МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ

НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ УНИВЕРСИТЕТ ИТМО

XII Конгресс молодых ученых
Сборник научных трудов

Том 2

Санкт-Петербург
2023

УДК 082.2, ББК – 94.3
XII Конгресс молодых ученых. Сборник научных трудов. Том 2. – СПб.: Университет ИТМО, 2023. – 486 с.

Издание содержит результаты научных работ молодых ученых, доложенные на XII Конгрессе молодых ученых по тематикам: биотехнологии, инфокоммуникационные технологии, информационная безопасность, искусственный интеллект и смежные с ней области, цифровая урбанистика.

ISBN 978-5-7577-0699-3
ISBN 978-5-7577-0701-3 (Том 2)

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, 2023
© Авторы, 2023

Редакционная коллегия сборника научных трудов
XII Конгресса молодых ученых. Том 2

Тематика биотехнологии
Председатель – Сазонова Е.К. кандидат технических наук, доцент факультета биотехнологий
Бычкова Е.С., доктор технических наук, доцент, доцент факультета биотехнологий
Еремеева Н.Б., кандидат технических наук, доцент факультета биотехнологий
Кригер О.В., доктор технических наук, профессор факультета биотехнологий

Тематика инфокоммуникационные технологии
Председатель – Грудинин В.А. кандидат технических наук, доцент факультета инфокоммуникационных технологий
Кынев С.М. кандидат физико-математических наук, заведующий лабораторией научно-образовательного центра фотоники и оптоинформатики
Хоружников С.Э. кандидат физико-математических наук, доцент, директор национального центра квантового интернета

Тематика информационная безопасность
Председатель – Заколдаев Д.А. доцент, кандидат технических наук, декан факультета безопасности информационных технологий
Комаров И.И., кандидат физико-математических наук, доцент, доцент факультета безопасности информационных технологий
Будько М.Ю. кандидат технических наук, доцент, доцент факультета безопасности информационных технологий
Коробейников А.Г. доктор технических наук, профессор, профессор факультета безопасности информационных технологий
Попов И.Ю. кандидат технических наук, доцент факультета безопасности информационных технологий
Филькова И.А. заместитель декана факультета безопасности информационных технологий

Тематика искусственный интеллект и смежные с ней области
Председатель – Ватьян А.С. кандидат технических наук, доцент факультета инфокоммуникационных технологий
Графеева Н.Г. кандидат физико-математических наук, доцент высшей школы цифровой культуры
Гусарова Н.Ф. кандидат технических наук, старший научный сотрудник, доцент факультета инфокоммуникационных технологий
Духанов А.В. доктор технических наук, доцент, доцент факультета цифровых трансформаций
Муромцев Д.И. кандидат технических наук, доцент, доцент факультета программной инженерии и компьютерной техники
Филатова А.А. инженер исследовательского центра в сфере искусственного интеллекта "Сильный искусственный интеллект в промышленности"

Тематика цифровая урбанистика
Председатель – Митягин С.А. кандидат технических наук, директор института дизайна и урбанистики
Горнова Г.В. доктор философских наук, доцент института дизайна и урбанистики
Чугунов А.В. кандидат политических наук, доцент, директор центра технологий электронного правительства
Шмелева И.А. кандидат психологических наук, доцент, доцент института дизайна и урбанистики.

ВВЕДЕНИЕ

Издание содержит результаты научных работ молодых ученых, доложенные на XII Конгрессе молодых ученых Университета ИТМО, проходившего с 3 по 6 апреля 2023 года, по тематикам: биотехнологии, инфокоммуникационные технологии, информационная безопасность, искусственный интеллект и смежные с ней области, цифровая урбанистика.

Конгресс проводится ежегодно с 2004 года в целях стимулирования научно-технической деятельности молодых ученых, приобретения ими опыта публичных выступлений, повышения научного уровня магистерских диссертаций и апробации выпускных квалификационных работ бакалавров и специалистов за 2023/2024 учебный год, а также в целях привлечения талантливых специалистов из высших учебных заведений в научные лаборатории Университета ИТМО, ознакомления с достижениями молодых ученых как профессионалов в различных отраслях, развития сотрудничества и привлечения абитуриентов в магистратуру и аспирантуру Университета ИТМО, совершенствования качества подготовки квалифицированных кадров, создания условий для развития научного творчества обучающейся молодежи, интеграции их в научно-технологичное пространство.

В рамках Конгресса молодых ученых Университета ИТМО проходят заседания секций, круглых столов с участием молодых ученых, выставка разработок молодых ученых.

БИОТЕХНОЛОГИИ

УДК 664-00

ОБОСНОВАНИЕ ВЫБОРА ИНГРЕДИЕНТОВ ДЛЯ ПРОИЗВОДСТВА ФУНКЦИОНАЛЬНОЙ СНЕКОВОЙ ПРОДУКЦИИ ДЛЯ ПРОФИЛАКТИКИ ОНКОЛОГИЧЕСКИХ ЗАБОЛЕВАНИЙ

Беспалая Д.С.[1] (студент)
Научный руководитель – доцент Еремеева Н.Б.
1 – Университет ИТМО
e-mail: darinabespalaya@gmail.com

Аннотация
В данной работе рассматриваются профилактические и противоопухолевые свойства экстракта растения Куркума (*Curcuma extract*), с возможностью дальнейшего использования его в качестве ингредиента в функциональном питании. На основе собранных данных был разработан сбалансированный по своему ингредиентному составу продукт, содержащий куркумин и другие ингредиенты с доказанной противоопухолевой активностью.

Ключевые слова
Куркумин, функциональный ингредиент, профилактика онкологических заболеваний, химиотерапия, снековый батончик, разработка рецептур, аминокислоты, витамины.

В современном мире остро стоит проблема распространения онкологических заболеваний. За последние 10 лет количество заболеваний, связанных со злокачественными новообразованиями, в России выросло на 20,5% [1]. В последние годы онкология поднялась на второе место по количеству диагностируемых заболеваний. В ближайшие десятилетия ожидается, что онкологические заболевания станут одной из ведущих причин всех заболеваний в мире. Важно отметить, что статистика выявляемости онкологических заболеваний растет не столько из-за увеличения количества заболеваний по сравнению с прошлым, сколько из-за большей доступности диагностики, её эффективности. Что, безусловно, является положительным событием. Но существует и обратная взаимосвязь: например, количество меланомы – злокачественной опухоли кожных покровов – выросло в связи с увеличившимся воздействием ультрафиолетового излучения от солнца как результатом разрушения озонового слоя планеты. Исходя из данной информации, становится понятно, что некоторые виды онкологии прогрессируют в своем количестве с течением времени. Исходя из собранной информации о том, что статистика по злокачественным новообразованиям увеличивается, а лечение все еще является малодоступным и дорогостоящим, необходимо найти эффективные и простые в реализации методы профилактики онкологических заболеваний. На данный момент на рынке функционального питания представлены продукты, содержащие ингредиенты с онкопротекторными свойствами. Систематический анализ данных продуктов по сбалансированности их состава, по массовой доле функциональных ингредиентов и по универсальности для питания разных групп населения выявил некоторые недостатки. Исходя из вышесказанного, необходимо предложить состав функционального продукта, превосходящий существующие по ряду основных показателей.

Основной целью работы является подбор и обоснование оптимального ингредиентного состава продукта функциональной направленности для профилактики онкологических заболеваний и облегчения течения химиотерапии и лучевой терапии - снекового батончика. В состав продукта входят арахисовая паста, овсяные хлопья крупного помола, чернослив, грецкий орех, сушеные ягоды малины, куркумин. Данный состав был выбран как окончательный на основе приведенного ниже обоснования.

В данной работе особое внимание уделено основному функциональному ингредиенту продукта – экстракту растения Куркума (*Curcuma extract*). Куркумин – это один из куркуминоидов, входящих в состав корня растений куркума (*Curcuma)* – род

травянистых многолетних растений из семейства имбирных (*Zingiberaceae*). Куркумин в виде порошка характерного желтого цвета используется в качестве пищевого красителя, входит в Кодекс Алиментариус (*Codex Alimentarius*) под кодом Е100. Куркумин представляет собой гидрофобный полифенол. Он взаимодействует с такими молекулами, как медиаторы воспаления, факторы роста, ферменты, белки-носители, ионы металлов, супрессоры опухолей, факторы транскрипции, онкопротеины и клеточные нуклеиновые кислоты [3]. Химическая структура куркумина, обуславливающая его способность связывать различные молекулы важна для его антиканцерогенных и опухолеподавляющих свойств, рассматриваемых в данной работе. Посредством различных клеточных и молекулярных механизмов куркумин уменьшает канцерогенез и рост опухолей [2]. В ряде исследований сообщалось, что куркумин увеличивает вероятность выживания пациентов и снижает уровень опухолевых маркеров.

Всесторонний систематический обзор, представленный в некоторых статьях, также подтверждает, что куркумин уменьшает побочные эффекты химиотерапии или лучевой терапии, что приводит к улучшению качества жизни пациентов [3]. Куркумин обладает антиоксидантным, противоартритным, антиамилоидным (уменьшает нейрональное повреждение при нейродегенеративных заболеваниях, например, болезнь Альцгеймера, которая имеет вероятность развития на фоне проведения химиотерапии), противоишемическим и противовоспалительным действиями. Он обладает защитной и терапевтической эффективностью против рака кожи, полости рта, легких, поджелудочной железы и кишечного тракта [3].

Выбор таких высокоэнергетических ингредиентов, как грецкий орех и чернослив, в составе обусловлен предполагаемым ухудшенным состоянием потребителя, проходящего химиотерапию и необходимостью восполнить организм недостающими в связи со сниженным аппетитом калориями. Также орехи являются источником фитостеринов – молекул, естественным образом связанных с холестерином, содержащихся в оболочке растительной клетки. Фитостерины предупреждают рак и атеросклероз [4].

Сушеные ягоды малины являются источниками флавонов (класс флавоноидов), полифенольных соединений, витамина С, органических кислот: яблочной, лимонной и янтарной. Флавоноиды способны ингибировать основные этапы развития раковых клеток. Влияние флавоноидов на этап инициирования (зарождения) раковой клетки связано с их доказанным антиоксидантным действием. Показано, что ряд флавоноидов регулируют активность протеинов клеточного цикла, что приводит к уменьшению возникновения раковых клеток [4].

Витамин С – аскорбиновая кислота – содержащаяся в ягодах малины также влияет на рост и развитие опухолевых клеток. Около 200 мг аскорбиновой кислоты в день уменьшает количество эпинефрина – гормона стресса, выделяемого надпочечниками. Витамин стимулирует антителообразование и благотворно влияет на иммунную систему. Доказано, что именно иммунитет ответственен за раннее распознавание мутировавших клеток в организме. Неограниченное деление мутировавших клеток может приводить к развитию злокачественных опухолей. Некоторые другие витамины, содержащиеся в ингредиентах составных частей продукта, также обладают противоопухолевой активностью – витамины Е, РР, B_9, D, А. Арахисовая паста в составе продукта укрепляет иммунитет и снижает уровень липопротеинов низкой плотности, ответственных за возникновение холестериновых бляшек на стенках сосудов. Уменьшает риск появления сахарного диабета, улучшает работу сердечно-сосудистой и нервной систем. Арахисовая паста также является источником белка (в 100 граммах содержится около 30% суточной нормы белка), витаминов РР, Е, B_9, а также меди и магния. В продукте не употребляется добавленный сахар. Этот выбор обусловлен желанием сделать продукт подходящим для диабетического питания. Наибольшее

количество злокачественных новообразований появляется в более позднем возрасте. Большинство случаев заболеваемости возникает в возрасте более 60 лет: среди мужчин – 71,3%, а среди женщин – 65,8% [1]. Похожую статистику можно наблюдать, говоря о сахарном диабете 2 типа: В типичных случаях он развивается в возрасте старше 40 лет. Более того, по статистическим данным IDF (International Diabetes Federation) сахарный диабет является прогрессирующим заболеванием. С 2019 по 2021 года число случаев выявления сахарного диабета выросло на 74 миллиона. Примерно 10 % всех людей страдают диабетом.

В готовом продукте был проведен расчет аминокислотного состава белка, аминокислотного скора (формула 1), содержания жирных кислот, макро- и микроэлементов, витаминов, а также энергетической ценности продукта. Полученные результаты представлены в таблицах 1-5.

$$АКС = \frac{[НАК]продукт}{[НАК]эталон} * 100,\ \%, \quad (1)$$

где [НАК] продукта – содержание определенной незаменимой аминокислоты в составе белковой части исследуемого продукта, г/100 белка; [НАК] эталона – содержание определенной незаменимой аминокислоты в составе идеального (эталонного) белка, г/100 белка (табл. 1).

Таблица 1

Аминокислотный состав белка в продукте

Наименование незаменимой аминокислоты	Содержание, мг/г белка		АКС, %
	эталонный	исследуемый	
Изолейцин	40	49	122
Лейцин	70	95	136
Лизин	55	50	92
Гистидин	35	35	100
Метионин + цистеин	60	43	71
Треонин	40	44	109
Триптофан	10	21	207
Валин	50	65	130
Фенилаланин + тирозин	16	132	825

Таким образом, лимитирующими аминокислотами являются: лизин, метионин + цистеин (табл. 2).

Таблица 2

Содержание жирных кислот в продукте

Наименование продукта	На 44 г (порция продукта)		
Снековый батончик для функционального питания	Насыщенные	Мононенасыщенные	Полиненасыщенные
	3,7	9,7	7,9

Полученное соотношение НЖК: МНЖК: ПНЖК = 3,7 : 9,7 : 7,9 (4:10:8). Данное соотношение можно сравнить с установленной нормой потребления жирных кислот: 30 % насыщенных, 60 % мононенасыщенных и 10 % полиненасыщенных (3:6:1) [5].

В рационе современного человека наблюдается дефицит поступления полиненасыщенных жирных кислот в организм. Исходя из полученных данных о соотношении жирных кислот в продукте, можно сделать вывод, что разработанный продукт подходит для восполнения данного дефицита (табл. 3-5).

Таблица 3

Содержание макро- и микроэлементов в продукте

Наименование продукта	Содержание макро - и микроэлементов, мг/ 100г продукта								
	K	Ca	Mg	P	Fe	Cu	Mn	Zn	Se
Снековый батончик для функционального питания	519,4	54,3	135,6	304,8	3,7	0,4	1,4	2,0	0,9
	Доля восполнения суточной потребности в макро - и микроэлементах при употреблении 100г продукта, %								
	20,8	5,4	33,9	38,1	36,8	43,2	67,7	16,9	1,8

Таблица 4

Содержание витаминов в продукте

Наименование продукта	Содержание витаминов, мг/100 г продукта								
	B_1	B_2	B_5	B_6	B_9	C	PP	H	E
Снековый батончик для функционального питания	0,3	0,1	0,4	0,35	0,1	1,7	5,1	0,01	4,9
	Доля восполнения суточной потребности в витаминах при употреблении 100г продукта, %								
	20,3	5,7	8,1	17,4	25,5	1,9	25,6	21,9	32,7

Таблица 5

Биохимический состав 100 г сырья

Наименование сырья	Белки, г	Жиры, г	Углеводы, г
Арахисовая паста	24,06	49,94	21,57
Хлопья овсяные (Геркулес)	2,4	4	14,8
Куркума	7,83	9,88	43,83
Грецкий орех	15,23	65,21	7,01
Малина сушеная	1,2	0,56	5,44
Чернослив	2,3	0,7	57,5

Расчет пищевой ценности продукта. Расчет белковой составляющей в порции арахисовой пасты:

$$\text{Масса арахисовой пасты} * \frac{\text{количество белка в 100 г}}{100\ \text{г}} \quad (2)$$

$$14 * \frac{24,06}{100} = 3,3684\text{г белка на порцию арахисовой пасты (14 г)}$$

Расчеты по остальным белковым составляющим продукта, а также углеводам и жирам были проведены аналогично. Исходя из полученных результатов была рассчитана энергетическая ценность порции продукта (44 г).

$$4\ \text{ккал} * 4,3853\ \text{г белка в порции продукта} + 9\ \text{ккал} * 9,3742\ \text{г жира} + 4\ \text{ккал} * 8,3547\text{г углеводов} = 135\ \text{ккал на порцию} = 564,84\ \text{кДж на порцию}$$

Энергетическая ценность 100 г продукта - 306 ккал; 1283,7 кДж.

Также было определено общее содержание фенольных соеденений с помощью модифицированной версии метода Фолина-Чокалтеу. Результаты выражали в моль эквивалента галловой кислоты в л экстракта (моль ГК/л). Получили 558 мг галловой кислоты на 100 г исходного продукта. Содержание флавоноидов в объектах измеряли с использованием модифицированного метода. Содержание флавоноидов – 56 мг катехина на 100 г исходного сырья, что составляет более половины суточной потребности в катехинах для взрослого человека.

По итогам данной работы и проведенных в ходе нее анализов и расчетов, можно прийти к выводу, что разработанный продукт сбалансирован по своему ингредиентному составу. Расчет пищевой и биологической ценности позволяет готовить о том, что продукт является источником многих незаменимых аминокислот, а также некоторых витаминов и микроэлементов. Отсутствие добавленного сахара позволяет вводить продукт в рацион людям, которым необходимо ограничить его потребление. Небольшой вес конечной порции (44 г) дает возможность употреблять нормированное количество ккал. Состав снековового батончика, разработанного в данном исследовании, является оптимальным и отвечает требованиям, предъявляемым для диабетического питания.

Литература

1. [Электронный ресурс]. URL: https://rosstat.gov.ru (дата обращения: 25.01.2023).
2. Mansouri K. Clinical effects of curcumin in enhancing cancer therapy: A systematic review / K. Mansouri, S. Rasoulpoor, A. Daneshkhah et al. // BMC Cancer. 2020. Vol. 20. P. 791. doi: 10.1186/s12885-020-07256-8.
3. Banik U. Curcumin: the spicy modulator of breast carcinogenesis / Banik U., Parasuraman S., Adhikary A.K., et al. // BMC Cancer. 2017. Vol. 36 (98). https://doi.org/10.1186/s13046-017-0566-5.
4. Зверев Я.Ф. Противоопухолевая активность флавоноидов. Бюллетень сибирской медицины / 2019. 18 (2). С. 181–194. doi: 10.20538/1682-0363-2019-2-181-19.
5. Поздняковский В.М. Гигиенические основы питания, качество и безопасность пищевых продуктов / Учебник. 5-е изд., испр. и доп. Новосибирск: изд-во Сиб. унив., 2007. 455 с.

УДК 504.06

ОСОБЕННОСТИ РАЗРАБОТКИ ЭКОЛОГИЧЕСКИХ ДЕКЛАРАЦИЙ ДЛЯ РЕСУРСОСБЕРЕГАЮЩИХ КОМПЛЕКСНЫХ ТЕХНОЛОГИЙ ПЕРЕРАБОТКИ БИОСЫРЬЯ

Гаврецкая А.В.[1] (студент), **Шестель А.А.**[1] (студент)
Научный руководитель –д.т.н., профессор Куприна Е.Э.[1]
1 – Университет ИТМО
e-mail: gavretskaya@yandex.ru

Работа выполнена в рамках темы НИР №620142 «Ресурсосберегающие экологически безопасные технологии биоактивных пищевых ингредиентов с заданными свойствами из биоресурсов и вторичных сырьевых источников растительного и животного происхождения для обеспечения здорового питания населения РФ с использованием пищевой комбинаторики».

Аннотация
В работе рассмотрены основные виды экологических маркировок, их особенности и принципы работы с ними. Разработана экологическая декларация III типа на основе оценки жизненного цикла рыбного полуфабриката с повышенной пищевой ценностью. Определены основные особенности в составлении экологической декларации для комплексных технологий переработки биосырья.

Ключевые слова
Негативное воздействие, окружающая среда, производство, границы системы, функциональный продукт.

Целью экологических деклараций является повышение спроса и производства товара, оказывающего меньшее негативное воздействие на окружающую среду, с помощью предоставления достоверной информации о продукте [1]. В том числе в рамках экономических взаимодействий B2C и B2B. Выделяют три типа экологических маркировок (рис. 1). В зависимости от целей преследуемых компанией используют разные типы экологических деклараций [1, 2].

Рис. 1. Типы экологических маркировок

Тип I – экологическая маркировка. Данный тип является добровольным. Особенностью является то, что программа оценки разрабатывается третьей стороной. Она же выдает лицензию на использование экологической маркировки на продуктах, исходя из анализа их жизненного цикла. Данный вид экологический маркировки наиболее востребован и жестко контролируется.

Второй тип экологической маркировки – экологическое заявление продавца или экологическая самодекларация. Оно не требует сертификации независимой третьей

стороной, в связи с чем имеет минимальные требования к проверке. Но также вызывает и меньшее доверие у потребителей и партнеров.

Последний, третий, тип – это экологическая декларация. Это добровольная программа, в ходе которой независимый орган или отрасль экономики разрабатывает требования, основанные на оценке жизненного цикла продукции, определяет формы участия независимых третьих сторон, устанавливает параметры для экологической декларации типа III и служит для сравнения продукции разных категорий [3]. Также экологические декларации могут разрабатываться самими производителями, но затем необходимо подтверждение независимой третьей стороной.

Разработка экологических деклараций проводится на основе расчетов оценки жизненного цикла [4]. В рамках данной работы расчет ОЖЦ осуществлялся при применении специализированного программного обеспечения и базы данных MIPS Вуппертальского института климата. Разрабатывалась экологическая декларация на функциональный продукт питания с использованием белкового гидролизата. Процесс производства представляет собой комплексную схему переработки биосырья с получением рыбного полуфабриката с повышенной пищевой ценностью.

