДЕОНАБЛЮДЕНИЯ НА ОСНОВЕ АВТОНОМНЫХ И ОБЫЧНЫХ КАМЕР	162
Захарова Ю.Е. СВЕТОВЫЕ ХАРАКТЕРИСТИКИ КАК ИНСТРУМЕНТ СОЗДАНИЯ ВИЗУАЛЬНЫХ АКЦЕНТОВ В ПРОСТРАНСТВЕ ГОРОДА В ТЕМНОЕ ВРЕМЯ СУТОК	165
Капустина Е.В., Гресько В.Р., Сергеев М.М. ВЛИЯНИЕ ЛЕГИРОВАНИЯ ПЕРЕХОДНЫМИ МЕТАЛЛАМИ НА ФОТОПОТЕМНЕНИЕ В ХАЛЬКОГЕНИДНЫХ ПЛЁНКАХ	169

Каракотов Р.Р., Анфиногентов С.А., Кузнецов А.А. ИССЛЕДОВАНИЕ КОЛЕБАНИЙ В СОЛНЕЧНОЙ ПЛАЗМЕ ПУТЁМ ПОСТРОЕНИЯ ПРОСТРАНСТВЕННО- ВРЕМЕННЫХ ДИАГРАММ	174
Карзов М.М., Пименов А.Ю., Никандров Г.В., Маргарянц Н.Б. ОЦЕНКА ЭНЕРГЕТИЧЕСКИХ ХАРАКТЕРИСТИК ЛАЗЕРНОГО СКАНИРУЮЩЕГО ДАЛЬНОМЕРА ДЛЯ БЕСПИЛОТНОГО ЛЕТАТЕЛЬНОГО АППАРАТА	179
Кутько Г.А. ПРИНЦИПЫ ФОРМИРОВАНИЯ ГОРОДСКИХ СОБЫТИЙНЫХ ПРОСТРАНСТВ СРЕДСТВАМИ ИСКУССТВЕННОГО ОСВЕЩЕНИЯ.....	183
Николаева Н.В. ВЛИЯНИЕ ПРИЕМОВ АРХИТЕКТУРНОГО ОСВЕЩЕНИЯ НА ВЫЯВЛЕНИЕ ДОМИНАНТНОСТИ ОБЪЕКТОВ ГОРОДСКОЙ СРЕДЫ	187
Смотровая Е.Е., Михайлова О.С., Сыренова Т.Е., Ежова В.В. ПОДГОТОВКА ИЗОБРАЖЕНИЙ, ПОЛУЧЕННЫХ КАМЕРОЙ ВСЕГО НЕБА, ДЛЯ КОМПЛЕКСНОГО АНАЛИЗА ВОЛНОВЫХ СОБЫТИЙ В МАГНИТОСФЕРЕ	191
Хлынов Р.Д., Рыжова В.А., Чан В.Т. ПОРТАТИВНАЯ ОПТИКО-ЭЛЕКТРОННАЯ СИСТЕМА ДЛЯ МОНИТОРИНГА ОПТИЧЕСКИ НЕОДНОРОДНЫХ СРЕД.....	196