Для проведения ОЖЦ были определены границы системы. Среди особенностей при разработке экологической декларации для комплексных технологий переработки является получение сырья из вторичных ресурсов, которые имеют отрицательную добавочную стоимость.

Также подобные производства позволяют минимизировать отходы, что является несомненным плюсом в рамках оценки влияния производимой продукции на окружающую среду.

На рисунке 2 приведена технологическая схема производства. Из нее следует, что в процессе производства активно задействованы отходы рыбопереработки: кость и кожа форели. Также из данной схемы можно заметить, что белковый гидролизат образующийся при получении покрытия, и являющийся в данном случае отходом отправляется в параллельный производственный цикл, где является компонентом рыбного фарша.

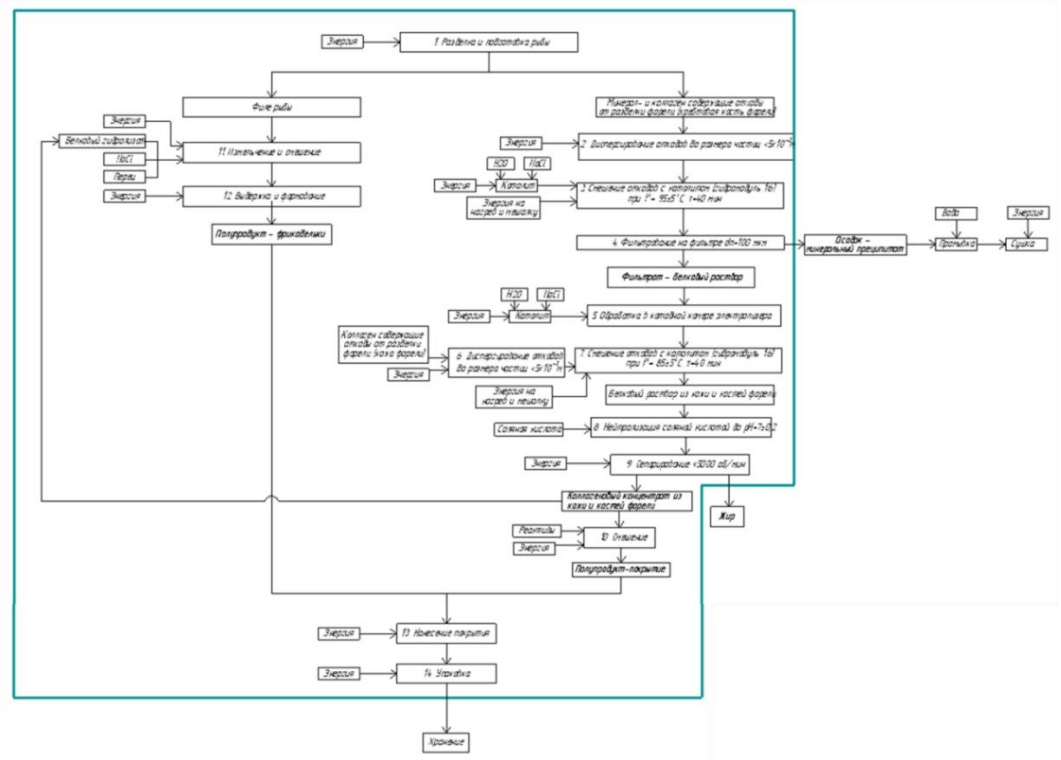


Рис. 2. Границы системы

Особенностью результатов ОЖЦ является распределение показателей по категориям воздействия. Как правило, наибольшее негативное воздействие подобных предприятий рыбной переработки оказывается на водные ресурсы (рис. 3). Это вызвано тем, что ОЖЦ учитывает и ресурсы, которые были затрачены на вылов рыбы, а также учитывают изъятые биотические ресурсы.

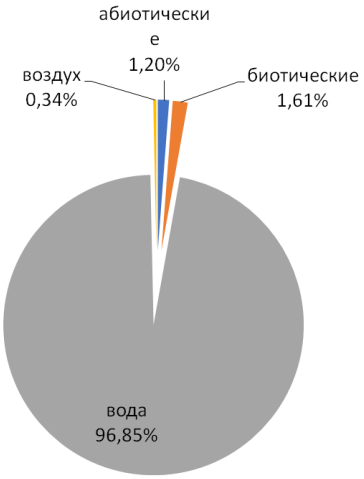


Рис. 3. Распределение воздействий

Также все выходные потоки распределяют по категориям воздействий на окружающую среду. Классически выделяют следующие категории воздействия:

1. Эвтрофикация (ПЭ).
2. Закисление (ПЗ).
3. Глобальное потепление (ПГП).
4. Образование фотохимического смога (ПФОО).
5. Образование побочного продукта.

Особенность данного ОЖЦ заключается в том, что мы учитываем выработку электроэнергии на ТЭЦ, которое обеспечивается за счет сжигания природного газа. В связи с этим стадии потребляющие большое количество электроэнергии наносят значительный ущерб окружающей среде за счет выделения парниковых газов при производстве электроэнергии. Такими стадиями являются – стадия 3 – смешение отходов с католитом, стадия 5 – обработка в катодной камере, стадия 7 – смешение с католитом. Соотношение оказываемого влияния можно видеть на рисунке 4.

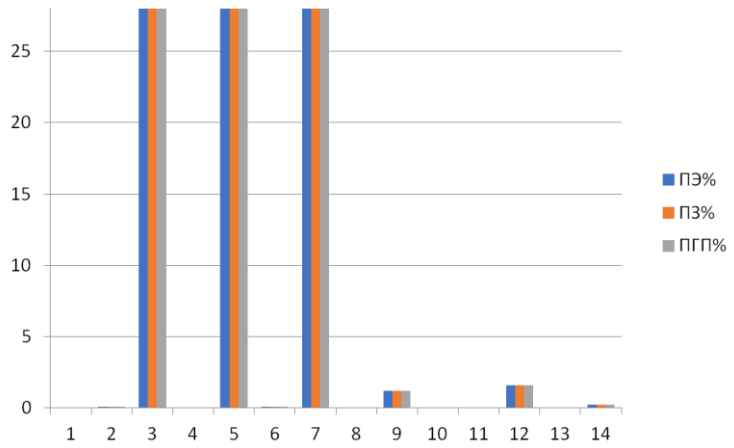


Рис. 4. Распределение воздействий по стадиям

Особенностью данного производства является использование электролиза в технологической схеме. При замене данных стадий на альтернативные категории воздействия изменятся, однако, оказываемое негативное влияние увеличится за счет образования едких сточных вод. Таким образом, несмотря на высокие энергозатраты, данная технологическая схема является наиболее экологичной для производства.

В ходе работы определены основные факторы, оказывающие влияние на окружающую среду. Разработана экологическая декларация для функционального продукта питания с использованием белкового гидролизата, которая приведена ниже. Для дальнейшей регистрации экологическая декларация должна быть верифицирована независимой экспертной организацией.

Экологическая декларация продукции

Руководящий документ: ГОСТ Р 51956-2002

Экологическая декларация на функциональные продукты питания с использованием белкового гидролизата.

Информация о производителе: ИТМО, факультет Биотехнологий
Информация о продукте: рыбный функциональный продукт питания, изготавливаемый с использованием белкового гидролизата
Функциональная единица: 10 кг рыбного функционального продукта питания
Принцип установки границ системы: «от ворот до ворот»
Стадии ЖЦ:

- Разделка и подготовка рыбы
- Диспергирование отходов
- Обработка отходов в катодной камере электролизера
- Нейтрализация и сепарирование
- Приготовление покрытия
- Получение фарша
- Формирование полуфабрикатов
- Нанесение покрытия
- Упаковка продукта

Дата составления декларации: февраль 2023 г.
Основные экологические показатели:

Потребление энергетических ресурсов	2,96 кВт-ч электроэнергии
Потребление водных ресурсов	32,56 кг
Закисление	113,17 мол.H^+
Эвтрофикация	31258,18 г O_2
Выбросы парниковых газов	780146,35 г-экв. CO_2
Образование отходов	5,25 кг

Литература

1. ГОСТ Р ИСО 14025-2012 Этикетки и декларации экологические. Экологические декларации типа III. Принципы и процедуры. М.: Стандартинформ. 2014. 24 с.
2. ГОСТ Р ИСО 14024-2000 Этикетки и декларации экологические. экологическая маркировка типа I. Принципы и процедуры. М.: Издательство стандартов. 2001. 12 с.
3. Казанцева А.Н., Маликова О.И. Экологическая маркировка как инструмент перехода к устойчивому развитию [Электронный ресурс]. Режим доступа: https://cyberleninka.ru/article/n/ekologicheskaya-markirovka-kak-instrument-perehoda-k-ustoychivomu-razvitiyu (дата обращения: 11.04.2023).
4. ГОСТ Р ИСО 14040-2010 Экологический менеджмент. Оценка жизненного цикла. Принципы и структура. М.: Стандартинформ. 2010. 18 с.

УДК 664.959

ТЕХНОЛОГИЯ ПОЛУЧЕНИЯ РЫБНЫХ ПОЛУФАБРИКАТОВ СО СБАЛАНСИРОВАННЫМ СОСТАВОМ ДЛЯ ГЕРОНТОЛОГИЧЕСКОГО ПИТАНИЯ

Данилюк М.А.[1] (аспирант)
Научный руководитель – д.т.н., профессор Ишевский А.Л.[1]
1 – Университет ИТМО
e-mail: madaniliuk@itmo.ru

Аннотация
Питание населения всегда являлось важной стратегической политикой любого государства. При разработке продуктов питания должны учитываться особенности групп населения, для которых они разрабатываются, пищевая ценность, качество и безопасность. Становиться актуальным разработка рыбных полуфабрикатов со сбалансированным составом для геронтологического питания, так как рыбные продукты содержат незаменимые аминокислоты в необходимых количествах, а также отличаются хорошей усвояемостью.

Ключевые слова
Рыбные полуфабрикаты, питание населения, геронтологическое питание, рыба, форель.

Питание населения является важной стратегической политикой любой страны. В последнее время растет внимание к рациону населения в том числе к питанию пожилых людей. Большое количество работ в отечественной и зарубежной литературе посвящено разработке принципов правильного и сбалансированного геронтологического питания. Учитывая возрастающее количество заболеваний, стрессовых факторов, неправильного питания населения, одной из наиболее важных проблем является расширение ассортимента и обеспечение населения высококачественными диетическими продуктами питания. По рекомендациям ВОЗ рыбу, особо богатую полиненасыщенными жирными кислотами, людям старшего поколения необходимо употреблять два-три раза в неделю. Полиненасыщенные жирные кислоты представляют собой профилактическое средство от сердечно-сосудистых заболеваний [1]. Рыбные продукты содержат незаменимые аминокислоты в необходимых количествах, а также отличаются хорошей усвояемостью, что является немаловажным фактором в геронтологическом питании. Рыбные продукты представляют собой источник полноценного белка [2, 3, 4].

В соответствии со статистическими данными, приведенными РОССТАТ за 2021 год, потребление морепродуктов в Российской Федерации достигает 22 кг на душу населения. К сожалению, в сложившейся в последнее время экономической ситуации произошло резкое подорожание всей рыбной продукции, что в свою очередь делает ее малодоступной для некоторых слоев населения, включая пожилых людей, поэтому появляется необходимость в выработке рыбных полуфабрикатов с уменьшенной стоимостью.

Для реализации удешевления рыбных полуфабрикатов представляется перспективным использовать переработку гидробионтов с использованием ресурсосберегающей технологии для приготовления рыбного фарша, а также использование частей рыбы после филетирования, чтобы удешевить сырьевые ресурсы и максимально сохранить полезные свойства рыбы, так как в остатках после филетирования все еще содержится приличное количество мяса, кожи и подкожного жира, богатого полиненасыщенными жирными кислотами, рыбьи кости, которые перемалываются до однородной субстанции. При получении полуфабрикатов с использованием данной технологии, добавляя к фаршу филе рыбы фарш, полученный после филитирования, в них увеличивается количество полиненасыщенных жирных кислот, однако, снижается количество белка, что в свою очередь ведет к уменьшению

пищевой ценности получаемых рыбных полуфабрикатов, ухудшаются функционально-технологические свойства фарша, такие как: влагосвязывающая способность (ВСС), жироудерживающая способность (ЖУС), влагоудерживающая (ВУС).

Современные тенденции усовершенствования технологий производства рыбных продуктов направлены на создание сбалансированных продуктов по биологической и пищевой ценности, а также способствующие улучшению и сохранению здоровья населения.

Одним из путей создания рыбных продуктов со сбалансированным содержанием белка, обогащенных полиненасыщенными жирными кислотами, а также для улучшения вкусовых качеств продукции, увеличения биологической ценности рыбной продукции является использование в рецептуре молочных и сывороточных белков, так как они представляют собой полноценные белки и являются легкоусвояемыми.

В исследовании использовалась форель, которая является полужирной рыбой, так как в соответствии с основами геронтологического питания жирные сорта рыб не могут быть рекомендованы к употреблению лицами старшего поколения [5].

Исследованы разные смеси фарша контроля с добавлением 25%, 50%, 75% ресурсосберегающего фарша. В фарше филе форели массовая доля жира составляет 5,1%, ОМЕГА-3 0,52 г на 100 г фарша, а ОМЕГА-6 0,88 г на 100 г фарша, а в ресурсосберегающем фарше массовая доля жира - 14,9%, а ОМЕГА-3 и ОМЕГА-6 в 3 раза больше, чем в фарше филе форели. Таким образом добавление ресурсосберегающего фарша помогает увеличить количество полиненасыщенных жирных кислот в рыбных полуфабрикатах, поэтому они могут быть рекомендованы для геронтологического питания.

Для улучшения функционально-технологических свойств смесей фарша использовались сывороточные и яичные протеины в различной концентрации, так как они позволяют улучшить органолептические свойства готовой продукции, образуют стабильную водо-жировую эмульсию, улучшают связывание воды и жира, а также повышают биологическую ценность.

Определены функционально-технологические свойства фарша, полученного по ресурсосберегающей технологии. Рассчитаны наилучшие соотношения составления ресурсосберегающего фарша с фаршем контролем, приготовленным из филе форели. На рисунке 1 показаны функционально-технологические свойства фаршей и цена, где контроль – это фарш филе форели, смесь фаршей 1 – добавление 25% ресурсосберегающего фарша, смесь фаршей 2 – добавление 50% ресурсосберегающего фарша, смесь фаршей 3 – добавление 50% ресурсосберегающего фарша и ресурсосберегающий фарш 100%.

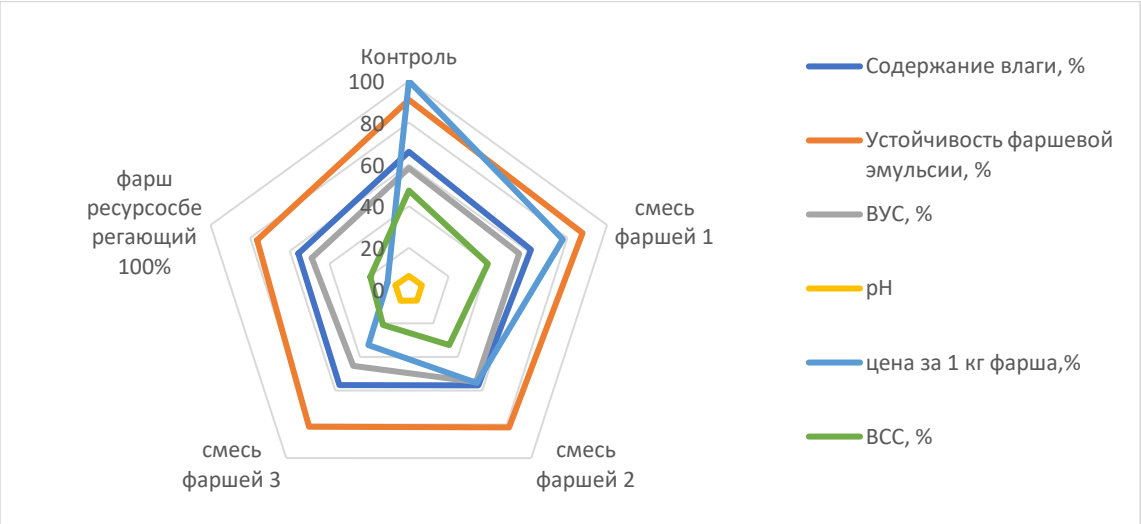


Рис. 1. Показатели фаршей

На рисунке 2 представлены результаты органолептического тестирования готовых полуфабрикатов для различных соотношений фарша с ресурсосберегающим.

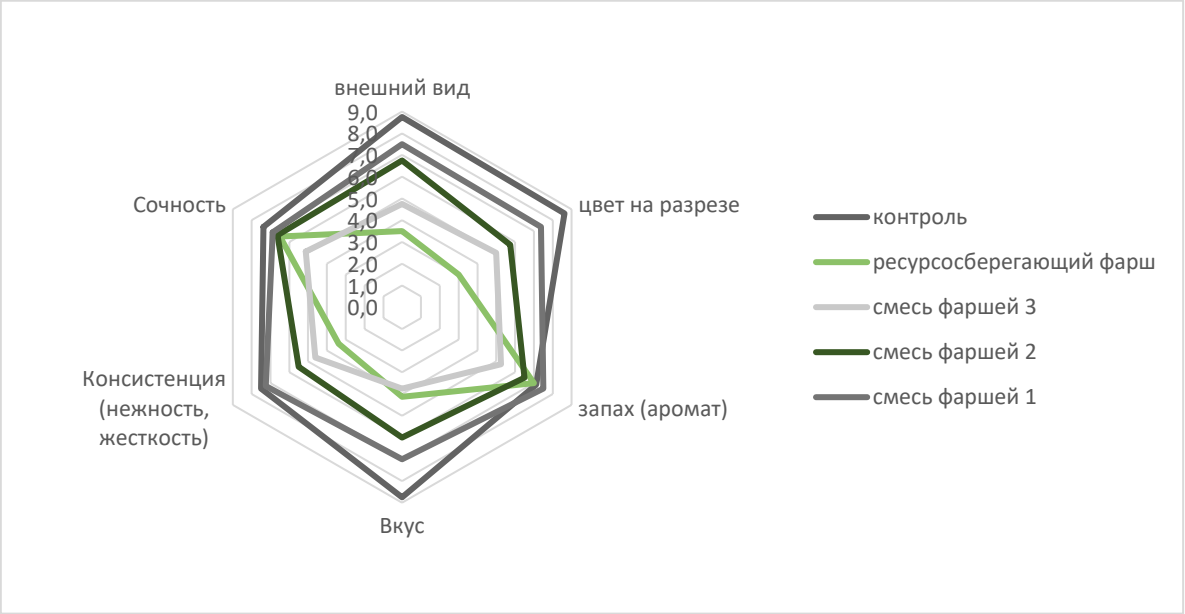


Рис. 2. Органолептические показатели полуфабрикатов

Исследованы лучшие смеси фаршей с добавлением яичного белка и концентрата молочного белка (5%, 10%, 15%, 20%), измерены их функционально-технологические и органолептические показатели.

Результаты исследования показали, что рыбные полуфабрикаты с использованием ресурсосберегающей технологии производства рыбного фарша могут повышать содержание полезных незаменимых жирных кислот ОМЕГА-3 без значительного ухудшения качества продукции. Также, с внесением дополнительных молочных или яичных белков, увеличивается пищевая и биологическая ценность разрабатываемых полуфабрикатов. В полуфабрикатах, содержащих в составе 25% ресурсосберегающего фарша, повышается содержание ОМЕГА-3 и ОМЕГА-6, при этом основные функционально-технологические показатели полуфабрикатов остаются близкими к контрольным образцам. Добавление к фаршу, в состав которого входит 25% ресурсосберегающего фарша, яичного белка в концентрации 5% улучшает аминокислотный скор смеси, органолептические и функционально-технологические показатели.

Литература

1. Calviello G., Serini S. Dietary Omega-3 polyunsaturated fatty acids ans cancer, Springer. 2010. p. 261. ISBN 978-90-481-3578-3.
2. Касьянов Г.И., Иванова Е.Е., ОдинцовА.Б., Студенцова Н.А. Технология переработки рыбы и морепродуктов: Учеб. Пособ.-Ростов-на-Дону: изд.центр «МарТ». 2001. 416 с.
3. Колаковский Э. Технология рыбного фарша. М.: Агропромиздат. 1991. 220 с.
4. Юдина С.Б. Технология геронтологического питания. М.: Лань. 2017. 232 с.
5. Погожева А.В. «Ешь, пей, молодей. Уникальные принципы геродиететики – здорового питания в пожилом возрасте». АСТ. Москва. 2015. с. 208. ISBN 978-5-17-090419-8.

УДК 663.94

ТЕХНОЛОГИЯ ФУНКЦИОНАЛЬНОГО КОФЕЙНОГО ПРОДУКТА С ИСПОЛЬЗОВАНИЕМ ИНУЛИНА

Дерягина А.С.[1] (студент)
Научный руководитель – к.т.н., доцент, ординарный доцент Смотраева И.В.[1]
1 – Университет ИТМО
e-mail: deryagina_asya@mail.ru

Аннотация
В данной работе рассмотрена технология получения напитка, обогащенного высокомолекулярным полисахаридом – инулином. Благодаря своим свойствам такой напиток будет оказывать благоприятный эффект на одну или несколько физиологических функций, корректировать процессы метаболизма в организме человека, повышать его защитные функции и снижать риск развития заболеваний.

Ключевые слова
Корень лопуха, инулин, зеленый кофе, ультразвуковая экстракция, физико-химические показатели.

Лопух большóй (лат. *Arctium láppa*) - вид многолетних травянистых растений семейства Астровые (лат. *Autetceece*). Листья лопуха большого содержат ценный комплекс биологически активных соединений, такие как флавоноиды, дубильные вещества, полисахариды, органические кислоты. Корни лопуха накапливают высокий процент инулина во время их роста. Поскольку инулин – это легкогидролизуемый полисахарид, то гидролизаты корня лопуха могут быть основой безалкогольных напитков, выполняя функцию подсластителя [1]. В связи с этим одно из актуальных направлений современных исследований – разработка безалкогольных напитков на основе инулина лопуха. В России предприятия по производству лопуха находятся в центральной части страны.

Инулин представляет собой полифруктозан, который может быть получен в виде аморфного порошка или в виде кристаллов. Построен из остатков d-фруктозы, связанных β-2,1-связями, и d-глюкозы связанных α-1,2-связями. Помимо лопуха инулин содержится во многих растениях (рис. 1), в промышленном значении наиболее важны *Cichorium intybus* (цикорий) и *Helianthus tuberosus* (топинамбур) [2].

Рис. 1. Содержание инулина в растительном сырье, %

Цель данного исследования – разработать технологию кофейного напитка с использованием сушеного корня лопуха, основой для напитка будет использоваться молотый зеленый кофе.

В ходе разработки функционального напитка, определялись наиболее подходящие параметры выделения водорастворимых полисахаридов (%) из листьев лопуха большого

с помощью ультразвуковой экстракции, проводилась физико-химическая характеристика готового продукта, а также его органолептическая оценка.

В последние годы экстракция с помощью ультразвука используется как новый метод извлечения химических соединений из растений, который имеет преимущества по сравнению с традиционными методами. Метод ультразвуковой экстракции обеспечивает чистоту процесса, повышает эффективность экстракции за счет увеличения проникновения растворителя в клетки растений посредством кавитации [3].

Главной целью экстрагирования было получение наибольшего выхода водорастворимых полисахаридов. Для интенсификации процесса извлечения ВРПС применяли ультразвуковую ванну YAXUN YX2100 объемом 600 мл. Объектами исследования экстрагирования инулина стали 4 образца с параметрами и их выходами, представленными в таблице 1.

Таблица 1

Параметры и выход ВРПС

№ образца	Мощность, Вт	Температура, °С	ВРПС, %
1	30	70	13,91
2	30	90	12,21
3	50	70	9,77
4	50	90	8,44

Рисунок 2 иллюстрирует зависимость выхода водорастворимых полисахаридов от температуры воды. Наибольший выход инулина получился при температуре экстрагирования 70°С.

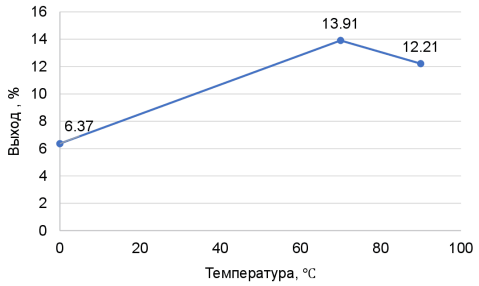


Рис. 2. Зависимость выхода инулина от температуры экстрагента

На рисунке 3 представлена зависимость выхода водорастворимых полисахаридов от мощности. Из графика можно сделать вывод, что при мощности 30 Вт получается максимальный выход инулина.

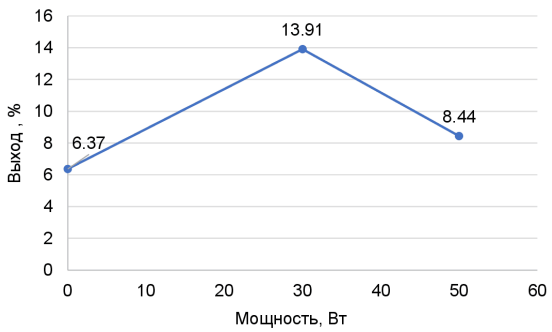


Рис. 3. Зависимость выхода инулина от мощности экстрагента

Основа напитка состояла из 5 г молотого зеленого кофе и 250 мл воды с температурой 98-100°С. После экстрагирования полученный порошок инулина добавлялся в определенном соотношении в раствор с кофе и заваривался 15 минут.

В готовом напитке исследовались такие физико-химические показатели, как массовая доля сухих веществ с помощью автоматического рефрактометра АЛР-3, pH напитка и массовая концентрация титруемых кислот с помощью гидроксида натрия. Полученные результаты представлены в таблице 2.

Таблица 2

Физико-химические показатели готовых напитков

Показатели	СВ, %	pH	Титруемая кислотность, г/100 $см^3$
Кофейный напиток, обогащенный инулином	0,8	5,6	0,8

В завершение была проведена органолептическая оценка, профилограмма которой изображена на рисунке 4. Напиток на основе зеленого кофе обладает травянисто-терпким ароматом и кисловатым вкусом. Благодаря добавлению корня лопуха, в напитке прослеживается травянисто-ореховый привкус.

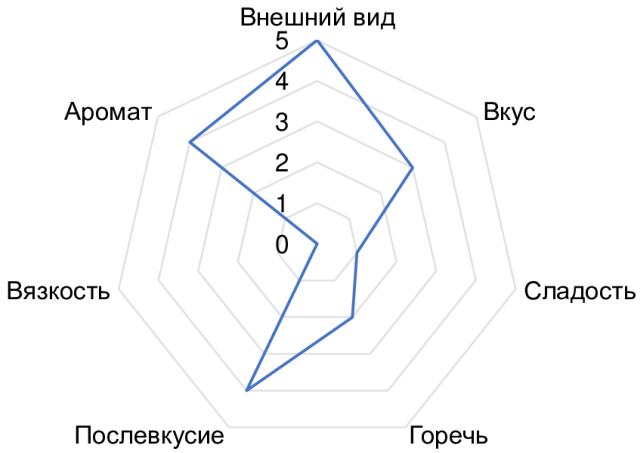


Рис. 4. Профилограмма напитка

Таким образом, был приготовлен напиток, обогащенный инулином. В готовом напитке были исследованы его физико-химические показатели и произведена его органолептическая оценка. В процессе выполнения экспериментальной части были установлены наилучшие параметры экстрагирования. В дальнейшем планируется улучшать такие органолептические показатели напитка как вкус, аромат и послевкусие, путем добавления имбиря.

Литература

1. Силаева М.А., Арбузова К.Р., Келешева В.Р., Соколова О.С., Шаненко Е.Ф., Ермолаева Г.А. Безалкогольный напиток на основе автолизата инулина топинамбура // Пиво и напитки. 2016. № 4. С. 16–18.

2. Сербаева Э.Р., Якупова А.Б., Магасумова Ю.Р., Фархутдинова К.А., Ахметова Г.Р., Кулуев Б.Р. Инулин: природные источники, особенности метаболизма в растениях и практическое применение. // Биомика. 2020. Т. 12. №1. С. 57–79.
3. Дьякова Н.А. Разработка и валидация экспрессной методики количественного определения водорастворимых полисахаридов в корнях лопуха обыкновенного (Arctium lappa L.) // Химикофармацевтический журнал. 2015. Т. 49. №9. С. 35–38.

УДК 577.164.16

СРАВНЕНИЕ РАСТИТЕЛЬНЫХ И БАКТЕРИАЛЬНЫХ ИСТОЧНИКОВ ВИТАМИНА B12

Жернякова А.В.[1]
Научный руководитель – к.т.н., доцент Бараненко Д.А.[1]
1 – Университет ИТМО

Работа выполнена в рамках темы НИР №620145 «Роль биологически активных веществ природного происхождения в развитии и нутритивной профилактике неинфекционных заболеваний».

Аннотация
В работе рассматриваются растительные и бактериальные источники витамина B12. Изучены параметры выращивания пропионовокислых бактерий (*Propionibacterium shermanii)* и последующей экстракции цианокобаламина. Рассмотрены параметры экстракции витамина из водоросли рода Arthrospira. Проведен сравнительный анализ выбранных источников витамина B12.
Ключевые слова
Цианокобаламин, растительное питание, витамин B12, пропионовокислые бактерии, спирулина.

Вегетарианство приобрело популярность в последние годы, поскольку люди стали больше осознавать влияние своего выбора продуктов питания на окружающую среду, благополучие животных и их здоровье. В то время как вегетарианские диеты могут обеспечить организм множеством питательных веществ, в них часто не хватает необходимого витамина витамина B12.

Витамин B12 (или цианокобаламин) нужен для образования эритроцитов, развития нейронов и синтеза ДНК. Его нехватка может приводить к накоплению гомоцистеина (нейротоксичного соединения), анемии, утомляемости и ухудшению памяти. Особенно актуальна проблема B12- дефицитной анемии у веганов. При питании исключительно растительными продуктами, восполнение кобаламина – обязательно. Включение в рацион пищевых добавок и обогащенных продуктов – неотъемлемая часть растительного питания.

Растительные источники витамина B12 не считаются надежным источником витамина, поскольку они могут быть не в активной форме или не могут эффективно усваиваться организмом. В этом контексте важно понимать различные растительные источники витамина B12 и их биодоступность, чтобы обеспечить адекватное потребление этого важного питательного вещества людьми, придерживающимися растительной диеты [1].

Целью работы было исследовать и сравнить растительные и бактериальные источники витамина B12. В качестве объектов исследования были выбраны такие бактериальные источники, как *Propionibacterium shermanii;* и растительные -водоросли *Arthrospira.*

Пропионовокислые бактерии (ПКБ) широко используются в пищевой и химической промышленности. Для культивирования ПКБ и извлечения цианокобаламина необходимо оптимизировать несколько параметров. Во-первых, pH питательной среды поддерживается на уровне 6,5-7,5, так как именно в этом диапазоне ПКБ растут оптимально. Во-вторых, температура поддерживается на уровне 30-37°C, чтобы способствовать эффективному росту. В-третьих, содержание источника углерода в среде, такой как глюкоза или лактат, и источник азота, такой как дрожжевой экстракт или пептон, для поддержки роста бактерий. Наконец, при экстракции цианокобаламина используют кислотное осаждение и хроматографические методы для очистки витамина.

Эти методы включают использование кислоты для высвобождения цианокобаламина из бактериальных клеток и последующую хроматографию для выделения и очистки витамина. В целом, оптимизация этих параметров имеет решающее значение для эффективного культивирования ПКБ и последующей экстракции цианокобаламина. В среднем, выделение цианокобаламина из пропионовокислых бактерий составляет 60–70 мкг/мл. Ниже приведена схема технологии получения витамина B12 [2] (рис. 1).

Рис. 1. Технологическая схема получения витамина B12 [2]

Спирулина – это сине-зеленая водоросль, которую используют как БАД из-за ее высокой питательной ценности и потенциальной пользы для здоровья. На базе лаборатории мы исследовали 3 наименования сухой, пресованной в таблетки, спирулины на наличие в ней витамина B12.

Использование экстрагированного B12 в пищевой продукции накладывает определенные ограничения на методы экстракции. Так, добавление цианида калия, бензилового спирта, хлороформа недопустимо в данном случае. Приемлемыми методами являются: высокотемпературная водная экстракция и высокотемпературная спиртовая экстракция [3].

Сублимированная биомасса обрабатывалась 80% этанолом. Суспензию кипятили в течение 20 мин при 80°C. Экстракт центрифугировали при 1000 g в течение 10 мин. Надосадочную жидкость выпаривали под вакуумом и использовали для анализа витамина B12. Результаты Хромато-масс-спектрометрии представлены на рисунках 2-4.

MS Quantitative Table

ID#	Name	Ret. Time	m/z	Conc.	Area	Unit
1	B12	1,312	678,00>147,00	0,00024	43	mkg/ml
Total					43	

Рис. 2. Содержание витамина B12 в Спирулине1

Масса навески образца – 2 г
Объем экстракта после окончания экстракции – 23 мл
Объем спирта, внесенного до начала экстракции – 40 мл
C(в12) = (0,00024мкг/мл*23мл)/2 г = 0,0028 мкг/г (на сухой вес спирулины).

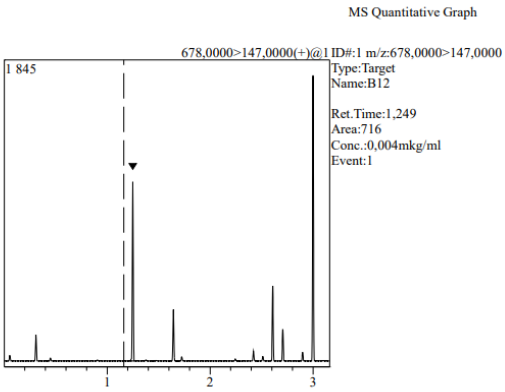


Рис. 3. Содержание витамина B12 в Спирулине2

Масса навески образца – 2 г
Объем экстракта после окончания экстракции – 30 мл
Объем спирта, внесенного до начала экстракции – 40 мл
C(в12) = (0,00396 мкг/мл*30 мл)/2г = 0,0594 мкг/г (на сухой вес спирулины).

MS Quantitative Table

ID#	Name	Ret. Time	m/z	Area	Unit	Conc.
1	B12	1,263	678,00>147,00	30	mkg/ml	0,00017
Total				30		

Рис. 4. Содержание витамина B12 в Спирулине3

Масса навески образца – 2 г
Объем экстракта после окончания экстракции - 32 мл
Объем спирта, внесенного до начала экстракции – 40 мл
C(в12) = (0,00017 мкг/мл*32 мл)/2 г = 0,0027 мкг/г (на сухой вес спирулины).

На рисунке 5 представлены для сравнения источники витамина B12. За контроль взят показатель говяжьей печени.

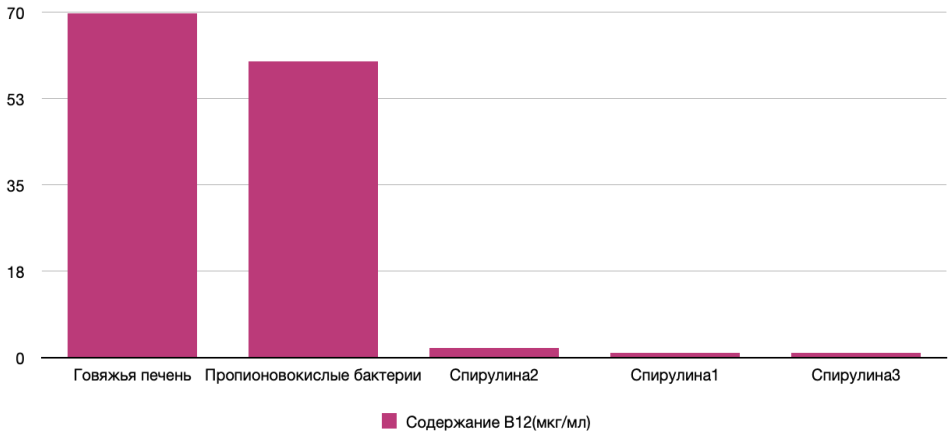


Рис. 5. Диаграмма. Содержание витамина B12 в разных источниках

Из показателей диаграммы можно видеть, что наиболее близким по показателям источником неживотного витамина B12 является пропионовокислые бактерии. Из чего следует, что бактериальные источники цианокобаламина являются перспективными для дальнейших исследований.

В данной работе были описаны параметры культивации пропионовокислых бактерий, а также методы экстракции выделяемого ими цианокобаламина. Дан сравнительный анализ различных источников витамина B12. Исследована возможность его экстракции и дальнейшее применение в пищевой промышленности.

Литература

1. Watanabe F., Yabuta Y., Bito T., Teng F. Vitamin B12-Containing Plant Food Sources for Vegetarians. Nutrients. 2014. 6(5):1861-1873.
2. Разговоров. П.Б. Технология получения биологически активных веществ: учеб. пособие / Разговоров П.Б. Иван. гос. хим.-технол. ун-т. Иваново. 2010. 72 с.
3. Andrade L. M. et al. Chlorella and spirulina microalgae as sources of functional foods //Nutraceuticals, and Food Supplements. 2018. Т. 6. №. 1. С. 45-58.

УДК 637.04

АНАЛИЗ ВОЗМОЖНОСТИ ПРОИЗВОДСТВА ГОРОХОВОГО ИЗОЛЯТА В РОССИИ

Иващенко Л.В.[1] (магистрант 1 курса)**, Амбарцумов Т.Г.**[1] (магистрант 1 курса)**,**
Захарова У.Е.[1] (магистрант 1 курса)
Научный руководитель – к.т.н., доцент Орлова О.Ю.[1]
Консультант – ассистент ФТМИ Бойцова Ю.С.[1]
1 – Университет ИТМО
e-mail: leviva0063@gmail.com

Работа выполнена в рамках темы НИР №622151 «Обеспечение коммерциализации научно-прикладных проектов».

Аннотация
В данной научной статье рассматривается анализ возможности производства горохового изолята в России. Гороховый изолят представляет собой высокобелковый продукт, получаемый из гороха, который становится все более популярным в качестве альтернативы животным белкам. Рост спроса на растительные белки обусловлен экологическими, экономическими и здравоохранительными факторами. Статья начинается с обзора текущего состояния рынка горохового изолята, а также основных производителей и потребителей данного продукта. Затем проводится анализ сырьевой базы для производства горохового изолята в России, включая оценку доступности и качества гороха, а также инфраструктуры для его переработки. Далее рассматриваются технологические аспекты производства горохового изолята, включая существующие и перспективные технологии, а также ограничения.
Ключевые слова
гороховый изолят, растительный белок, агропромышленный комплекс, технологии переработки.

Исследование возможностей получения альтернативных растительных белков является важным направлением в научных исследованиях, которое может иметь ряд положительных последствий для человечества и нашей планеты в целом. Существуют несколько подходов к решению проблему белкового голодания, но наиболее популярным и перспективным является выделения белковых компонентов из растительных продуктов. Крупные производители успешно внедряют линейки продуктов с повышенным содержанием белка, к примеру в состав протеиновых батончиков чаще всего входит именно гороховый изолят. Так как простота его получения и возможности применения неограничен, производители экспериментируют и добавляют изолят в печенье, питательные смесь [1]. Горох по своему аминокислотному составу близок к животным белкам, поэтому является вполне перспективным продуктом, а учитывая географическое особенности России, которые позволяют культивировать горох, можно говорить о высоком потенциале развития производств по получения гороховых изолятов. Вследствие этих факторов следует изучить рынок глубокой переработки гороха и выделить тренды глубокой переработки гороховых продуктов.

Проблема ограниченности пищевых ресурсов в последние годы встаёт более остро, человечество все чаще стало задумываться о том, чем будет питаться в будущем. Поэтому на данный момент ведутся активные работы по поиску и созданию альтернатив традиционным продуктам питания. Для обеспечения жизнедеятельности человека необходимо закрывать потребности в трёх базовых компонентов: белки, жиры, углеводы. Три этих компонента являются энергетический жизненно важными для организма человека. Поэтому учёные находится в поиске способов получения этих компонентов из не традиционных продуктов. Поэтому одной задачей является выделение этих компонентов из продуктов, для создания оптимальных рецептур, и концентрированных полуфабрикатов. Благодаря выделению конкретного компонента из продукта, к примеру белка, появляется возможность не зависеть от традиционных

источников сырья. Необходимо проанализировать потенциал рынка России с точки зрения возможностей получения горохового изолята и возможности его применения в локальных продуктах, а также экспорта.

Белки могут быть получены из разных продуктов, одним из перспективных вариантов получения белков является выделение из их из растительных продуктов питания. Есть множество продуктов богатых белком, но, когда речь идёт о массовости продукта, необходимо учитывать себестоимость получения и возможность масштабирования технологий, под эти задачи хорошо подходит горох. Этот продукт богат белками, имеет взвешенный аминокислотный состав, что позволяет говорить о возможности использования гороха в питании человека. Согласно данным, экспертно-аналитический центр агробизнеса, объемы культивируемого в России гороха увеличиваются, но не все сорта гороха целесообразно подвергать глубокой переработке с целью выделения белка [2]. Большинство западных производителей экстрагируют белок из желтого сорта гороха. Этот сорт хорошо подходит для растительной еды и спортивного питания. Желтый горох содержит более высокое количество белка по сравнению с другими сортами, такими как зеленый или черный горох. Кроме того, желтый горох имеет более мягкую структуру и лучше растворяется в воде, что делает его более подходящим для получения изолята.

Гороховый изолят – это высококачественный источник белка, который может быть использован в различных областях питания людей и животных. В области питания людей гороховый изолят может быть использован в качестве основного источника белка или как дополнительный компонент в различных продуктах. Он может быть добавлен в белковые коктейли, напитки, смузи, мюсли, йогурты, сыры, мясные и растительные блюда, хлеб, кондитерские изделия и другие продукты питания [3]. Гороховый изолят также может быть использован в качестве заменителя мяса в вегетарианских и веганских блюдах, что делает его отличным выбором для тех, кто следит за своей диетой. Кроме того, гороховый изолят может быть использован в пищевой промышленности в качестве функционального ингредиента для улучшения текстуры, вязкости, консистенции и стабильности продуктов.

Гороховый изолят может быть использован для производства различных видов кисломолочной продукции, таких как йогурт, кефир, творог и другие, производство которых является одной из самых развитых отраслей пищевой промышленности России. Экстрагируемый гороховый белок содержит высокую концентрацию белка, который может быть использован для увеличения питательной ценности кисломолочной продукции [4]. Кроме того, гороховый изолят не содержит лактозу, что делает его подходящим для производства кисломолочных продуктов для людей, страдающих от непереносимости лактозы. Применение изолята для производства кисломолочной продукции может улучшить вкусовые и текстурные характеристики продукта. Получаемый в результате "порошок", может быть использован в качестве заменителя молочного белка, что позволяет уменьшить содержание жира и калорийности кисломолочной продукции, что особенно важно для людей, следящих за своим питанием [5]. Кроме того, он может быть использован для создания новых видов кисломолочной продукции, таких как напитки на основе йогурта или кефира [6]. Эти продукты могут быть обогащены витаминами и минералами, что делает их более полезными для здоровья.