Чеботарев С.В., Брусницын А.В., Филиппов И.М., Быстрянцева Н.В. ИССЛЕДОВАНИЯ АЛГОРИТМОВ УПРАВЛЕНИЯ ОСВЕЩЕНИЕМ МНОГОПОЛЬЗОВАТЕЛЬСКОГО ПРОСТРАНСТВА НА ОСНОВЕ АНАЛИЗА ПРОСТРАНСТВЕННОГО РАСПРЕДЕЛЕНИЯ ЯРКОСТИ	201
ТРАНСЛЯЦИОННЫЕ ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ	204
Абрамовский И.С. ИСПОЛЬЗОВАНИЕ КОРПУСНОЙ ЛИНГВИСТИКИ В СОЦИАЛЬНО- ПОЛИТИЧЕСКИХ ИССЛЕДОВАНИЯХ	205
Алейников С.А. АЛГОРИТМЫ ИДЕНТИФИКАЦИИ ЦИФРОВЫХ КОММУНИКАТИВНЫХ АКТОВ.....	208
Баглык Е.В., Смолин А.А. ЭФФЕКТИВНОСТЬ РАСПОЛОЖЕНИЯ БУРГЕР-МЕНЮ НА МОБИЛЬНЫХ ВЕРСИЯХ САЙТОВ	213
Беген П.Н. ПРИМЕНЕНИЕ ТЕХНОЛОГИЙ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА В ТЕМАТИЧЕСКОЙ КЛАСТЕРИЗАЦИИ ТЕКСТОВ	217
Бейлин М.Т., Осипов Н.А. ИССЛЕДОВАНИЕ ТРЕБОВАНИЙ, ПРЕДЪЯВЛЯЕМЫХ МЕТОДАМИ АНАЛИЗА БОЛЬШИХ ДАННЫХ, К ОРГАНИЗАЦИИ ВЫЧИСЛИТЕЛЬНОЙ СЕТИ ИНФОКОММУНИКАЦИОННОЙ СИСТЕМЫ	220
Волков А.А., Тесля Н.Н. ИНСТРУМЕНТЫ АВТОМАТИЧЕСКОГО ПОЛУЧЕНИЯ ОНТОЛОГИИ ИЗ РЕЛЯЦИОННЫХ БАЗ ДАННЫХ	226
Волков А.А., Тесля Н.Н. ОБЗОР МЕТОДОВ ДОСТУПА К ОТКРЫТЫМ ДАННЫМ В СИСТЕМЕ УМНОГО ГОРОДА	230