В целом, использование горохового изолята для производства кисломолочной продукции имеет множество преимуществ, таких как улучшение питательной ценности, вкусовых и текстурных характеристик продукта, а также возможность создания новых видов продукции. Это делает гороховый изолят перспективным продуктом для производства кисломолочной продукции на российском рынке [7].

Глубокая переработка гороха в России находится на начальной стадии развития. Однако, уже сейчас в России есть компании, которые занимаются переработкой гороха

в продукты высокой степени очистки, к примеру компания "Уралхим" запустила в тестовом режиме установку по переработке гороха в изолят [8, 9]. В качестве одного из главных продуктов переработки гороха выделяются растительные белки, которые могут быть использованы для производства пищевых продуктов и добавок, таких как мясные заменители, сухие завтраки, батончики, каши и другие. Кроме того, горох может быть использован для производства крахмала, белково-витаминных комплексов, диетических и функциональных продуктов. Одной из самых популярных технологий при производстве растительных белков из гороха является изоляция белка методом ферментативного гидролиза. Этот процесс заключается в обработке гороха ферментами, которые разрушают клеточные стенки гороха и выделяют белковую фракцию, которая затем очищается и сушится. Важно отметить, что переработка гороха является трудоемким и затратным процессом, что может сказаться на стоимости произведенных продуктов. Однако, увеличение интереса к растительным продуктам и более сбалансированным диетам, а также перспективы развития сектора растительных продуктов в России, могут стимулировать развитие глубокой переработки гороха в будущем.

Гороховый изолят – это высококачественный источник белка, который может быть использован в различных областях питания людей и животных. Его использование позволяет улучшить качество пищевых продуктов и кормов, а также создать новые продукты с повышенной питательной ценностью. Кроме того, гороховый изолят является экологически чистым продуктом, который может быть использован в качестве альтернативы другим источникам белка, что делает его полезным и перспективным продуктом на российском рынке.

Литература

1. Астанина В.Ю. Применение белковых препаратов чечевицы в технологии мясных продуктов: Автореф. дис. канд. техн. наук/ Астанина В.Ю. Воронеж. 1998. 22 с.
2. Всемирный день зернобобовых: российские предприниматели берут курс на горох.: Натур продукт [Электронный ресурс]. 2021. URL: https://np-mag.ru/zdorovaya-eda/poleznye-produkty/rossijskie-predprinimateli-berut-kurs-na-goroh/ (дата обращения 25.12.2022).
3. Кудряшов В.Л. Импортозамещающая технология производства пищевых концентрата и изолята белка, крахмала и клетчатки из гороха с применением баромембранных процессов // Научно- техническая информация. 2019. N 1(39). С. 46-50.
4. Растительные альтернативы мясных продуктов: тренды и перспективы.: Мясной эксперт [Электронный ресурс]. 2021. URL: https://meat-expert.ru/articles/515-rastitelnye-alternativy-myasnykh-produktov-trendy-i-perspektivy / (дата обращения: 25.12.2022).
5. Калорийность Горох колотый, вареный, без соли. Химический состав и пищевая ценность. Sstale [Электронный ресурс]. 2022. URL: https://stale.ru/boletus/kaloriinost-goroh-kolotyi-varenyi-bez-soli-himicheskii-sostav-i-pishchevaya/ (дата обращения: 25.12.2022).
6. Greenwise - драйверы Российского фудтеха: Greenwise [Электронный ресурс]. 2022. URL: https://greenwise.ru/about (дата обращения 25.12.2022).
7. О компании Tashir Food: Не мясо [Электронный ресурс]. 2022. URL: https://notmeat.ru (дата обращения: 25.12.2022).
8. Biofoodlab о компании: Biofoodlab [Электронный ресурс]. 2022. URL: https://biofoodlab.com/blog/ (дата обращения 25.12.2022).
9. «Уралхим» посмотрит на горох по-новому: Уралхим мясо [Электронный ресурс]. 2022. URL: https://www.uralchem.ru/press/news/item25318// (дата обращения 25.12.2022).

УДК 664.69

РАЗРАБОТКА РЕЦЕПТУР БЕЗГЛЮТЕНОВЫХ МАКАРОННЫХ ИЗДЕЛИЙ

Йимер Г.А.[1]
Научный руководитель – к.т.н., доцент Баракова Н.В.[1]
1 – Университет ИТМО

Аннотация
Изучены свойства макаронного теста, приготовленного из 40 % кукурузной муки, 15 % нутовой муки и 15 % нативного и модифицированного кукурузного крахмала E1422 GranStarch MC. Было установлено, что применение модифицированного крахмала в качестве структурообразователя дало положительные результаты - тесто приобрело эластичность, стало более растяжимым.
Ключевые слова
Безглютеновые макаронные изделия**,** кукурузная мука, нутовая муки, модифицированный кукурузный крахмал

Введение. Макаронные изделия относятся к продуктам массового потребления и производятся из пшеницы твердых сортов. В состав пшеничной муки входит белок-глютен, который и создает структуру теста. Мука, используемая для приготовления безглютеновых продуктов, не содержит глютен. Обычно при разработке безглютеновых макаронных изделий используют кукурузную, льняную, рисовую, нутовую муку [1]. Кукурузная мука широко используется для приготовления безглютеновых продуктов. По содержанию ненасыщенных жирных кислот, пищевых волокон, витаминов и минералов наиболее питательной и полезной является льняная мука. Рисовая мука является источником крахмала, нутовая мука – источником белка. Нутовая мука может быть рекомендована для получения макаронных изделий с высокой пищевой ценностью, с повышенными физико-химическими, текстурными и органолептическими показателями [2].

Высокий потребительский спрос на макаронные изделия, относящиеся к здоровому питанию, повлиял на разработку макаронных изделий, содержащих нетрадиционные ингредиенты – пищевые волокна, минералы, витамины, натуральные пигменты и антиоксиданты [3, 4]. Рекомендуется также макаронные изделия обогащать непатогенными микроорганизмами. Добавление непатогенных микроорганизмов, например, спирулинов, в тесто для макаронных изделий придает готовому изделию цвет, аромат, а также повышает его пищевую ценность, поскольку микроорганизмы богаты микро- и макроэлементами [5].

Сформировать макаронное изделия из муки, не содержащей глютен, а содержащий только крахмал, невозможно, так как крахмал не обладает вязко-тегучими свойствами. Необходимо разрабатывать либо специальные технологические приемы, либо включать в рецептуры макаронных изделий, для создания каркаса теста, компоненты, которые являются структурообразователями.

К технологическим приемам, способствующими формированию каркаса для получения эластичного теста, относится температура замеса теста. Существуют разные виды замеса макаронного теста: холодный - при температуре воды 25 ℃, теплый замес - при температуре 55 ℃ и горячий замес – при температуре 90 ℃. При горячем замесе допускается использовать воду по температуре близкой к закипанию. При внесении горячей воды в муку крахмал клейстеризуется и образует в тесте гомогенную однородную гелеобразную связующую массу. При горячем замесе структура теста становится более связанной и пластичной. Недостатком такого приемы является ухудшение качества макаронных изделий.

Низкотемпературная заморозка (-150 °C) – отличный способ длительного хранения макаронных изделий без глютена. Более низкая температура замораживания снижает твердость, способность к перевариванию, когезивность и клейкость безглютеновых

макаронных изделий. Низкотемпературное замораживание может уменьшить при длительном хранении рост микроорганизмов - дрожжей и грибков [6].

Раскатывание – это еще один распространенный технологический прием для формирования эластичного теста.

К современному способу, обеспечивающему необходимую структуру теста, относится включение в рецептуру макаронных изделий различных структурообразователей – камедей, набухающих, окисленных или сшитых крахмалов [7].

Крахмалы с низким содержанием амилозы не дают желательных результатов, но крахмалы с высоким содержанием амилозы, не дают приемлемой степени желатинизации в ходе термического процесса, но обеспечивают дополнительную ретроградацию крахмала. Крахмалы, богатые амилозой, интересны с точки зрения здорового питания, поскольку они являются хорошим источником резистентного крахмала – питательной среды для микробиома человека [8].

При разработке рецептуры безглютеновых макарон необходимо учитывать, что влагосвязывающие свойства различных видов муки отличаются от водопоглотительных свойств пшеничной муки, поэтому необходимо проводить корректировку по количеству вносимой воды. Оценку качества теста для макаронных изделий проводят по следующим показателям: влажности (%), коэффициенту увеличения массы, по содержанию сухих веществ в варочной воде (%), сохранность формы (%), органолептическим показателям-вкус, цвет, аромат, потери сухих веществ при варке (%), увеличение веса и объема (%).

В текущем эксперименте использовали нативный кукурузный крахмал и термостабильный, химически модифицированный крахмал E1422 GranStarch MC из восковидной кукурузы. E1422 GranStarch MC используется в качестве загустителя, наполнителя, стабилизатора и эмульгатора. Модифицированный крахмал помогает регулировать консистенцию продуктов, используется как влагоудерживающий, желирующий и глазирующий агент. E1422 GranStarch MC – это крахмал, обработанный ангидридами уксусной кислоты и адипиновой кислотой. E1422 устойчив к ретроградации и разрезанию, механическому воздействию, оттаиванию-замораживанию и низкому pH. Кроме того, крахмал E1422 GranStarch MC термоустойчив и не теряет своих качеств при длительном хранении. Объектами исследования были кукурузная мука производства ООО «Гарнец», Россия, нутовая мука без глютена производства ООО «Гарнец», Россия, нативный кукурузный крахмал производитель ООО «ЧКЗ», Россия и термостабильный, химически модифицированный крахмал E1422 GranStarch MC, Китай.

Была изучена эластичность и растяжимость теста. изменение объема макаронных изделий при варке.

Тесто готовили с использованием 40 % кукурузной муки, 15 % нутовой муки и 15 % нативного кукурузного крахмала и термостабильного, химически модифицированного кукурузного крахмала. Замешивание теста проводили при комнатной температуре.

При замесе теста с использованием нативного крахмала тесто отличалось «крахмалистостью», крошилось, не получалось сформировать эластичный мякиш.

При замесе теста с добавлением модифицированного крахмала тесто получалось эластичным, не прилипало к столу, хорошо раскатывалось.

Приготовленное тесто варили 20 минут при 100 °C. Объем макаронных изделий до и после варки рассчитывали по формуле объема параллелепипеда. Объем исходного образца составлял 7 см3. Удельный объем макаронных изделий в процессе варки увеличивался до 7; 9,12; 10,49 и 10,76 см3 через 5, 10, 15 и 20 мин варки изделий соответственно.

Результат эксперимента показал перспективность применения модифицированного крахмала E1422 GranStarch MC при производстве безглютеновых макаронных изделий.

Литература

1. Никитин И.А. Возможности применения льняной муки и муки сорго для разработки рецептур безглютеновых макаронных изделий / Никитин И.А., Клоконос М.В., Муталлибзода Ш., Зейбд Д.А. //Современные аспекты производства и переработки сельскохозяйственной продукции: сборник трудов VI Международной научно-практической конференции. Краснодар. 2020. С. 313-318.
2. El-Sohaimy S.A., Brennan M., Darwish A.M. G. and Brennan C. Physicochemical, texture and sensorial evaluation of pasta enriched with chickpea flour and protein isolate//Ann. Agric. Sci. 2020.vol. 65. no. 1. pp. 28–34.
3. Cubadda R.E., Carcea M., Marconi E., and Trivisonno M.C. Influence of Gluten Proteins and Drying Temperature on the Cooking Quality of Durum Wheat Pasta// Cereal Chem. J. 2007.vol. 84. no. 1. pp. 48–55.
4. Nilusha R.A.T., Jayasinghe J.M.J.K., Perera O.D.A.N. and Perera P.I.P. Development of Pasta Products with Nonconventional Ingredients and Their Effect on Selected Quality Characteristics: A Brief Overview// Int. J. Food Sci. vol. 2019. pp. 1–10. Nov. 2019.
5. Özyurt G. et al./Evaluation of the Cooking Quality Characteristics of Pasta Enriched with Spirulina Platensis: Quality Characteristics of Pasta Enriched with Spirulina// J. Food Qual. 2015. vol. 38. no. 4. pp. 268–272.
6. Vieira A. et al./Effects of long-term frozen storage on the quality and acceptance of gluten-free cassava pasta//Heliyon, 2021. vol. 7. no. 8. p. e07844.
7. Осипова Г.А. Использование продуктов переработки овса в производстве макаронных изделий /Г.А.Осипова// Научные основы пищевых технологий. 2011. Том 7. №2. с. 23-29.
8. Bresciani A., Giordano D., Vanara F., Blandino M., and Marti A. /High-amylose corn in gluten-free pasta: Strategies to deliver nutritional benefits ensuring the overall quality// Food Chem. 2021.vol. 353. p. 129489.

УДК 691.175.664, 621.763, 54.03

ЗЕЛЁНАЯ ХИМИЯ КАК ОСНОВА ДЛЯ ПОЛУЧЕНИЯ ПОЛИУРЕТАНО-КОМПОЗИТНОГО МАТЕРИАЛА ИЗ БИО-СЫРЬЯ

Коновалов Д.С.[1] (аспирант**)**
Научный руководитель – д.х.н., доцент Зуев В.В.[1]
1 – Университет ИТМО

Аннотация
В работе представлены синтезы для получения полиуретана на основе из касторового масла с разным процентным добавлением микроопилок из березы для улучшения параметров: относительная прочность и удлинение на разрыв. В ходе полученных данных была установлена зависимость влияния наполнителя на физико-механические показатели получаемого материала, а также решена проблема высокого содержания гидроксильных групп в касторовом масле.
Ключевые слова
Полиуретан, композитные материалы, микроопилки, зеленая химия, возобновляемые материалы.

Актуальность

В последние десятилетия во всем мире отмечается всплеск спроса на применение материалов на основе биологических, воссоздаваемых ресурсов. Одновременно предпринимаются усилия по сокращению потребления нефти, это связано со снижением парниковых газов за счет уменьшения потребления нефтепродуктов. Использование возобновляемого сырья, является важным вкладом в устойчивое развитие производства полимеров. Эта стратегия основана на преимуществах природного потенциала синтеза и принципах Зеленой химии. Материалы, полученные из возобновляемого сырья, вызывают определенный интерес. Разработка полимеров, синтезированных из сельскохозяйственных продуктов, таких как крахмал, целлюлоза, сахара или лигнин, значительно увеличилась за последние два десятилетия [1]. Среди всех возможных природных источников полимеров растительные масла считаются одними из самых дешевых и распространенных в природе [2, 3]. Их можно использовать в качестве выгодной химической платформы для синтеза полимеров благодаря присущей им биоразлагаемости и низкой токсичности для человека и окружающей среды. В настоящее время предпринимаются многочисленные попытки предложить большое разнообразие химических методов получения материалов на основе растительных масел [4]. Широкий спектр химических методов, применимых к природным материалам, приводит к появлению множества различных мономеров и полимеров с множеством применений.

Полимеры, получаемые на основе растительного сырья, имеют возможность вторично перерабатываться. Материалы полученные на основе растительного, а точнее, возобновляемого сырья имеют более высокую экономическую выгоду. Но на практике, данные материалы имеют низкие эксплуатационные характеристики. Получаемые свойства из полиизоционатов на основе касторового масла являются недостаточными для массового внедрения, поэтому в качестве повышения механических и эксплуатационных свойств предлагается использовать березовые микроопилки.

Целью проекта является разработка методики и получение полиуретано-композитного покрытия с применением основ зеленой химии с содержанием био-сырья, как основа экологического стандарта для применения в условиях Санкт-Петербурга.

Задачи проекта:

1. Получение полиуретано-композитного материала из био-сырья.
2. Влияние микроопилок березы в качестве наполнителя на физико-механические и эксплуатационные параметры получаемого материала.

3. Стабилизация получаемого покрытия.
4. Разработка методики получения однородно-распределенного наполнителя по всей поверхности покрытия.

Основная часть

Был произведен синтез полиуретано-композитного покрытия на основе касторового масла и Puranole PM-200 в соотношении 1:2,74 по массе. К данной смеси вводилось 5,10,20,30% микроопилок предварительно вымаченных в касторовом масле. По окончанию полимеризации пленок было выявлено высокое содержание гидроксильных групп, что привело к реакции с большим выделением CO2, а также установлено плохое смесеобразование опилок, выраженное в виде комкования на поверхности пленки (рис. 1).

Рис. 1. Полиуретановые пленки с содержанием микроопилок

Данные факторы очень существенно повлияли на физико-механические показатели, что видно из рисунка 2, чем больше содержание микроопилок в составе, тем выше прочность по началу, но затем все показатели стремительно падают, за счет большего вспенивания. Связано это с содержанием гидроксильных групп в касторовом масле и опилках.

Содержание наполнителя, %	Прочность на разрыв, МПа	Относительное удлинение, %
0	0,8	35
5	3	60
10	2,7	46
20	3	17
30	1,4	19

Рис. 2. Физико-механические показатели покрытия с применением наполнителя от 5 до 30%, с вспениванием

После анализа полученных данных были применены следующие действия и постановка задач:

1. Получение однородной пленки без комкования микроопилок.
2. Равномерное распределение опилок по всей поверхности получаемой пленки.
3. Уменьшить содержание гидроксильных групп предварительно прокалив опилки при температуре 100С в течении 24ч.

Для решения первой задачи было решено воспользоваться доступным оборудованием – дисольвером и насадку в виде фрезы для того, чтоб наполнить касторове масло максимально возможным содержанием березовых опилок, предварительно подготовленных и высушенных, в ходе чего выяснить максимальное количество наполнителя (2:1). Далее получаемую массу для получения однородного наполнителя пропустить через трехваловую краскотерную машину (рис. 3).

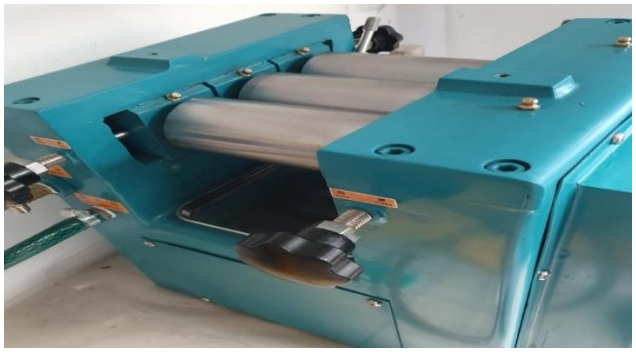


Рис. 3. Трехваловая краскотерная машина

Как результат, после данных механических воздействий данный наполнитель в ходе введения его в основную часть хорошо размешивался и распределялся по всей поверхности пленки. Были проверены физико-механические характеристики получаемого покрытия, в ходе которого было установлено, что содержание гидроксильных групп уменьшилось, но вспенивание осталось что видно на рисунке 4. Однако зависимость, получаемая при введении микроопилок, стала более логичной (рис. 5). Таким образом, было получено, что максимальная прочность достигается при введении 20% микроопилок, но также под микроскопом было видно большое содержание свободных ячеек CO2, что существенно повлияло на показатели физико-механики.

Рис. 4. Пленки с содержанием пред подготовленных микроопилок от 5 до 30%

Содержание наполнителя, %	Прочность на разрыв, МПа	Относительное удлинение, %
0	0,8	35,1
5%	1,6	35,1
10%	1,9	30,2
20%	3,7	27,3
30%	3,5	18,2

Рис. 5. Физико-механические показатели покрытия с пред подготовленным наполнителем от 5 до 30%, с вспениванием

Для решения задачи уменьшения содержания гидроксильных групп в касторовом масле были применены следующие решения:

1) введение полиметилсилаксан 200 для снижения поверхностного натяжения получаемой пленки для более простого выхода CO2;

2) провести дегидратацию касторового масла с применением гексометилдисилазан при температуре 150;

3) применить другой изоцианат PM700 или смесь с PM200.

4) ввести молекулярные сита;

5) уменьшить время жизни введением в касторовое масло катализатора.

Были получены следующие результаты, которые представлены на рисунках 6, 7.

Рис. 6. Полиуретановые пленки на основе касторового масла

Содержание наполнителя, %	Прочность на разрыв, МПа	Относительное удлинение, %
Касторовое масло	0,8	35,1
Смесь PM200 и PM700	0,5	41,8
Молекулярные сита	1,8	45
Катализатор с ПМС-200	2,8	59,7
Дегидратация	1,4	56
Молекулярные сита, ПМС-200, катализатор	2,5	53,77

Рис. 7. Физико-механические параметры полиуретановых пленок с применением разных способов уменьшения гидроксильных групп в материале

Таким образом, были получены полиуретано-композитные материалы из био-сырья, установлена зависимость влияния микроопилок березы в качестве наполнителя на физико-механические и эксплуатационные параметры получаемого материала. Разработана методика получения однородно-распределенного наполнителя по всей поверхности покрытия.

Литература

1. Lligadas G., Ronda J.C., Galia M., Cadiz V. Plant Oils as Platform Chemicals for Polyurethane Synthesis: Current-State-of-the-Art. Biomacromolecules. 2010 Nov 8/ 11(11). Pp. 2825-35.
2. doi: 10.1021/bm100839x.Biermann U., Friedt W., Lang S., Luhs W., Machmuller G., Metzger J.O., Klaas M.R.G., Schafer H.J., Scheider M.P. New Syntheses with Oils and Fats as Renewable Raw Materials for the Chemical Industry. Angewandte Chemie International Edition, 2000. 39. Pp. 2206-2224.