Двойникова А.А., Карпов А.А. РАСПОЗНАВАНИЕ КАШЛЯ С ПОМОЩЬЮ АНАЛИЗА СПЕКТРОГРАММ.....	235
Дырко А.В., Шмелева И.А. АНАЛИЗ ПРАКТИК И ПРОЕКТОВ В ОБЛАСТИ УСТОЙЧИВОГО ТУРИЗМА И ТУРИСТСКИХ ОБЪЕКТОВ В СЕВЕРНЫХ СТРАНАХ	240
Евсеева З.А., Гудым А.А., Шмелева И.А., Джумагулова А.Ф. ИССЛЕДОВАНИЕ ПРИМЕНИМОСТИ ЦИФРОВЫХ СЕРВИСОВ В ГЕНДЕРНО- ОРИЕНТИРОВАННОМ ПРОЕКТИРОВАНИИ ГОРОДСКОЙ СРЕДЫ	246
Ермолаева А.А., Карсаков А.А., Смолин А.А. ИССЛЕДОВАНИЕ ПАТТЕРНОВ ВЗАИМОДЕЙСТВИЯ ПОЛЬЗОВАТЕЛЕЙ С СЕРВИСАМИ РЕКОМЕНДАЦИЙ	253
Зайцева И.С. РЕИНЖИНИРИНГ БИЗНЕС-ПРОЦЕССА ОПЕРАТОРА СВЯЗИ: ПРЕДПРОЕКТНОЕ ИССЛЕДОВАНИЕ	259
Звягинцева Д.А., Иванов А.Б., Ульянов В.И. РАЗРАБОТКА МЕТОДА ВЫДЕЛЕНИЯ ПРИЗНАКОВ ИЗ МЕТАГЕНОМНЫХ ОБРАЗЦОВ, ОСНОВАННОГО НА РАСКРАСКЕ ГРАФА ДЕ БРЮИНА.....	263
Иванов А.Б., Ульянов В.И. ИЗВЛЕЧЕНИЕ МЕТАГЕНОМНЫХ ПРИЗНАКОВ И МАШИННОЕ ОБУЧЕНИЕ ДЛЯ ДИАГНОСТИРОВАНИЯ ВОСПАЛИТЕЛЬНЫХ ЗАБОЛЕВАНИЙ КИШЕЧНИКА	267
Казак В.А. ПРОЕКТИРОВАНИЕ СИСТЕМЫ ОЦЕНКИ ДОСТУПНОСТИ ОБЪЕКТОВ ГОРОДСКОЙ СРЕДЫ ПО ДАННЫМ ИЗ СОЦИАЛЬНЫХ СЕТЕЙ	271
Калинин П.С., Лисицинский С.Г., Орлов Г.М. ИССЛЕДОВАНИЕ ПРИМЕНЕНИЯ ДИАЛОГОВЫХ АГЕНТОВ В СФЕРЕ ЗДРАВООХРАНЕНИЯ.....	275
Касицына М.С., Осипов Н.А. ПРИМЕНЕНИЕ КОГНИТИВНЫХ ТЕХНОЛОГИЙ В ПОСТРОЕНИИ ИНДИВИДУАЛЬНОЙ ТРАЕКТОРИИ ОБУЧЕНИЯ.....	280
Кинёв И.Е., Гебель Г.В., Жданов Д.Д. ИССЛЕДОВАНИЕ НЕЙРОСЕТЕВЫХ ТЕХНОЛОГИЙ ДЛЯ СОЗДАНИЯ УСЛОВИЙ, СООТВЕТСТВУЮЩИХ ЕСТЕСТВЕННОЙ АККОМОДАЦИИ.....	285
Клокова Е.И., Авксентьева Е.Ю. МОНЕТИЗАЦИЯ НОВЫХ МОБИЛЬНЫХ ИГР.....	289
Козырева А.И. ОБЗОР ПРАКТИКИ ПРИМЕНЕНИЯ ТЕХНОЛОГИЙ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА ДЛЯ АНАЛИЗА ПОВЕДЕНИЯ ПОЛЬЗОВАТЕЛЕЙ ТУРИСТИЧЕСКИХ СЕРВИСОВ.....	292
Красноперова Ю.А., Соловьёва В.В. РАЗРАБОТКА МЕТОДА ВЫБОРА УЧАСТКОВ УЛИЧНО-ДОРОЖНОЙ СЕТИ ДЛЯ СОЗДАНИЯ ПЕШЕХОДНЫХ ОБЩЕСТВЕННЫХ ПРОСТРАНСТВ НА ПРИМЕРЕ САНКТ-ПЕТЕРБУРГА.....	297
Курбатов М.С., Зудилова Т.В. ИССЛЕДОВАНИЕ И РАЗРАБОТКА ПОДХОДОВ К СОЗДАНИЮ ЧАТ-БОТА НА ПРИМЕРЕ ИСУ ИТМО.....	301
Матвеев Г.А., Войтюк Т.Е. ПРИМЕНЕНИЕ БЛОКЧЕЙНА НА ГОСУДАРСТВЕННОМ УРОВНЕ И АНАЛИЗ ЕГО УСТОЙЧИВОСТИ.....	306