3. Montero L. de Espinosa, Meier M.A.R. Plant oils: the perfect renewable resource for polymer science?! European Polymer Journal. 2011. Volume 47, Issue 5, May 2011. Pp. 837-852.
4. Behr A., Gomes J.P. The refinement of renewable resources: new important derivatives of fatty acids and glycerol. European Journal of Lipid Science and Technology. 2010. 112 (1): Pp. 31 -50.

УДК 614.31:637

ПРИМЕНЕНИЕ МЕТОДА ХРОМАТО МАСС- СПЕКТРОМЕТРИИ ДЛЯ ОЦЕНКИ БЕЗОПАСНОСТИ РЫБЫ И РЫБНЫХ ПОЛУФАБРИКАТОВ

Наумова А.В.[1], Утсаль В.А.[1], Яккола А.Н.[1], Наумов И.А.[1], Островидова Е.В.[1]
Научный руководитель – Ишевский А.Л.[1]
1 – Университет ИТМО

Аннотация
Рыба и продукты ее переработки относятся к числу скоропортящихся продуктов. С увеличением сроков и температуры хранения происходят процессы окисления, и катализируются процессы разложения белка, которые проходят с участием тканевых и бактериальных ферментов. В настоящей работе приведен разбор методов наиболее распространенных для контроля качества и безопасности рыбы и рыбных полуфабрикатов.
Ключевые слова
Рыба, липиды, окисление, контроль качества, безопасность.

Одна из наиболее актуальных проблем продовольственного рынка – обеспечение качества, безопасности и повышения сроков хранения пищевых продуктов. Основная причина ухудшения качества и снижения сроков хранения продовольственных товаров – процессы, которые приводят к ухудшению качества и утрате безопасности, происходящие непосредственно в самом товаре. В процессе холодильного хранения в мороженой рыбе протекают биохимические процессы, которые приводят к окислению жира и гидролизу белков.

Для каждого вида рыбы происходят характерные изменения, например, для мороженой рыбы с высоким содержанием жира наиболее характерны изменения в результате окислительных процессов в жире, а для тощей рыбы – денатурационные изменения в белковой системе сырья. Скорость, с которой происходят эти изменения в большей степени зависит от химического состава, свойств мороженой рыбы, температурного режима и времени хранения. Окисление жира в рыбе происходит даже при очень низких температурах, даже ниже -30 С°. Одновременно с органолептическими изменениями в рыбе накапливаются продукты окисления жира, подобные метаболиты придают рыбе неприятный запах и прогорклый вкус. В результате окисления высших жирных кислот образуются разнообразные промежуточные продукты – карбоксильные соединения, например оксикислоты и соединения содержащие кислород, первоначальными продуктами окисления – перекисные соединения. Все процессы окисления осуществляются за счёт молекулярного кислорода. Кинетика реакций зависит от многих факторов, прежде всего от ненасыщенности жирных кислот. Присутствие в жирах и маслах свободных жирных кислот, особенно ненасыщенных, -моно и диглицеридов в значительной степени ускоряет процесс окисления. Жиры и жирные кислоты подвержены изменениям под влиянием микроорганизмов, а именно аэробных и анаэробных бактерий и грибов. На первой стадии расщепления жира происходит гидролиз, который осуществляется за счёт фермента липазы. Порча жиров охлажденных и замороженных продуктов, происходит в большей степени при участии психрофилов [1]. Следует учитывать, что многие жирорасщепляющие микроорганизмы являются психрофилами, поэтому способны развиваться при хранении пищевых продуктов при минусовых температурах. Сроки хранения мороженой рыбы зависят от ее вида и химического состава, исходного состояния, вида разделки, способа и режимов замораживания рода упаковки, температуры хранения и других факторов [1, 2].

Материалы и методы

В настоящей работе проведен сравнительный анализ методов исследования окислительной порчи на примере метода хромато-масспектрометрии. Исследования метаболитов идентификаторов порчи рыбы и рыбных полуфабрикатов. Исследование проведено на двух видах рыб хек мороженный без головы ТУ 9261-002-51561792-2005 и сельдь с головой мороженная ТУ 10.20.13-002-51561792-2018.

В качестве модели исследования был выбран фарш [3, 4]. Исследование было проведено методом газо-жидкостной хроматомасс-спектрометрии. Однородную массу, не более 1 г, шпателем переносили в чистый флакон, предварительно взвесив флакон без содержимого. Добавляем 2 мл ацетанитрила. Затем содержимое флакона перемешиваем на Vortexe, переносим жидкую фазу в чистый флакон. Экстракцию повторяем дважды, экстракты объединяем. Далее объединенный экстракт сушим безводным сульфатом натрия, переносим в чистую виалу, после чего полностью удаляем ацетнитрил под тягой. Затем добавляем тридейтерометиловый эфир тридекановой кислоты в количестве 10 мкл С = 2 мг/мл, сушим нагреваем при температуре 30С в течении 20 секунд, добавляем 50 мкл селилирующего агента N,O-бистриметилселил-трифторацетамид (BSTFA), нагреваем 2 минуты при 50 С°, после охлаждения разбавляем 100 мкл хлористого метилена. Вводили пробу в количестве 1 мкл [2, 3]. Эксперимент проводился ежедневно в течение 9 дней.

Выводы

В настоящее время активно ведется поиск новых методов пролонгирования сроков хранения рыбы и рыбных полуфабрикатов. Для того чтобы подобрать правильный состав биоразлагаемой полимерной пленки нужно точно установить временные рамки и метаболиты, которые определяют сроки и соответственно сам вид порчи продукта. В настоящей статье рассматривается исследование сроков хранения рыбы и рыбных полуфабрикатов методом микробной хроматомасс-спектрометрии. Анализ показал, что динамика порчи рыбных образцов одинакова. Найден ряд соединений, который образуется при порче рыбы и рыбных полуфабрикатов и предположительно являются маркерами процесса, что поможет при дальнейшем выборе композиционных составов безопасных биоразлагаемых полимерных пленок.

Литература

1. Du W.-X., Huang T., Kim J., Marshall M.R., Wei C. Chemical, microbiological, and aromascan evaluation of mahi–mahi fillets under various storage conditions // Agric J. Food Chem. 2001. V. 49. P. 527–534.
2. Sanceda N.G., Suzuki E., Ohashi M., Kurata T. Histamine behavior during the fermentation process in the manufacture of fish sauce // Agric J. Food Chem. 1999. V. 47. P. 3596–3600.
3. ГОСТ Р 8.795-2012 Методики идентификации химических веществ методом хромато-массспектрометрии.
4. ГОСТ ISO 3960-2020. Жиры и масла животные и растительные. Определение перекисного числа. Йодометрическое (визуальное)определение по конечной точке.

УДК 637.146

ОБОГАЩЕНИЕ ДЕСЕРТА НА ОСНОВЕ ВТОРИЧНОГО МОЛОЧНОГО СЫРЬЯ ПРОБИОТИЧЕСКИМ ШТАММОМ

Сергуненкова О.Б.[1] (студент)
Научный руководитель – к.т.н., доцент Еремеева Н.Б.[1]
1 – Университет ИТМО
soba.203@ya.ru

Аннотация
Проблема рационального и полного использования молочной сыворотки существует во всех странах. Решением является перерабатывать ее в качестве вторичного сырья. Идея комплексного подхода к переработке реализуется посредством ферментации ее пробиотическим штаммом. Так. повышается ее пищевая ценность, улучшаются органолептические и реологические характеристики. В результате изменение свойств позволяет использовать ее для разработки пробиотических десертов с целью расширения ассортимента продуктов из молочной сыворотки.
Ключевые слова
Десерт, молочный пудинг, пробиотик, рецептура, молочная сыворотка, вторичное молочное сырье.

В настоящее время каждому известно, что питание имеет сильное влияние на здоровье человека. Наряду с экологией, генетическими особенностями, общей гигиеной и поведением, каждый день человек сам создает себя сегодняшнего и будущего за счет выбора пищи. Нарушения питания могут быть важнейшими факторами в развитии серьезных заболеваний, таких как ожирение, сахарный диабет, онкологические заболевания и т. п. Эти заболевания приводят к сокращению продолжительности жизни, ухудшению работоспособности. В случае питания детей оно в главной степени решает будущую судьбу ребенка. Потребление необходимых макро- и микронутриентов в достаточном количестве будет фактором сохранения и улучшения здоровья, увеличения продолжительности активного периода жизни.

Молочные продукты всегда были неотъемлемой частью рациона человека. Они богаты витаминами, минеральными веществами, а также белками, в составе которых присутствуют незаменимые аминокислоты. Существуют совершенно различные продукты на молочной основе, удовлетворяющие разные вкусы потребителей всех возрастов. Так. молочная промышленность занимает устойчивое место на рынке, являясь одной из самых развитых в пищевой сфере. Развитое производство принимает в оборот большое количество сырья.

К сожалению, молочное производство является не безотходным. После выработки сыра и творога остается побочный продукт – молочная сыворотка. Несмотря на высокую биологическую ценность, проблема ее переработки существует до сих пор на заводах самых развитых стран. Ее решение является приоритетным направлением пищевой промышленности, к тому же оно позволит расширить ассортимент молочной продукции. Состав молочной сыворотки разнообразен, к биологически активным веществам продукта можно отнести минеральные вещества, витамины, органические кислоты, аминокислоты, углеводы, ферменты. Одним из способов переработки вторичного сырья является создание десертов [1].

Десерты молочной промышленности набирают популярность последние десятилетия. Их разнообразие по вкусу, консистенции, функциональным добавкам и применению подтверждают это. Наиболее популярные позиции молочных десертов: йогурт, кисель, мусс, крем, желе, коктейль, пудинг [2].

Целью данной работы является разработка рецептуры и технологии приготовления обогащенного пробиотическим штаммом пудинга на основе вторичного молочного

сырья. Одной из ведущих задач является получение высоких органолептических показателей у продукта. Остальные нацелены на получение стабильной структуры, сохраняющейся в течение долгого срока в условиях хранения при температуре (4 ± 2) °С.

Пудинг – знаменитый десерт родом из Великобритании. Молочный пудинг – ответвление от традиционного десерта, изготавливается на основе молока или молочного сырья. Он имеет желеобразную структуру, но, в отличие от желе, проявляет слабые свойства текучести. Его текстура нежная, однородная.

Пробиотики – микроорганизмы, способствующие биологической активности микрофлоры кишечника человека. Продукты, обогащенные пробиотиками, обладают лечебно-профилактическими свойствами. Их использование при заквашивании продукта повышает его пищевую, биологическую ценность, улучшает органолептические и реологические свойства. Существуют пробиотические штаммы молочнокислых бактерий, которые способны выделять в окружающую среду экзополисахариды, которые могут быть использованы как в качестве натуральных биозагустителей, улучшающих реологические характеристики кисломолочных продуктов, так и выступать в роли факторов адгезии полезных микроорганизмов на стенках кишечника [3].

В результате изучения литературы была найдена статья [4], в которой предложена рецептура и технология приготовления пудинга на основе молочной сыворотки (табл. 1).

Таблица 1

Рецептура пудинга на основе молочной сыворотки на 1000 кг продукта, предложенная в статье [4]

Компонент	Количество компонента, кг
Сыворотка молочная	805
Сахар	100
Желток	20
Ванилин	05
Крахмал кукурузный	35
Стабилизатор Grindsted SB 550 A	35
Итого	1000

Технология производства пудинга на основе молочной сыворотки, предложенная в статье [4] состоит из следующих этапов: приемка и оценка качества сырья, пастеризация сыворотки с сахаром и ванилью, растворение основных ингредиентов в небольшом количестве сыворотки, внесение в горячую сыворотку растворенной смеси, охлаждение смеси, розлив в тару, охлаждение, упаковка, маркировка, оценка качества готового продукта, хранение. При разработке рецептуры и технологии производства пудинга в данной работе за основу брали именно эти рецептуру и технологию изготовления молочного пудинга.

В работе использовалось следующее сырье: сыворотка сухая, сахар-песок, крахмал кукурузный Gletogel B, ванильный сахар производителя «Dr. Oetker», стабилизатор Palsgaard AcidMilk306 (5846). При составлении рецептуры уделялось большое внимание получаемой консистенции продукта. Не слишком плотной, нежной и тающей во рту.

Свежая сыворотка может храниться малое количество времени и должна быть сразу переработана. Одним из способов сохранения сыворотки является сушка. Отсутствие влаги в продукте исключает развитие патогенных микроорганизмов. Выбор пал на деминерализованную сухую сыворотку с уровнем деминерализации 70%. Деминерализация – процесс частичного удаления минеральных веществ. Происходит снижение концентрации солей, органических кислот, азотистых соединений в продукте,

за счет чего снижается его кислотность, а растворимость и органолептические показатели повышаются.

Крахмал был подобран из-за его структурообразующих и органолептических качеств. Например, при использовании картофельного крахмала Gletel BAW-40 структура продукта получалась слишком плотной для пудинга, и текстура не была однородной – при органолептическом анализе ощущались крупные частицы.

Ванильный сахар в пудинге играет роль вкусообразующей добавки.

Стабилизатор марки Palsgaard специализируется на производстве ферментированных продуктов. Препятствует отделению сыворотки и пенообразованию и придает продукту гладкую поверхность.

Таким образом, была составлена рецептура обогащенного пробиотиком пудинга на основе молочной сыворотки, которая представлена в таблице 2.

Таблица 2

Разработанная рецептура обогащенного пробиотиком пудинга на основе молочной сыворотки на 1000 кг продукта

Компонент	Количество компонента, кг
Сыворотка сухая деминерализованная	85,25
Вода	767,25
Сахар	100
Ванильный сахар	10
Крахмал кукурузный Gletogel B	22,5
Стабилизатор AcidMilk306 (5846)	15
Итого	1000

В качестве пробиотической закваски использовались следующие микроорганизмы: *Lactococcus lactis subsp. cremoris, Lactococcus lactis subsp. lactis biovar diacetylactis.*

Сама технология претерпела незначительные изменения. Состояла она из этапов проверки качества сырья, диспергирования сухих компонентов в воде, гомогенизации, пастеризации при 90 °С выдержкой 60 секунд, охлаждения, заквашивания, розлива в тару, сквашивания, охлаждения и хранения.

Все сухие компоненты предварительно смешивали. Диспергирование производили в воде, нагретой до температуры 40 °С. Гомогенизация проводилась до пастеризации при температуре 60 °С для получения однородной консистенции. Пастеризация проводилась при температуре 90 °С выдержкой 50–60 секунд. Затем смесь охлаждали до температуры заквашивания (43 ± 2) °С и вносили закваску. После перемешивания разливали в тару, сквашивание производилось уже в таре, чтобы не разрушать сгусток, в течение 6–7 часов при температуре при (43 ± 2) °С. Сквашенный продукт охлаждали до температуры (4 ± 2) °С в холодильной камере и оставляли для структурообразования на 1–3 часа. Готовый продукт оценивали по органолептическим и физико-химическим показателям.

Произвели сравнение показателей полученного пробиотического пудинга (ПП) и пудинга, произведенного по рецептуре из таблицы 1, (П1). Результаты сравнения приведены в таблицах 3 и 4.

Добавление процесса гомогенизации в технологию позволило избавиться от крупных частиц по всему объему продукта и получить однородную, нежную консистенцию. Закваска придала продукту приятную легкую кислинку, а также повысила его кислотность. Показатель предельного напряжения сдвига нужен был для определения вязкости. Чем меньше предельное напряжение сдвига, тем больше вязкость. Таким образом, образец П1 был сильно плотным, желеобразным, почти в 2 раза более вязким, чем образец ПП, хотя по массовой доли сухих веществ они были почти одинаковы.

Таблица 3

Сравнение органолептических показателей ПП и П1

Наименование показателя	Характеристика	
	ПП	П1
Внешний вид и консистенция	Глянцевитая поверхность. Однородная, нежная, в меру плотная консистенция	Густая, плотная консистенция, неоднородная – по всей массе пузырьки воздуха и комочки муки
Вкус и запах	Кисло-сладкий вкус, выраженный запах ванили	Очень сладкий, выраженный вкус и запах ванили
Цвет	Светло-желтый	Светло-желтый

Таблица 4

Сравнение физико-химических показателей ПП и П1

Наименование показателя	Значение показателя	
	ПП	П1
Активная кислотность, ед. pH	5,72	6,20
Предельное напряжение сдвига, Па	4,098	8,233
Массовая доля сухих веществ, %	21,98	21,28

Для того, чтобы оценить влияние закваски молочнокислых бактерий на органолептические, структурно-механические и другие физико-химические свойства молочного пудинга, были изготовлены два образца – пробиотический пудинг по разработанной рецептуре (ПП) и контрольный образец по той же рецептуре без закваски (контроль). Сравнение их показателей представлено в таблицах 5 и 6.

Таблица 5

Сравнение органолептических показателей ПП и контроля

Наименование показателя	Характеристика	
	ПП	Контроль
Внешний вид и консистенция	Глянцевитая поверхность. Однородная, нежная, в меру плотная консистенция	Глянцевитая поверхность. Однородная, нежная, менее плотная и более рыхлая консистенция
Вкус и запах	Кисло-сладкий вкус, выраженный запах ванили	Сладкий, выраженные вкус и запах ванили
Цвет	Светло-желтый	Светло-желтый, но более насыщенный

Таблица 6

Сравнение физико-химических показателей ПП и контроля

Наименование показателя	Значение показателя	
	ПП	Контроль
Активная кислотность, ед. pH	5,72	6,00
Предельное напряжение сдвига, Па	4,098	3,832
Массовая доля сухих веществ, %	21,98	22,66

Деятельность микроорганизмов повысила кислотность продукта, уменьшила количество сухих веществ, но в свою очередь повысила вязкость, сформировав сгусток по всей массе. Именно поэтому контроль не обладал той необходимой плотной консистенцией, он был также нежным и мягким, но более рыхлым. По внешнему виду они мало различались. Глянцевитой поверхностью обладали оба образца, только ПП по цвету был на тон светлее – из-за образования сгустка. В дополнение закваска придала кисло-сладкий вкус десерту, сравнимый с лимонным щербетом. Десерт сохранял все свойства и изначальный внешний вид в течение 7 дней в условиях хранения при температуре (4 ± 2) °С.

По оценке дегустаторов из разных возрастных и социальных групп, полученный пробиотический пудинг получил средний балл 4,6 из 5. Оценивались следующие органолептические показатели по 5-балльной гедонической шкале: внешний вид, консистенция, вкус, цвет и запах. Результаты представлены на рисунке.

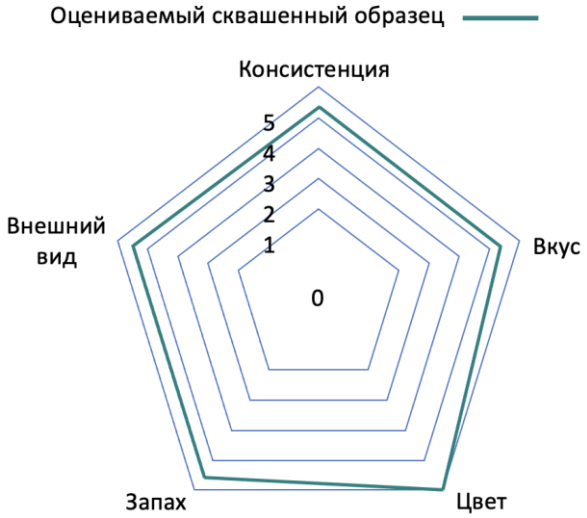


Рисунок. Органолептическая оценка

В ходе работы была разработана рецептура пудинга заквашенного пробиотическим штаммом на основе вторичного молочного сырья, удовлетворяющий требования потребителя по органолептическим показателям. Таким образом, производство такого десерта сможет увеличить оборот перерабатываемой сыворотки, уменьшить ее выбросы, а также расширить ассортимент молочной продукции.

Полученные результаты сравнения с контрольным вариантом без закваски позволили определить, что использование закваски пробиотического штамма молочнокислых бактерий благоприятно влияет на структурно-механические свойства молочного пудинга.

Литература

1. Храмцов А.Г., Нестеренко П.Г. Технология продуктов из молочной сыворотки: Учеб. пособие. 2004. С. 587.
2. Мусина О.Н. Новые молочные продукты для здорового питания // Переработка молока. 2016. № 1. С. 32-35.
3. Маркелова В.В. Разработка технологий пробиотических продуктов из молочной сыворотки, ферментированной экзополисахаридпродуцирующими штаммами L.Acidophilus: автореф. канд. техн. наук. СПб. 2013. С. 16.
4. Уткина О.С., Вострикова С.С., Калашникова А.Ф. Технология производства и оценка качества пудинга на основе молочной сыворотки// Вестник Ижевской государственной сельскохозяйственной академии. 2020. № 4 (64). С. 58-64.

УДК 691.175.5/.8

ИМПЕДАНСНЫЕ ХАРАКТЕРИСТИКИ ПОЛИМЕРНОГО ВОЛОКНИСТОГО КОМПОЗИТНОГО МАТЕРИАЛА

Ситникова А.О.[1] (студент), **Капралова В.М.**[1] (доцент)
Научный руководитель – к.ф.-м.н., доцент Капралова В.М.[1]
1 – Санкт-Петербургский политехнический университет Петра Великого
e-mail: alolsitn@gmail.com

Аннотация
Исследован диэлектрический отклик волокон сополимера ПВДФ-ТрФЭ, модифицированных путем покрытия проводящим полимером (полипирролом), на среды с различным значением кислотности. Методом импедансной спектроскопии измерены частотные зависимости ёмкости и тангенса угла диэлектрических потерь для исходного образца и синтезированного композита. По полученным данным построены годографы, на основании анализа которых сделаны выводы об изменении типа проводимости волокон сополимера ПВДФ-ТрФЭ по действием обработки в средах с различным значением кислотности.

Ключевые слова
Полимерные материалы, композиты, импедансная спектроскопия, диэлектрические свойства, отклик на кислотность среды.

Полимерные материалы обладают рядом уникальных свойств, благодаря которым они имеют очень широкий спектр применений в различных областях и уже используются на практике в течение многих лет. Тем не менее, интерес к их исследованиям не угасает, поскольку, во-первых, периодически синтезируется новые материалы данного класса соединений, а во-вторых, благодаря использованию различных способов модификации свойств уже известных материалов возникают новые возможности их практического использования. В частности, особый интерес к исследованиям полимерных материалов обусловлен возможностью их применения в медицинских целях, так как многие из них являются биосовместимыми. Исследование диэлектрических свойств полимерных материалов методом импедансной спектроскопии позволяет сделать выводы об особенностях их молекулярной структуры, что необходимо для понимания релаксационных процессов, происходящих на молекулярном уровне, что, в свою очередь, способствует увеличению количества способов возможного практического использования конкретного полимера [1].

Сополимер поливинилиденфторида с трифторэтиленом (ПВДФ-ТрФЭ) представляет интерес для исследований, так как является биосовместимым пьезоэлектриком. Благодаря сочетанию пьезоэлектрического эффекта и биосовместимости в данном сополимере его применяют для регенерации нервных окончаний [2]. Покрытие волокон ПВДФ-ТрФЭ проводящим полимером полипиррлом (ППир) позволяет увеличить электропроводность исходного материла и в результате расширить спектр его применения [3]. По этой причине исследование свойств композита на основе пьезоэлектрического и проводящего полимеров, в том числе и их диэлектрических свойств, представляет несомненный интерес и является необходимым шагом на пути к их широкому практическому использованию. Однако на данный момент имеющаяся в мировой научной литературе информация о возможностях варьирования свойств подобного композиционного материала в зависимости от внешних воздействий, например, от варьирования кислотности среды, является явно недостаточной, что и инициировало проведение данного исследования.

Целью данной работы являлось исследование частотных зависимостей диэлектрических характеристик сополимера ПВДФ-ТрФЭ, синтезированного в виде волокон и модифицированного ППир, в зависимости от кислотности среды, а также

построение и анализ годографов для исследованных образцов. В дальнейшем результат работы может быть использован для усовершенствования состава материала с целью применения его в сенсорах, реагирующих на изменение кислотности среды [4].

В проведенном исследовании использовались волокна сополимера ПВДФ-ТрФЭ, полученные методом электроспиннинга из 15% раствора данного материала и смеси растворителей: N-диметилформамида и ацетона в соотношении 6:4. В ходе эксперимента поверхность сополимера была покрыта полипирролом методом полимеризации *in situ* в кислой среде. С помощью измерителя иммитанса Е7-20 были экспериментально исследованы диэлектрические характеристики в диапазоне частот от $f = 120$ Гц до $f = 1$ МГц как для исходного сополимера ПВДФ-ТрФЭ, так и для композитов, полученных после его модификации ППир. Для получения данных о влиянии значений pH среды на свойства композита исследуемые образцы выдерживались в растворах аммиака (NH_4OH, pH = 11.2) или соляной кислоты (HCl, pH = 1.6) в течение суток, а затем высушивались при комнатной температуре в течение двух суток. С целью повышения точности экспериментальных данных было проведено по три серии измерений каждой характеристики. В результате были получены частотные зависимости емкости (C), тангенса угла диэлектрических потерь ($\mathrm{tg}\delta$), модуля импеданса ($|Z|$) и фазового угла (φ) для исходного образца и двух образцов композита после их выдержки в средах с различным значением кислотности. На основании экспериментальных данных для всех исследованных образцов были построены и проанализированы годографы.

Как видно из графика, представленного на рисунке 1, значение тангенса угла диэлектрических потерь исходного материала близко к $\mathrm{tg}\delta \approx 0$ во всем частотном диапазоне, что свидетельствует о том, что исходный материал является диэлектриком. При этом после обработки модифицированного материала в кислоте значения тангенса угла диэлектрических потерь значительно увеличиваются, вплоть до 13 на частоте в 120 Гц, а после обработки в растворе щелочи тангенс угла диэлектрических потерь уменьшается до 1.5 на той же частоте. В результате, значения тангенса угла диэлектрических потерь для образцов, подвергнутых обработке в растворе щелочи, во всем исследованном частотном диапазоне приближаются к значениям, полученным для исходного (непокрытого полипирролом) волокна. Исходя из полученных данных можно сделать вывод, что модификация и последующая обработка проводящего композита в кислой среде значительно увеличивает проводимость исходного материала. Это связано с тем, что ППир синтезируется в проводящей форме, и композит в целом также становится электропроводящим. При этом обработка композита в щелочи снижает его проводимость практически до уровня, соответствующего немодифицированному ПВДФ-ТрФЭ. Таким образом, увеличение значения pH среды приводит к сильному уменьшению проводимости модифицированных волокон.

Как видно из частотных зависимостей емкости исследованных образцов, представленных на рисунке 2, значение емкости исходного сополимера ПВДФ-ТрФЭ составляет около 40 пФ и практически не зависит от частоты. Обработка полученного после модификации поверхности исходного образца ППир композита в средах с различным значением кислотности приводит к изменению характера частотной зависимости ёмкости – в обоих случаях значения емкости уменьшаются с ростом частоты. При этом после обработки композиционного материала в кислоте значения емкости в диапазоне частот от 120 Гц до 10 кГц значительно увеличиваются (на низких частотах вплоть до 160 пФ) и только в области высоких частот приближаются к значениям, наблюдаемым в немодифицированном ПВДФ-ТрФЭ. Обработка композита в щелочной среде приводит к незначительному увеличению значений емкости в области низких (до $f \approx 1$ кГц), а при больших частотах они становятся даже меньше, чем для исходного (немодифицированного ППир) материала, уменьшаясь вплоть до значения 7 пФ на частоте 1 МГц.

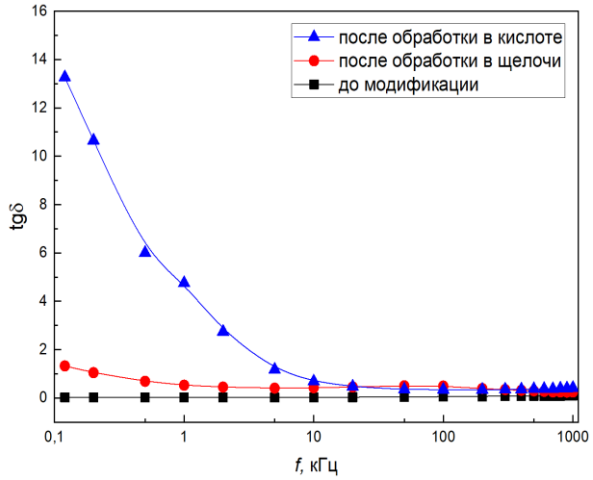


Рис. 1. Зависимость тангенса угла диэлектрических потерь от частоты

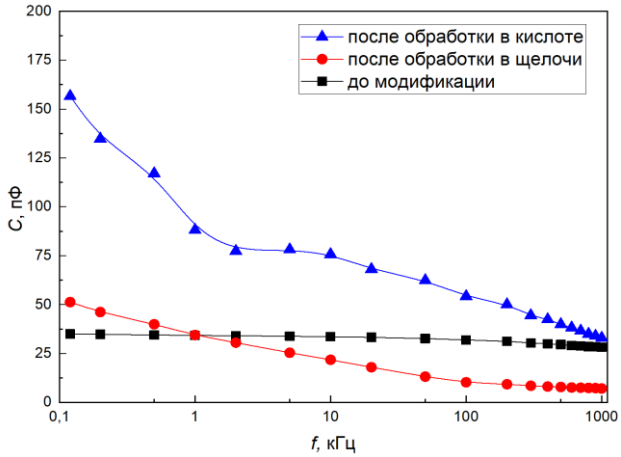


Рис. 2. Зависимость емкости от частоты

Таким образом, значения емкости синтезированного композита после его обработки в средах с различным значением кислотности сильно отличаются во всем частотном диапазоне (различие составляет от 3 до 5 раз). Это указывает на возможность использования композиционного материала на основе ПВДФ-ТрФЭ и ППир в качестве рабочего элемента датчика, реагирующего на кислотность среды, значение которой может быть определено по измеряемому значению емкости.

На основании полученных данных о полном импедансе исследованных образцов и фазовом угле построены годографы для исходных волокон сополимера ПВДФ-ТрФЭ (рис. 3), а также после их модификации ППир и последующей обработки в средах с различным значением кислотности (рис. 4 и рис. 5).

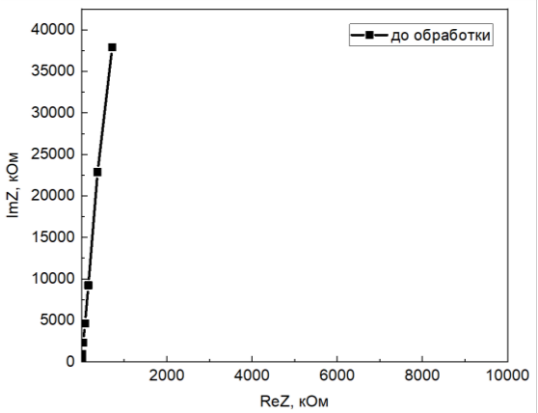


Рис. 3. Годограф для исходных волокон ПВДФ-ТрФЭ

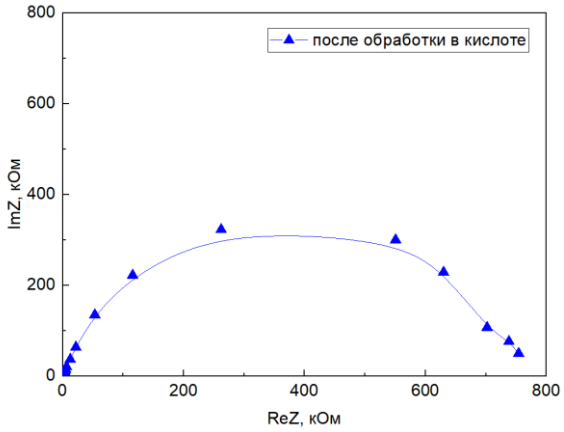


Рис. 4. Годограф для композита на основе ПВДФ-ТрФЭ и ППир
после обработки в кислой среде

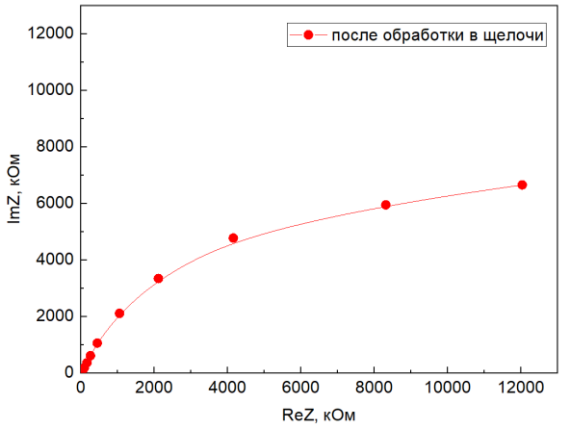


Рис. 5. Годограф для композита на основе ПВДФ-ТрФЭ и ППир
после обработки в щелочной среде

Как видно из рисунка 3, зависимость мнимой части сопротивления от вещественной для исходного (немодифицированного ППир) материала является линейной. Это качественно соответствует импедансной диаграмме диффузионного элемента типа Варбурга в случае бесконечной длины диффузии. Однако в полной мере вид полученных годографов импедансу Варбурга не соответствует, поскольку фазовый угол, сохраняя свою постоянную величину, имеет значение, существенно превышающее $|\varphi| \approx 45°$. Полученные результаты свидетельствуют о том, что проводимость исследуемых образцов только частично обусловлена диффузией частиц вещества. Отметим, что эквивалентной схемой элемента Варбурга является последовательное подключение сопротивления и конденсатора.

По годографу, полученному после обработки композита в кислоте (рис. 4), можно сделать вывод, что в области высоких частот (вблизи нуля и примерно до середины диапазона реальной части полного сопротивления) он совпадает с диаграммой Коул-Коула и частично соответствует полуокружности, но в области низких частот (ReZ > 400 кОм) тип проводимости приближается к индуктивному. Тем не менее фазовый угол φ все равно остается отрицательным, $\varphi = -3°$, поэтому полностью тип проводимости не меняется. Эквивалентной схемой композита в области низких частот является параллельное соединение резистора и конденсатора.

Из зависимости мнимой части полного импеданса от вещественной, полученной для композита, подвергнутого обработке в растворе щелочи (см. рис. 5), можно сделать

вывод об изменении характера проводимости в области низких частот на принципиально противоположный, соответствующий емкостному типу.

Таким образом, в работе проведены экспериментальные исследования и анализ диэлектрического отклика волокон сополимера ПВДФ-ТрФЭ, модифицированных полипирролом, на среды с различным значением кислотности. Обнаружено, что варьирование значения pH среды приводит к изменению как значений диэлектрических характеристик исследованного композита, так и характера их частотных зависимостей. В ходе анализа полученных данных были построены и проанализированы годографы, свидетельствующие об изменении характера проводимости в области низких частот после обработки модифицированных материалов в средах с различным значением кислотности.

Литература

1. Буянова Е.С., Емельянова Ю.В. Импедансная спектроскопия электрохимических материалов: учебное пособие. Екатеринбург: Изд-во УрГУ 2008. 70 с.
2. Kucinska-Lipka J., Gubanska I., Janik H., Sienkiewicz M. Fabrication of polyurethane and polyurethane based composite fibres by the electrospinning technique for soft tissue engineering of cardiovascular system. Gdansk University of Technology, Chemical Faculty, Department of Polymer Technology. 2015. P. 170–171.
3. Капралова В.М., Сапурина И.Ю., Сударь Н.Т., Третьяков А.А., Gryshkov O., Glasmacher B. Импедансная спектроскопия электропроводящих композиционных материалов на основе микроволокон сополимера поливинилиденфторида с трифторэтиленом, модифицированных полипирролом // Письма в журнал технической физики. 2021. Т. 47. Вып. 1. С. 51–54.
4. Maalouf R., Fournier-Wirth C., Coste J., Chebib H., Salikali Y., Vittori O., Errachid A., Cloarec J., Martelet C., Jaffrezic-Renault N. Label-Free Detection of Bacteria by Electrochemical Impedance Spectroscopy: Comparison to Surface Plasmon Resonance. Analytical Chemistry. 2007. V. 79. № 13. P. 4879–4886.

УДК 615.322

β-ЦИКЛОДЕКСТРИН – СИСТЕМА ДОСТАВКИ ПОЛИФЕНОЛОВ НОВОГО ПОКОЛЕНИЯ

Фофанов В.С.[1,2], Мишин К.К.[2], Скворцов Н.В.[2]
Научный руководитель – к.х.н. Попова Е.В.[1,2]
1 – Университет ИТМО
2 – ФГУП НИИ гигиены, профпатологии и экологии человека ФМБА России

Аннотация
Большой интерес исследователей в последнее десятилетие вызывает пряно-ароматическое растение семейства имбирных – куркума. Однако основным препятствием для применения его в лечебной практике является химическая нестабильность, нерастворимость в воде и ограниченное тканевое распределение. В данной работе рассмотрена методика комплексообразования β-циклодекстрина с куркумином. Проведена оценка эффективности данного метода на основе фармакокинетического исследования на крысах.

Ключевые слова
Куркумин, циклодекстрин, комплексообразование, биодоступность, тест растворение, фармакокинетика.

Введение

Куркумин – активный компонент многолетнего растения Curcuma Longa, который обладает разноплановой фармакологической активностью. В последние десятилетия его активно рассматривают как антиоксидант, противовоспалительное и противораковое средство [1]. Препараты куркумина проявляют выраженную противоопухолевую активность в отношении клеток опухоли молочной железы человека MCF-7, клеток лейкоза человека K562, клеток HeLa, клеток гепатомы мыши H22, клеток фибросаркомы человека HT1080, клеток рака прямой кишки, клеток аденокарциномы легкого человека A549. Основные механизмы действия, с помощью которых куркумин проявляет свою уникальную противораковую активность, включают индукцию апоптоза и ингибирование пролиферации и инвазии опухолей путем подавления различных клеточных сигнальных путей [2]. Несмотря на преимущества, главным ограничением для применения куркумина является его низкая растворимость в воде, а соответственно низкая пероральная биодоступность. Помимо этого, он плохо поглощается клетками, так как будучи гидрофобной молекулой прикрепляется к жирным ацильным цепям мембранных липидов, из-за чего практически не проникает в цитоплазму. Одним из методов преодоления данных ограничений является использование различных систем доставки.

В данной работе в качестве носителя был выбран β-циклодекстрин. Циклодекстрины – класс природных макроциклических олигосахаридов, состоящих из определенного числа D-глюкопиранозных звеньев, связанных α-(1-4)-связями и обладающих специфической функциональной активностью [3]. Благодаря гидрофильной оболочке и липофильному ядру они способны образовывать нековалентные комплексы по типу «гость – хозяин», повышая растворимость, биодоступность и химическую стабильность загруженных лекарственных средств [4]. Подобный способ формирования комплекса характеризуется непрерывной ассоциацией и диссоциацией вещества из носителя [5]. При этом ключевую роль в формировании комплекса играет размер полости носителя, который должен быть подходящим для включения молекулы определенного размера.

β-циклодекстрин широко используется в фармацевтической разработке ввиду его широкой доступности и оптимального размера полостей, подходящих для включения самого широкого спектра веществ [6]. Кроме того, известно, что β-циклодекстрины

могут не только повышать растворимость в воде липофильных веществ путем комплексообразования с ними, но и повышать их биодоступность за счет транспортировки непосредственно к барьерной мембране с последующим прохождением и диссоциацией [7]. Наибольший фрамацевтический интерес представляют гидроксипропилпроизводные β-циклодекстрина (HP-β-CD) и γ-циклодекстрина (HP-γ-CD), случайно метилированный β-циклодекстрин (RM-β-CD), сульфобутиловый эфир β-циклодекстрин (SBE-β-CD), монохлортриазиновый β-циклодекстрин (MCT-β-CD), и так называемые разветвленные CDs, такие как мальтозил-β-CD (G2-β-CD) и т. д. [8]. Циклодекстрины внесены в список GRAS (общепризнанные безопасными) Управлением по санитарному надзору за качеством пищевых продуктов и медикаментов (FDA) [9]. Перорально принимаемые циклодекстрины считаются нетоксичными ввиду их низкой абсорбции из желудочно-кишечного тракта.

Для инкапсулирования пищевых биоактивных соединений могут применяться различные методы, такие как эмульсия, экструзия, соосаждение, коацервация, сушка распылением, лиофильная сушка. Правильно выбранный метод инкапсулирования оказывает существенное влияние на комплексообразование. Эффективность того или иного метода зависит в первую очередь от природы препарата и циклодекстрина. Удаление растворителя из комплекса лекарственное средство–циклодекстрин также может влиять на эффективность комплексообразования [10].

Ранее группой Solans и соавторами была продемонстрирована возможность повышения растворимости и устойчивости пальмитата витамина А за счет образования комплекса с β-циклодекстрином без использования органических растворителей [11]. В другом исследовании было установлено улучшение фармакокинетических параметров различных полифенольных соединений и повышение биодоступности за счет образования комплекса с циклодекстринами [12]. В работе группы Капустина были обозначены перспективы практического применения β-циклодекстринов в пищевой промышленности в особенности для наноструктурированных комплексов с пряно-ароматическими соединениями и пищевыми красителями [13].

Целью работы было изучение перспективности применения циклодекстрина в качестве системы доставки куркуминоидов и оценка их биодоступности на животных моделях.

Материалы и методы. Получение комплекса

Для получения комплексов был выбран метод диспергирования. В агатовой ступке растирали куркумин и β-циклодекстрин в мольном соотношении 1:3 в течение 20 минут до образования однородной смеси. Аналогично повторялось для смеси с соотношением куркумин β-циклодекстрин 1:9. Для сравнения был выбран 3 образец, куркумин без добавления β-циклодекстрина.

Испытание тест растворение проводят в две стадии

Первая стадия (кислотная). По 750 мл хлористоводородной кислоты раствора 0,1 М помещают в каждый из 6 сосудов для растворения. Доводят температуру среды растворения до (37 ± 0,5) °С. Помещают по 1 капсуле в каждый из 6 сосудов для растворения, включают мотор перемешивающего устройства.

Отбирают аликвоту через 30 мин, 60 мин, 90 мин, 120 мин и сразу же продолжают процесс растворения в щелочной среде.

Отобранную аликвотную часть раствора анализируют с помощью спектрофотометрии.

Вторая стадия (щелочная). В каждый из 6 сосудов для растворения прибавляют по 250 мл натрия фосфата раствора 0,2 М ($Na_3PO_4 \times 12H_2O$), температура которого составляет (37 ± 0,5) °С (перемешивающее устройство аппарата продолжает работать).

Доводят pH среды растворения до 6,80 ± 0,05 с помощью хлористоводородной кислоты раствора 2 М или натрия гидроксида раствора 2 М. Отбирают аликвоту через 165 мин, 180 мин, 240 мин, 300 мин. Отобранную аликвотную часть раствора анализируют с помощью спектрофотометрии [14].

Фармакокинетическое исследование

Для опыта были выбраны самцы крыс линии Wistar, вес которых на момент операции составлял 350 - 370 гр. Им давали, пищевые гранулы ad libitum и чистую питьевую воду. Температура в помещении была установлена на уровне 21°C ±2°C при относительной влажности 55% ±10%. Циклы освещения автоматически контролируются при 12: 12-часовом цикле свет / темнота, установленном на 100 Люкс.