Меркушев А.Е., Максимова Т.Г. РАЗРАБОТКА КЛАССИФИКАТОРА ОБЛАЧНЫХ УСЛУГ ДЛЯ СИСТЕМЫ УПРАВЛЕНИЯ ЗАТРАТАМИ И ПЛАНИРОВАНИЯ ОБЛАЧНОЙ ИТ-ИНФРАСТРУКТУРЫ ПРЕДПРИЯТИЙ.....	310
Мещеряков И.Д. СИСТЕМА АВТОМАТИЧЕСКОГО ПОДБОРА ПАРАМЕТРОВ АЛГОРИТМОВ ДИАРИЗАЦИИ.....	314
Низомутдинов Б.А., Углова А.Б. РАЗРАБОТКА МЕТОДА ОЦЕНКИ ДОСТУПНОСТИ ГОРОДСКОЙ СРЕДЫ С ИСПОЛЬЗОВАНИЕМ ДАННЫХ СОЦИАЛЬНЫХ СЕТЕЙ.....	319
Огурцова Е.В. ТЕХНОЛОГИИ ВЫЯВЛЕНИЯ В ТЕКСТАХ ИНДИКАТОРОВ СОЦИАЛЬНОГО САМОЧУВСТВИЯ	323
Парошин Я.Д., Платонов Н.С., Чёрный П.П., Маркина Т.А. ИССЛЕДОВАНИЕ ВОЗМОЖНОСТИ ОБЪЕДИНЕНИЯ СИСТЕМ "УМНЫЙ ДОМ" РАЗЛИЧНЫХ ПРОИЗВОДИТЕЛЕЙ	327
Пац К.М., Виноградова Е.А., Сергушичев А.А. ПРИМЕНЕНИЕ СТРУКТУРНЫХ ДЕСКРИПТОРОВ ДЛЯ ПОИСКА ЗАКОНОМЕРНОСТЕЙ В 3D-СТРУКТУРАХ РЕЦЕПТОРА ВИТАМИНА D	332
Попов В.В., Иванов А.Б., Ульянов В.И. СРАВНЕНИЕ МЕТОДОВ ВЫДЕЛЕНИЯ ПРИЗНАКОВ ИЗ МЕТАГЕНОМОВ ДЛЯ РЕШЕНИЯ ЗАДАЧИ КЛАССИФИКАЦИИ ВОСПАЛИТЕЛЬНЫХ ЗАБОЛЕВАНИЙ КИШЕЧНИКА.....	336
Ролинский С.О., Двойникова А.А., Карпов А.А. АНАЛИТИЧЕСКИЙ ОБЗОР МЕТОДОВ ИЗВЛЕЧЕНИЯ ТЕКСТОВЫХ ТРАНСКРИПЦИЙ ИЗ РЕЧЕВЫХ ВЫСКАЗЫВАНИЙ.....	341
Рудалева Е.А., Попов Д.В. РАЗРАБОТКА МЕТОДИКИ ИССЛЕДОВАНИЯ ОШИБОК ДИАРИЗАЦИИ ПО ЛЕКСИКО-ГРАММАТИЧЕСКИМ ПАРАМЕТРАМ.....	346
Рождественский Д.А., Смолин А.А. ПРОЕКТИРОВАНИЕ ПОЛЬЗОВАТЕЛЬСКОГО ИНТЕРФЕЙСА ПРИЛОЖЕНИЯ С ИСПОЛЬЗОВАНИЕМ АДАПТИВНОЙ ДОПОЛНЕННОЙ РЕАЛЬНОСТИ ДЛЯ ПРОИЗВОДСТВА МЕДИАКОНТЕНТА	352
Ружевич М.С., Мынбаев К.Д. ИССЛЕДОВАНИЕ ОПТИЧЕСКИХ ХАРАКТЕРИСТИК ОБРАЗЦОВ $\text{Hg}_{0.7}\text{Cd}_{0.3}\text{Te}$, ВЫРАЩЕННЫХ МЕТОДОМ МПЭ.....	356
Рыжков Н.М. , Суренков Н.С. АВТОМАТИЗАЦИЯ АГРЕГАЦИИ АНАЛИТИЧЕСКИХ ДАННЫХ О ЛЕЧЕНИИ ОНКОЛОГИЧЕСКИХ ЗАБОЛЕВАНИЙ	359
Рюмина Е.В., Карпов А.А. АНАЛИТИЧЕСКИЙ ОБЗОР КОРПУСОВ ДЛЯ АВТОМАТИЧЕСКОГО ОЦЕНИВАНИЯ ПСИХОФИЗИЧЕСКИХ СОСТОЯНИЙ ЧЕЛОВЕКА.....	363
Рюмина Е.В., Иванько Д.В., Карпов А.А. ОБЗОР МУЛЬТИМОДАЛЬНЫХ КОРПУСОВ ДЛЯ ИССЛЕДОВАНИЯ ВЛИЯНИЯ ЭМОЦИОНАЛЬНОГО СОСТОЯНИЯ ДИКТОРА НА АВТОМАТИЧЕСКОЕ РАСПОЗНАВАНИЕ ФРАЗ ПО ГУБАМ.....	366