Все препараты куркумина растворяли в 2% растворе крахмала в объёме 1 мл:

1. Доза для препарата CUR -40 мг на крысу.
2. Доза для препарата CUR/βCD (1:3) – 160 мг на крысу.
3. Доза для препарата CUR/βCD (1:9) – 400 мг на крысу.

Препараты вводились внутрижелудочно с использованием специального зонда анатомической формы. Предварительно за два часа до эксперимента у крыс был отобран корм.

Через промежутки времени, равные: 30 мин; 60 мин; 120 мин; 180 мин; 240 мин; 300 мин; 360 мин и 24 часа. Содержание куркуминоидов определяли методом ВЭЖХ.

Объемы отобранных проб были равны 200 микролитрам и замещались тем же объемом гепаринизированного физиологического раствора внутривенно подопытному животному.

Результаты

Известно, что куркуминоиды представляют собой смесь куркумина и двух его производных, деметоксикуркумина и бисдеметоксикуркумина (рис. 1) [15]. Ароматические группы и изопреноидный линкер в структуре молекул куркуминоидов обеспечивают их гидрофобность. Особенности строения куркуминоиды хорошо растворимы в метаноле, этаноле, диметилсульфоксиде и ацетоне.

Рис. 1. Химическая структура куркуминоидов

Полученные методом диспергирования комплексы исследовали методами ИК Фурье спектрометрии, ЯМР спектроскопии и спектрофотометрическими методами. При этом, как правило, используется не один, а одновременно несколько подтверждающих друг друга методов.

ИК-спектроскопия с преобразованием Фурье часто используется для оценки формирования комплексов. Поскольку при формировании комплекса происходят

значительные изменения связей, можно наблюдать характерные для включенной молекулы пики. В случае физической смеси куркумина и носителя наблюдалось наложение друг на друга их спектров без каких-либо изменений.

Поскольку ЦД не проявляют значимого УФ-поглощения, метод спектрофотометрии удобен для количественного исследования включения и высвобождения куркумина из комплекса. В ходе формирования комплексов с куркумином и исследования его высвобождения в тесте растворения в имитаторах сред желудочно-кишечного тракта наблюдали характерный пик поглощения на длине волны 430 нм.

После усреднения полученных значений в ходе теста растворения и градуировки видим схожие зависимости для обоих образцов CUR/βCD (1:3) и CUR/βCD (1:9). Максимальная концентрация достигается спустя 240 минут после начала опыта. Активный рост концентрации начинается после 120 минут эксперимента, что совпадает с подщелачиванием среды. С помощью теста растворение сложно с достаточной достоверностью определить наиболее удачное соотношение куркумин-циклодекстрин. Данная зависимость представлена на рисунке 2.

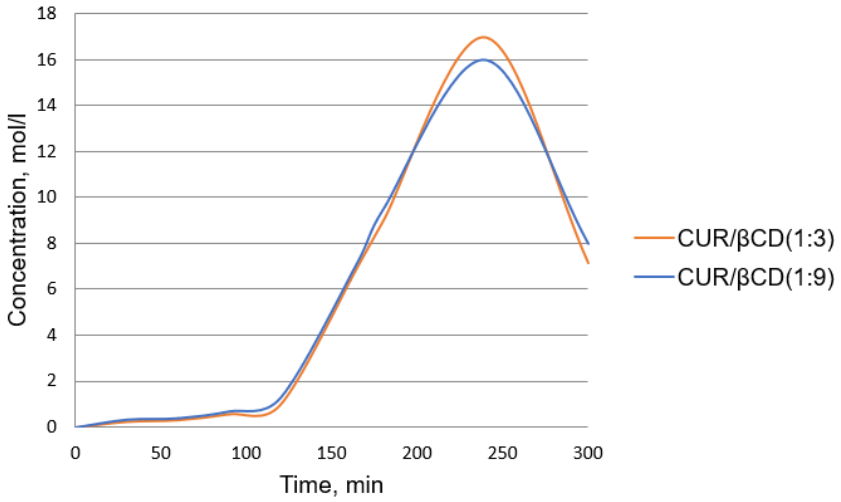


Рис. 2. Зависимость концентрации растворенного куркумина
от времени, прошедшего с начала опыта

Фармакокинетические исследования чистого куркумина выявили плохую абсорбцию и стремительный метаболизм, что серьезно снижает его биодоступность [16]. Фармакокинетическое исследование полученных комплексов показало, что обе его формы с различным соотношением куркумин/β-циклодекстрин имеют более высокие пиковые концентрации по сравнению с чистым куркумином. Пиковые концентрации достигаются в другой временной промежуток, что связано с различными значениями уровня pH у крыс и у человека. В данном случае pH, необходимый для активного растворения комплекса, пик проявлялся уже в передней части желудка животных согласно литературным данным, представленным в таблице 1.

Таблица 1

Показатели pH в различных частях пищеварительного тракта у человека и крыс

Вид животных	Желудок		Двенадцатиперстная кишка	Ободочная кишка	Прямая кишка	Фекалии
	Передняя часть	Задняя часть				
Человек	1,3-2,0	5,6	5,6-7,9	7,2-7,5	6,8-7,6	6,0-8,0
Крыса	3,0-5,0	3,3	5,0-7,1	5,1-6,8	5,9	6,9

В данной работе рассматривалась как сумма куркумина, диметоксикуркумина и бисдиметоксикуркумина, так и определялась концентрация отдельно каждого куркуминоида. Так, через 90 мин после начала опыта концентрация суммы куркуминоидов в случае комплекса с соотношением 1:3 более чем в 3 раза превышает концентрацию чистого вещества. Данная зависимость продемонстрирована на рисунке 3.

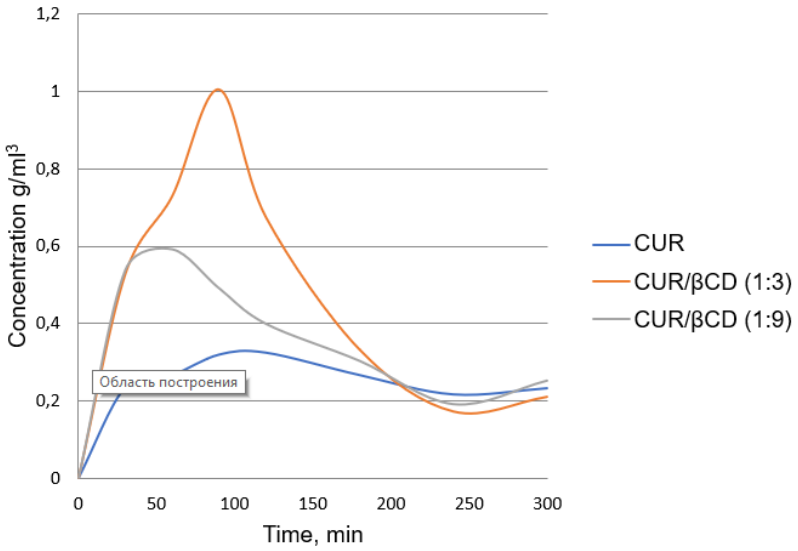


Рис. 3. Зависимость концентрации суммы куркуминоидов от времени в организме крыс

Содержание отдельных куркуминоидов в крови животных через 90 минут после приема представлено в таблице 2. Очевидно, что смесь с носителем почти в 6 раз биодоступнее, чем чистый куркумин.

Таблица 2

Сравнение концентраций куркуминоидов для различных образцов

Доза на животное	Куркумин, г/мл^3	Диметокси-куркумин, г/мл^3	Бисдиметоксикуркумин, г/мл^3
40 мг экстракта куркумина	0,138	0,163	0,018
160 мг смеси Куркумин + β-cyclo-Dextrin (1:3)	0,769	0,213	0,024
400 мг смеси Куркумин + β-cyclo-Dextrin (1:9)	0,283	0,187	0,021

Заключение

В данной работе были исследованы комплексы куркумина с β-циклодекстрином, полученные методом диспергирования. Тест растворение показал, как ведет себя комплекс в модели желудочно-кишечного тракта человека. Важно отметить, что повышение концентрации характерно именно в слабощелочной среде. Фармакокинетическое исследование, проведенное на биообъектах, показало эффективность применения комплекса CUR/βCD (1:3) по сравнению с чистым куркумином. В качестве дальнейших исследований планируется тестирование других методик комплексообразования куркумина с β-циклодекстрином, таких как соосаждение, замешивание и др.

Литература

1. Tomeh M.A., Hadianamrei R., Zhao X. A review of curcumin and its derivatives as anticancer agents //International journal of molecular sciences. 2019. Т. 20. №. 5. С. 1033.
2. Куннумаккара А.Б., Бордолой Д., Падмавати Г., Мониша Дж., Рой Северная Каролина, Прасад С., Аггарвал Б.Б. Куркумин, золотой нутрицевтик: многоцелевое применение при множественных хронических заболеваниях. бр. Дж. Фарм. 2017. 174. 1325-1348.
3. Иванова Л.А. и др. Получение комплексов циклодекстринов с биологически активными веществами //Пищевая промышленность. 2011. №.10. С. 67-69.
4. Gidwani B., Vyas A. A comprehensive review on cyclodextrin-based carriers for delivery of chemotherapeutic cytotoxic anticancer drugs //BioMed research international. 2015. Т. 2015.
5. Stella V.J. et al. Mechanisms of drug release from cyclodextrin complexes //Advanced drug delivery reviews. 1999. Т. 36. №. 1. С. 3-16.
6. Challa R. et al. Cyclodextrins in drug delivery: an updated review //Aaps Pharmscitech. 2005. Т. 6. С. E329-E357.
7. Haimhoffer Á. et al. Cyclodextrins in drug delivery systems and their effects on biological barriers //Scientia Pharmaceutica. 2019. Т. 87. №. 4. С. 33.
8. Vyas A., Saraf S., Saraf S. Cyclodextrin based novel drug delivery systems //Journal of inclusion phenomena and macrocyclic chemistry. 2008. Т. 62. С. 23-42.
9. Arya P., Raghav N. In-vitro studies of Curcumin-β-cyclodextrin inclusion complex as sustained release system //Journal of Molecular Structure. 2021. Т. 1228. С. 129774.
10. Jacob S., Nair A.B. Cyclodextrin complexes: Perspective from drug delivery and formulation //Drug development research. 2018. Т. 79. №. 5. С. 201-217.
11. Vilanova N., Solans C. Vitamin A Palmitate-β-cyclodextrin inclusion complexes: Characterization, protection and emulsification properties //Food chemistry. 2015. Т. 175. С. 529-535.
12. Сокуренко М.С. и др. Использование циклодекстринов с целью повышения биодоступности полифенольных соединений //Евразийский Союз Ученых. 2016. №. 3-4 (24). С. 155-156.
13. Капустин М.А. и др. Методы получения нанокомплексов биологически активных веществ с циклическими олигосахаридами, анализ их физико-химических свойств и использование в пищевом производстве //Труды Белорусского государственного университета. 2016. Т. 11. №. ч 1. С. 73-100.
14. ОФС.1.4.2.0014.15 Растворение для твердых дозированных лекарственных форм [Электронный ресурс]. - URL: http://pharmacopoeia.ru (дата обращения: 01.06.2018).
15. Капустин М.А. и др. Куркуминоиды: получение, свойства и применение. Сообщение 2. Антиоксидантная и антимутагенная активность куркуминоидов //Современная наука и инновации. 2023. №. 3. С. 105-118.
16. Осикбаева С.О., Гончарова Т.Г., Оразгалиева М.Г. Противоопухолевое действие куркумина и карнозоловой кислоты (обзор литературы) //Онкология и радиология Казахстана. 2019. Т. 4. №. 54. С. 45-50.

УДК 547.97

АНТОЦИАНЫ И БЕТАЛАИНЫ КАК ИНДИКАТОРЫ ПОРЧИ МЯСНЫХ ПРОДУКТОВ

Фролова К.О.[1] (аспирант)
Научный руководитель – к.т.н., доцент Еремеева Н.Б.[1]
1 – Университет ИТМО
e-mail: ksenya.fr.good98@gmail.com

Работа выполнена при финансовой поддержке Российского научного фонда (РНФ), проект №23-26-00056.

Аннотация
В работе рассмотрены антоцианы и беталаины как пигменты растительного происхождения, отмечена их чувствительность к изменению pH. Проведен анализ последних разработок на тему интеллектуальной упаковки для мясных продуктов, выделены исследования. Выделен и проанализирован ряд исследований, показывающих перспективность применения антоцианов и беталаинов в качестве индикаторов порчи мясных продуктов в составе интеллектуальной упаковки.

Ключевые слова
Антоцианы, беталаины, индикаторы pH, мясные продукты, интеллектуальная упаковка.

Мясные продукты являются важной составляющей человеческого рациона. Они относятся к скоропортящимся продуктам. Микробиологическая порча, окисление липидов и автолитические процессы, протекающие под действием ферментов, представляют собой три основных механизма порчи мяса в процессе переработки и хранения [1]. Мясные продукты являются благоприятной средой для жизнедеятельности разнообразной микрофлоры (как естественной, так и случайной), в том числе патогенных микроорганизмов. В результате метаболизма микроорганизмов мясо может подвергаться ослизнению, плесневению, кислотному брожению, гниению и другим процессам. При этом в зависимости от условий и микрофлоры образуются летучие соединения азота, сероводород, диоксид углерода и прочие соединения, меняющие pH продукта [2].

В настоящее время активно исследуется возможность применения различных форм и композиций интеллектуальных упаковок для хранения скоропортящихся пищевых продуктов, в том числе мяса, рыбы и морепродуктов [3]. Интеллектуальная упаковка представляет собой систему, способную обеспечивать мониторинг состояния упакованных пищевых продуктов и условий внутри упаковки с целью предоставления потребителю информации о качестве продукта. К основным категориям интеллектуальной упаковки относят индикаторы, датчики и носители данных [4].

Индикатор способен давать информацию о качестве и безопасности пищевых продуктов за счет визуально определимого изменения цвета в процессе хранения. Цветовые индикаторы реагируют с веществами, образующимися в результате роста микроорганизмов в упакованных пищевых продуктах, меняют цвет в результате изменения pH [5]. Одна из разновидностей индикаторов, индикатор pH, состоит из двух частей: основы и красителя, чувствительного к pH. В качестве красителей применяются пигменты растительного происхождения, в том числе антоцианы, куркумин, беталаины, ализарин, шиконин, каротиноиды [4, 5].

Цель данной работы заключается в обосновании применения антоцианов и беталаинов в качестве индикаторов в составе интеллектуальной упаковки.

Антоцианы представляют собой гликозиды антоцианидина, основу которого составляет 2-фенилбензопирилий, или катион флавилия. В настоящий момент идентифицировано 23 антоцианидина, в которых варьируется количество и положение

гидроксильных групп и/или групп метилового эфира [6]. Антоцианы являются пигментами, обуславливающими различные оттенки цвета у плодов, листьев, цветов и запасающих органов растений. К источникам антоцианов относятся виноград, краснокачанная капуста, черника, черноплодная рябина, брусника и др. Концентрация антоцианов в большинстве фруктов и овощей составляет от 0,1 до 1% по массе [6].

Изменение окраски антоцианов в зависимости от pH связано со структурными модификациями, происходящими вследствие амфотерной природы антоцианов (рис. 1). При низких значениях pH антоцианы существуют в виде катионов флавилия, характеризующихся красным цветом. При повышении pH до 4-5 образуются бесцветные псевдооснование карбинол и халкон, а интенсивность красного понижается. В слабокислых условиях (pH 6-7) преобладают сине-фиолетовые нейтральные хиноиды. В ходе увеличения pH до 8 образуются синие анионные хиноидные основания. При дальнейшем повышении pH происходит деградация замещающих групп антоцианов, в результате чего образуется желтый халкон [5].

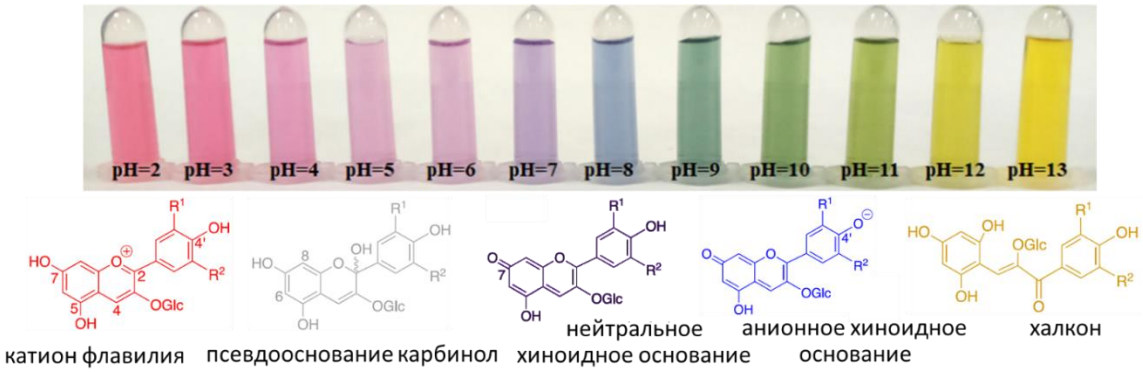


Рис. 1. Зависимость окраски раствора антоцианов от pH на примере экстракта фиолетового сладкого картофеля (фото взяты из [7])

Беталаины относятся к N-гетероциклическим пигментам, являющимся производными беталамовой кислоты. Беталаины обнаружены в 17 семействах растений порядка Гвоздичноцветные и встречаются у некоторых высших грибов [8]. Известно небольшое количество съедобных источников беталаинов, среди которых красная и желтая свекла, цветной мангольд, зерновой или листовой амарант и плоды кактусов родов *Opuntia* и *Hylocereus* [9]. В красной свекле концентрация беталаинов составляет в среднем 130 мг/100 г сырого веса [6].

Беталаины подразделяют на две группы на основании структурных характеристик и светопоглощающих свойств:

- красно-фиолетовые бетацианины;
- желтые бетаксантины [10].

В случае бетацианинов беталамовая кислота конденсируется с производными цикло-3,4-дигидроксифенилаланина (цикло-ДОФА). Бетаксантины представляют собой беталамовую кислоту с заместителями из различных аминокислот и ряда аминных соединений.

При высоких значениях pH в беталаинах происходят структурные изменения, вызывающие изменение цвета (рис. 2). Естественный pH беталаинов в дистиллированной воде составляет 3,85, что придает ей красный цвет. Раствор беталаинов имеет красную окраску в диапазоне pH 3-7. При увеличении pH до 8-9 раствор становится оранжевым, а при pH 10-12 – желтым, что объясняется постепенной деградацией беталаинов до предшественников: беталамовой кислоты и цикло-ДОФА в случае бетацианинов или аминных соединений в случае бетаксантинов [5].

Рис. 2. Зависимость окраски раствора беталаинов от pH на примере экстракта кожуры красной питайи (фото взяты из [11])

За последнее время был проведен ряд исследований по разработке интеллектуальной упаковки для мясных продуктов, в которых в качестве индикатора применялись антоцианы [12, 13], беталаины [14, 15], а также их смесь [16, 17]. Для подтверждения интеллектуальных свойств упаковки авторы проводили эксперименты, во время которых применяли разработанные материалы при хранении мясных продуктов. В ходе экспериментов контролировали качество хранящихся продуктов и следили за изменениями в интеллектуальной упаковке в ответ на процессы, протекающие в упакованных продуктах. При вынесении заключения о способности разработанной упаковки выявлять порчу пищевого продукта важным пунктом является возможность определить изменение цвета индикатора, соответствующее испорченному состоянию, невооруженным глазом.

Рассмотрим подробнее несколько исследований. Ализаде-Сани и коллеги [13] разработали пленку на основе метилцеллюлозы и хитозановых волокон с антоцианами, выделенными из экстракта барбариса в качестве индикатора. Пленка показала отчетливое изменение цвета с повышением pH от 2 до 14: от темно-красного до бледно-персикового (при pH 4-5), светло-розового (при pH 6-7) и оттенков желтого (при pH 8-14). Авторы проводили эксперимент, в котором использовали разработанную пленку как ярлык-индикатор для мониторинга свежести мяса ягненка, помещенного в упаковку и хранимого при комнатной температуре в течение 72 часов. Было отмечено, что значение pH мяса в ходе эксперимента повысилось с 5,8 до 7,5, при этом цвет пленки изменился с темно-красного до бледно-персикового (рис. 3, слева). Таким образом, подтверждается, что пленка визуально идентифицирует процесс порчи мяса, однако требуется дальнейшее исследование.

Гуо и др. [15] исследовали пленку на основе пектина, полученного из кожуры арбуза, с добавлением различных концентраций бетацианинов, выделенных из свеклы (1, 2 и 3% по массе). Результаты изучения зависимости цвета пленки от pH продемонстрировали, что пленка проявляет заметное изменение цвета при различных значениях pH. Цвет пленки изменился от темно-красного до розового с увеличением pH от 3 до 7, затем с бордового на желтый – при повышении pH от 8 до 10 (рис. 3, справа). Для исследования возможности применения разработанной пленки для мониторинга качества говядины проводили следующий эксперимент. Образцы мяса помещали в чашки Петри, на крышки которых с внутренней стороны прикрепляли пленку, чашки запаковывали полиэтиленовой пленкой и выдерживали при 4 °C и постоянном значении относительной влажности 75% в течение 8 дней. Состояние мяса определяли по показателю общего содержания летучих азотистых оснований, который регламентируется китайскими государственными стандартами. В результате эксперимента за 8 дней показатель повысился с 6,73 до 18,71 мг/100 г продукта

(критическое значение составляет 20 мг/100 г), при этом цвет пленки изменился с красного на желто-коричневый (рис. 3, справа). Соответственно, пленка способна визуально сигнализировать о порче продукта, но необходимо четкое определение оттенков пленки, соответствующих удовлетворительному по качеству и испорченному мясу.