Самородова М.Э., Чирковский А.Д. ВЛИЯНИЕ АКУСТИЧЕСКИХ УСЛОВИЙ НА РАБОТУ СИСТЕМЫ ОБНАРУЖЕНИЯ АТАК ПОВТОРНОГО ВОСПРОИЗВЕДЕНИЯ ПРИ ГОЛОСОВОЙ АУТЕНТИФИКАЦИИ.....	370
Селезнева Я.М., Зенкин А.М. МОДЕЛИ И МЕТОДЫ ОБЕСПЕЧЕНИЯ КИБЕРБЕЗОПАСНОСТИ СИСТЕМ ЦИФРОВОЙ ЭКОНОМИКИ НА ОСНОВЕ МАШИННОГО ОБУЧЕНИЯ.....	376
Сергеев А.И., Зудилова Т.В. ПОДХОДЫ К РАЗРАБОТКЕ ПРИЛОЖЕНИЙ С NOSQL И SQL СУБД.....	380
Фокина С.Г. ИССЛЕДОВАНИЕ ИНФОРМАЦИОННЫХ ПРОЦЕССОВ ПОДДЕРЖКИ ВРАЧЕБНОЙ ПРАКТИКИ ВРАЧА СТАЦИОНАРА.....	384
Хабирова А.Р. СРАВНИТЕЛЬНЫЙ АНАЛИЗ ФУНКЦИОНАЛЬНОСТИ ОТКРЫТОЙ ЧАСТИ ГОСУДАРСТВЕННЫХ ИНФОРМАЦИОННЫХ СИСТЕМ В ЧАСТИ ПУБЛИКАЦИИ ДАННЫХ ОБ ОБЪЕКТАХ ГОРОДСКОЙ СРЕДЫ.....	389
Чагина П.А., Ананченко И.В. АНАЛИЗ ВОЗМОЖНОСТИ ПРИМЕНЕНИЯ МАШИННОГО ОБУЧЕНИЯ И НЕЙРОННЫХ СЕТЕЙ ДЛЯ ПРОЕКТИРОВАНИЯ И ОПТИМИЗАЦИИ ТОРГОВЫХ РОБОТОВ.....	393
Чижик А.В., Мельникова С.А., Ковальчук С.В. ПРИЛОЖЕНИЯ МОБИЛЬНОГО ЗДРАВООХРАНЕНИЯ: ИССЛЕДОВАНИЕ МОТИВАЦИИ РУССКОГОВОРЯЩИХ ПОЛЬЗОВАТЕЛЕЙ.....	396
Шаров Д.Д., Бурак П.В., Стафеев С.К. РАЗРАБОТКА ПРИЛОЖЕНИЯ ДЛЯ ИНТЕРАКТИВНЫХ ГОЛОГРАФИЧЕСКИХ ВЫСТАВОК.....	402
Шнейдер Л., Шмелева И.А. УСТОЙЧИВАЯ МОБИЛЬНОСТЬ: ИСПОЛЬЗОВАНИЕ ЭЛЕКТРО-САМОКАТОВ В ЕВРОПЕЙСКИХ ГОРОДАХ: ПРОБЛЕМЫ И ПЕРСПЕКТИВЫ.....	405
Шостина А.Д., Иванов А.Б., Ульянцев В.И. МЕТАШЕРМАНТ: ИСПОЛЬЗОВАНИЕ HI-C СВЯЗЕЙ ПРИ ВЫДЕЛЕНИИ ГЕНОМНОГО КОНТЕКСТА ИЗ МЕТАГЕНОМОВ.....	411
Эседов Н.Э., Жданов Д.Д., Жданов А.Д. СТАНОВЛЕНИЕ ЕСТЕСТВЕННОЙ КОНВЕРГЕНЦИИ ЧЕЛОВЕЧЕСКОГО ЗРЕНИЯ В ВИДЕОПРОЗРАЧНЫХ СИСТЕМАХ СМЕШАННОЙ РЕАЛЬНОСТИ.....	414
Яшихина Е.В. РАСПОЗНАВАНИЕ ИМЕНОВАННЫХ СУЩНОСТЕЙ С ПОМОЩЬЮ DOCCANO.....	419

**Альманах научных работ
молодых ученых
Университета ИТМО
Том 2**

Редакционно-издательский отдел Университета ИТМО

Зав. РИО

Дизайн обложки

Вёрстка

Подписано к печати 22.07.2022

Заказ № 4675 от 22.07.2022

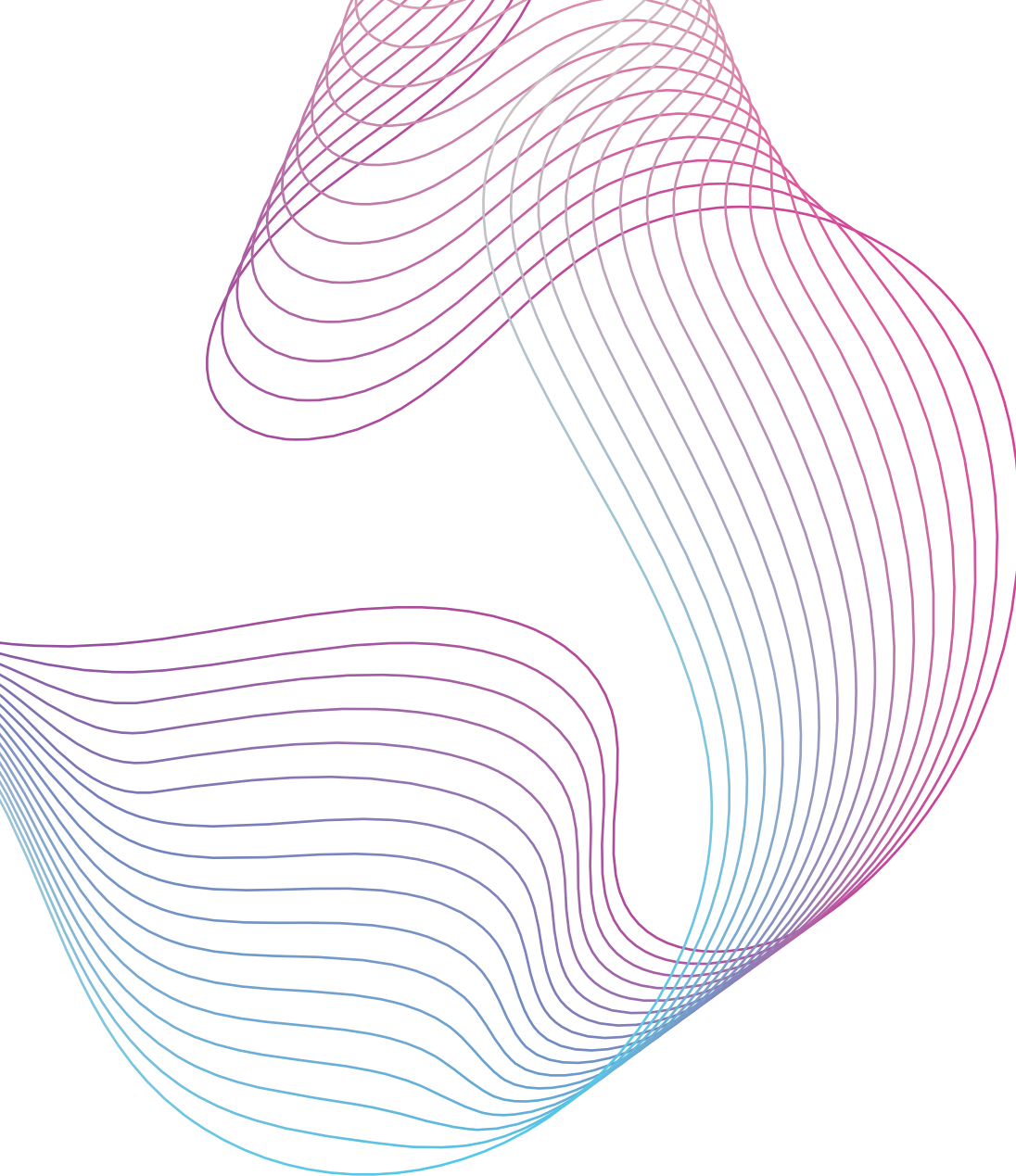
Тираж 100 экз.

Печатается в авторской редакции

Н.Ф. Гусарова

Н.Ю. Комаровская

К.Д. Бутылкина



ISBN 978-5-7577-0671-9



9 785757 706719

**Редакционно-издательский отдел
Университета ИТМО**

197101, Санкт-Петербург, Кронверкский пр., 49