Рис. 3. Результаты эксперимента по мониторингу свежести мяса посредством применения индикаторной пленки с антоцианами барбариса в составе (слева) [13] и пленки с введением бетацианинов из свеклы (справа) [15]

Лю и коллеги [17] разработали пленку на основе поливинилового спирта и Na-карбоксиметилцеллюлозы с экстрактами антоцианов краснокочанной капусты и бетацианинов кожуры красной питайи в соотношениях 1:3, 2:2 и 3:1. Пленки продемонстрировали следующее изменение цвета с повышением pH от 2 до 12: от красного (при pH 2-3) до бледно-розового (при pH 4-5), кремового (при pH 6-7), зеленоватого (при pH 8-10) и желтого (при pH 11-12). Ученые исследовали возможность применения пленок с целью мониторинга свежести мяса, для чего образцы свинины помещали в упаковку с прикрепленным на крышку кусочками пленки, укупоривали и выдерживали при 25 °C в течение 28 часов. Каждые 4 часа определялись цветовые параметры пленки, а также показатели мяса – содержание летучих азотистых оснований и pH (начальные значения 5,61 мг/100 г продукта и 6,05 соответственно). В соответствии с китайскими государственными стандартами критическим значением содержания летучих азотистых оснований для свинины является 15 мг/100 г продукта, что было достигнуто образцами между 20 и 24 часами хранения. Таким образом, после 24 часов хранения мясо было признано испорченным, что визуально идентифицировалось изменением цвета пленок, которые приобрели отчетливый зеленый оттенок (рис. 4). Лю и др. отметили, что наиболее заметные изменения цвета показала пленка, содержащая антоцианы и бетацианины в соотношении 3:1 (рис. 4, CPA-3A1B), соответственно, данная пленка признана наиболее подходящей для мониторинга свежести свинины.

Более того, на основании экспериментальных данных авторы разработали шкалу свежести свинины. Данные, полученные при хранении образцов от 0 до 4 часов, отнесли к группе "свежий продукт", данные от 8 до 20 часов – к группе "продукт средней свежести", оставшиеся данные определили в группу "испорченный продукт". Результаты исследования имеют большое практическое значение и перспективны для внедрения в промышленность. Следующим этапом может быть проведение подобного эксперимента и разработка шкалы свежести продукта при холодильном хранении.

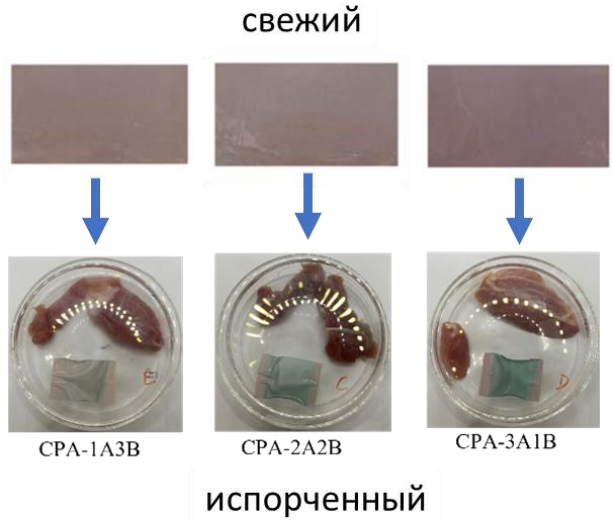


Рис. 4. Результаты эксперимента по мониторингу свежести мяса посредством применения индикаторной пленки совместным введением антоцианов из краснокочанной капусты и бетацианинов из кожуры красной питайи [17]

На основе анализа опубликованных на данный момент исследований можно сделать вывод о том, что антоцианы и беталаины подходят для применения в качестве индикаторов за счет своей pH-чувствительности. Можно отметить, что на индикаторные свойства пигментов влияет их источник и концентрация в интеллектуальной упаковке.

Предлагается разработка интеллектуальной биоразлагаемой упаковки для мясных продуктов с использованием антоцианов и беталаинов в качестве индикаторов. Источниками получения антоцианов выбраны красный виноград и черноплодная рябина, беталаины будут выделяться из красной свеклы.

Литература

1. Dave D., Ghaly A.E. Meat Spoilage Mechanisms and Preservation Techniques: A Critical Review // American Journal of Agricultural and Biological Sciences. 2011. Vol. 6. № 4. P. 486–510.
2. Мишанин Ю.Ф. Биотехнология рациональной переработки животного сырья: учебное пособие для вузов. Санкт-Петербург: Лань. 2021. 720 с.
3. Tahir H.E. et al. Smart films fabricated from natural pigments for measurement of total volatile basic nitrogen (TVB-N) content of meat for freshness evaluation: A systematic review // Food Chemistry. 2022. Vol. 396. P. 133674.
4. Ma Y. et al. Properties and Applications of Intelligent Packaging Indicators for Food Spoilage // Membranes. 2022. Vol. 12. P. 477.
5. Priyadarshi R., Ezati P., and Rhim J.-W. Recent Advances in Intelligent Food Packaging Applications Using Natural Food Colorants // ACS Food Science & Technology. 2021. Vol. 1. № 2. P. 124–138.
6. Delgado-Vargas F., Jiménez A.R., Paredes-López O. Natural pigments: carotenoids, anthocyanins, and betalains – characteristics, biosynthesis, processing, and stability // Critical reviews in food science and nutrition. 2000. Vol. 40. № 3. P. 173–289.
7. Zhao M. et al. Monitoring dynamic changes in chicken freshness at 4 °C and 25 °C using pH-sensitive intelligent films based on sodium alginate and purple sweet potato peel extracts // International Journal of Biological Macromolecules. 2022. Vol. 216. P. 361–373.
8. Sadowska-Bartosz I., Bartosz G. Biological Properties and Applications of Betalains // Molecules. 2021. Vol. 26. P. 2520.
9. Azeredo H.M.C. Betalains: properties, sources, applications, and stability – a review // International Journal of Food Science and Technology. 2009. Vol. 44. P. 2365–2376.

10. Socaciu C. (ed.) Food colorants: chemical and functional properties / 1st edition. – Boca Raton: CRC Pressю 2007. 648 p.
11. Qin Y. et al. Development of active and intelligent packaging by incorporating betalains from red pitaya (*Hylocereus polyrhizus*) peel into starch/polyvinyl alcohol films // Food Hydrocolloids. 2020. Vol. 100. P. 105410.
12. Chumee J. et al. Colorimetric biofilm sensor with anthocyanin for monitoring fresh pork spoilage // Materials Today: Proceedings. 2022. Vol. 65. P. 2467–2472.
13. Alizadeh-Sani M. et al. pH-responsive color indicator films based on methylcellulose/chitosan nanofiber and barberry anthocyanins for real-time monitoring of meat freshness // International Journal of Biological Macromolecules. 2021. Vol. 166. P. 741-750.
14. Kanatt S.R. Development of active/intelligent food packaging film containing *Amaranthus* leaf extract for shelf life extension of chicken/fish during chilled storage // Food Packaging and Shelf Life. 2020. Vol. 24. P. 100506.
15. Guo Z. et al. Active-intelligent film based on pectin from watermelon peel containing beetroot extract to monitor the freshness of packaged chilled beef // Food Hydrocolloids. 2021. Vol. 119. P. 106751.
16. Qin Y. et al. Comparison of the physical and functional properties of starch/polyvinyl alcohol films containing anthocyanins and/or betacyanins // International Journal of Biological Macromolecules. 2020. Vol. 163. P. 898–909.
17. Liu D. Novel colorimetric films based on polyvinyl alcohol/sodium carboxymethyl cellulose doped with anthocyanins and betacyanins to monitor pork freshness // Food Chemistry. 2023. Vol. 404. P. 134426.

УДК 504.06

РАЗРАБОТКА ЭКОЛОГИЧЕСКОЙ ДЕКЛАРАЦИИ НА ПРИМЕРЕ РЫБНЫХ ПОЛУФАБРИКАТОВ СО СЪЕДОБНЫМ ПОКРЫТИЕМ

Шестель А.А.[1] (студент), **Гаврецкая А.В.[1]** (студент)
Научный руководитель – д.т.н., профессор Куприна Е.Э.[1]
1 – Университет ИТМО
e-mail: shestel@mail.ru

Работа выполнена в рамках темы НИР №620142 «Ресурсосберегающие экологически безопасные технологии биоактивных пищевых ингредиентов с заданными свойствами из биоресурсов и вторичных сырьевых источников растительного и животного происхождения для обеспечения здорового питания населения РФ с использованием пищевой комбинаторики».

Аннотация
В работе рассмотрена разработка экологической декларации для рыбных полуфабрикатов. Определена функциональная единица и границы продукционной системы. Проведен инвентаризационный анализ жизненного цикла и оценка воздействия на окружающую среду. По результатам проведения оценки жизненного цикла составлены рекомендации. На основе проведенного исследования составлена экологическая декларация.
Ключевые слова
Оценка жизненного цикла, окружающая среда, полуфабрикат, негативное воздействие, экологическая маркировка.

В 2015 году мировым сообществом были разработаны цели устойчивого развития, которые должны способствовать благополучному развитию и повышению качества жизни общества. Одной из них является Цель 12: обеспечение перехода к рациональным моделям потребления и производства. Среди задач необходимых для ее реализации основными являются рациональное и эффективное использование природных ресурсов и снижение образования отходов [1]. Для этого необходимо проведение исследования производственных циклов и потоков, а так же планирование модернизации производства. Этому способствует разработка экологических деклараций, в том числе проведение оценки жизненного цикла (ОЖЦ).

Первым этапом необходимым для проведения оценки жизненного цикла является выделение границ системы [2]. В рамках данной работы рассматривалось производство рыбных полуфабрикатов со съедобным биодеградируемым покрытием. Этапы производства и границы системы приведены на рисунке 1. Затем было необходимо выбрать функциональную единицу. В рамках данного производства была выбрана функциональная единица – 10 кг рыбного полуфабриката. Затем был произведен расчет материального баланса по каждой стадии производственного процесса с учетом образующихся стоков и выбросов.

Далее для расчета показателя MIPS была составлена таблица, приведенная на рисунке 1. Данный показатель позволяет определить количество ресурсов, затрачиваемых на получение 1 кг продукта. Для проведения расчета использовались коэффициенты MI-Factor, разработанные Вуппертальским институтом климата, окружающей среды и энергетики [3]. Было рассчитано, что для получения 1 кг рыбного полуфабриката со съедобным покрытием необходимо 429,3 кг ресурсов.

На рисунке 2 показано распределение влияния используемых веществ и полупродуктов на такие категории как абиотические, биотические, воздушные и водные ресурсы. Так как данное производство относится к рыбоперерабатывающей отрасли преобладание негативного воздействия на водные ресурсы ожидаемо.

Name substance/pre-product	Unit	Amount	Abiotic material		Biotic material		Water		Air	
			MI-Factor kg/unit	kg/unit Main product	MI-Factor kg/unit	kg/unit Main product	MI-Factor kg/unit	kg/unit Main product	MI-Factor kg/unit	kg/unit Main product
Форель	кг	15,02	2,80	42,06	4,70	70,59	271,00	4070,42	0,83	12,47
Электричество	кВт-ч	0,34	2,67	0,92	0,00	0,00	37,92	13,08	0,64	0,22
Вода питьевая	кг	2,52	0,08	0,20	0,00	0,00	2,20	5,54	0,01	0,03
Вода техническая	кг	2,40	0,01	0,02	0,00	0,00	2,20	5,28	0,01	0,02
NaCl	кг	0,12	1,24	0,15	0,00	0,00	2,29	0,27	0,02	0,00
HCl	кг	0,00	3,03	0,00	0,00	0,00	40,70	0,02	0,38	0,00
Хитозан	кг	0,00	1,30	0,00	5,00	0,00	19,28	0,02	3,08	0,00
Альгинат натрия	кг	0,01	1,18	0,01	0,00	0,00	6,30	0,03	0,29	0,00
Масло подсолнечное	кг	0,22	3,15	0,69	2,54	0,56	51,04	11,23	0,73	0,16
Хлорид кальция	кг	0,00	2,46	0,00	0,00	0,00	11,65	0,00	0,09	0,00
Лимонная кислота	кг	0,00	1,00	0,00	1,00	0,00	7,00	0,01	0,01	0,00
Мука	кг	0,41	1,68	0,69	2,12	0,87	42,85	17,57	1,76	0,72
Перец	кг	0,01	0,80	0,00	1,53	0,01	9,43	0,05	0,15	0,00
Картон	кг	0,40	1,86	0,74	0,75	0,30	93,56	37,42	0,33	0,13
Σ				45,49		72,34		4160,96		13,76
Total		4292,54								

Рис. 1. Расчет показателя входных потоков

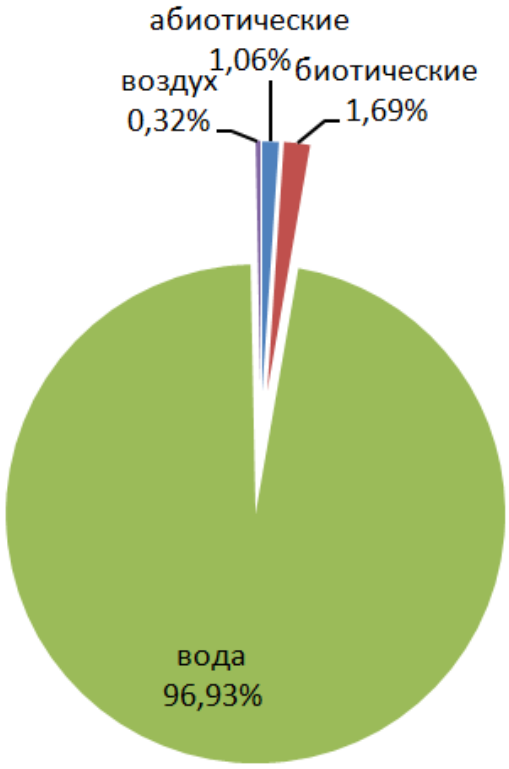


Рис. 2. Распределение воздействий

На следующем этапе выходным потокам, выявленным при проведении инвентаризационного анализа, были присвоены категории воздействия. Они представлены в таблице.

После определения категорий воздействия рассчитывался вклад каждой стадии в них. После распределения данных значений по производственным стадиям был получен график, изображенный на рисунке 3.

Таблица

Категории воздействий

Параметр	Содержание параметра
Категория воздействия	Эвтрофикация
Результаты инвентаризационного анализа жизненного цикла	Выбросы оксида азота
Показатель категории	г – эквивалент O_2
Характеристический коэффициент	Потенциал эвтрофикации г - эквивалент O_2
Категория воздействия	Закисление
Результаты инвентаризационного анализа жизненного цикла	Выбросы оксида азота, соляной и лимонной кислот
Показатель категории	мол. H^+
Характеристический коэффициент	Потенциал закисления мол. H^+
Категория воздействия	Образование фотохимического смога
Результаты инвентаризационного анализа жизненного цикла	Костная пыль, хлор
Показатель категории	г – эквивалент этана
Характеристический коэффициент	Потенциал фотохимического образования озона г – эквивалент этана
Категория воздействия	Глобальное потепление
Результаты инвентаризационного анализа жизненного цикла	Выбросы диоксида углерода, выбросы оксида азота (I)
Показатель категории	г – эквивалент CO_2
Характеристический коэффициент	Потенциал глобального потепления г – эквивалент CO_2
Категория воздействия	Образование побочного продукта
Результаты инвентаризационного анализа жизненного цикла	Образование побочных продуктов
Показатель категории	Образование белкового гидролизата и жира
Характеристический коэффициент	Показатель образования побочной продукции, г

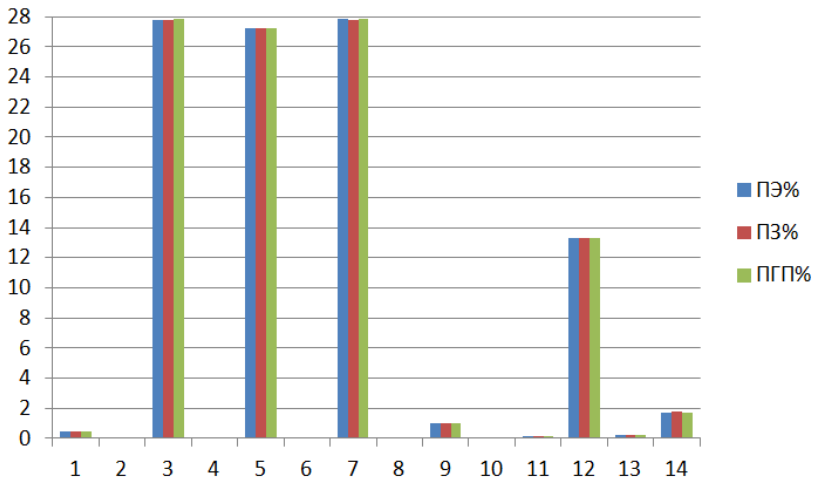


Рис. 3. Соотношение воздействий, оказываемых производственными стадиями ПЭ – потенциал эвтрофикации; ПЗ – потенциал закисления; ПГП – потенциал глобального потепления. Все величины приведены в процентах. 1 – разделка и подготовка рыбы; 2 – диспергирование отходов; 3 – смешивание сырья с католитом; 4 – фильтрование; 5 – обработка в катодной камере электролизера; 6 – диспергирование коллагенсодержащих отходов; 7 – смешивание с католитом; 8 – нейтрализация; 9 – сепарирование; 10 – смешение; 11 – измельчение и смешение; 12 – выдержка и формование; 13 – нанесение покрытия; 14 - упаковка

Из данной диаграммы видно, что наибольшее воздействие оказывают стадии 3,5,7. Данные стадии связаны с работой электролизера, который отличается высоким энергопотреблением.

После обработки всех полученных данных, на основе жизненного цикла продукции, разрабатывалась экологическая декларация, внешний вид которой представлен на рисунке 4. Она включает в себя информацию о продукции, в том числе о стадиях жизненного цикла и о величинах основных экологических показателей, которые рассчитаны на функциональную единицу.

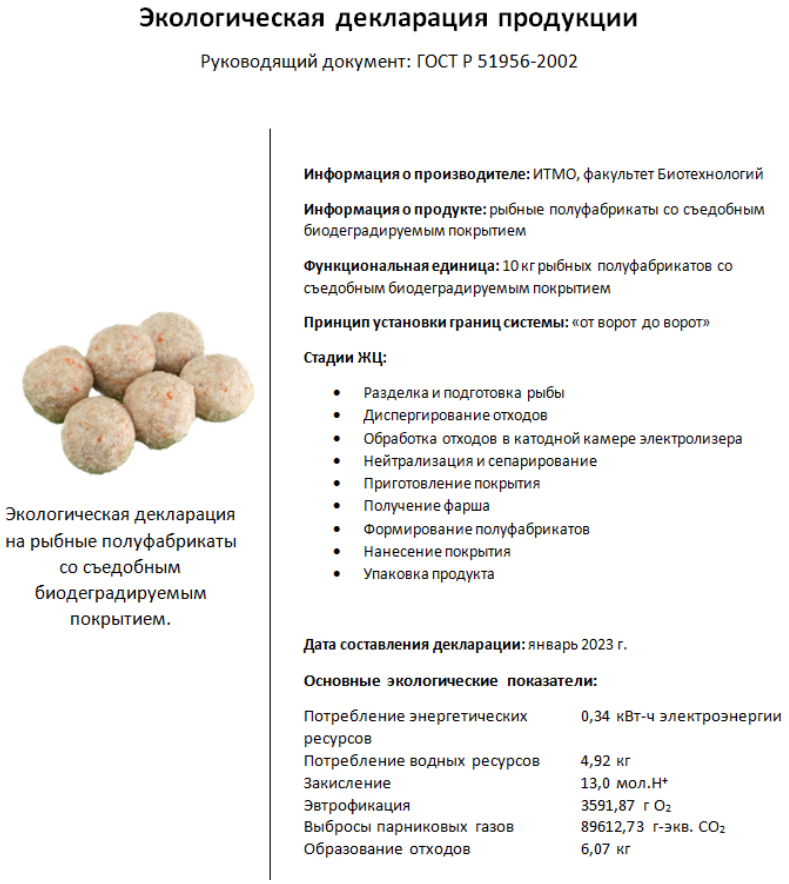


Рис. 4. Экологическая декларация рыбных полуфабрикатов со съедобным биодеградируемым покрытием

Для разработки экологическая декларация рыбного полуфабриката со съедобным биодеградируемым покрытием была выполнена оценка жизненного цикла продукции. Проведен инвентаризационный анализ и оценка воздействия на окружающую среду. Для снижения негативных воздействий рекомендуется:

- применение энергосберегающего оборудования;
- реализация образующихся побочных продуктов;
- применение оборотной тары.

Литература

1. Цели в области устойчивого развития [Электронный ресурс]. Режим доступа: https://www. un.org /sustainabledevelopment/ ru/sustainable- development-goals/ (дата обращения: 10.04.2023).
2. ГОСТ Р ИСО 14040-2010. Экологический менеджмент. Оценка жизненного цикла. Принципы и структура. Введен 01.06.2010. М.: Стандартинформ. 2010. 18 с.
3. Wuppertal Institute [Электронный ресурс]. Режим доступа: https://wupperinst.org/ (дата обращения: 13.04.2023).

ИНФОРМАЦИОННАЯ БЕЗОПАСНОСТЬ

УДК 004.056

INFORMATION SECURITY RISKS AT AUTOMATED INDUSTRIAL INFORMATION SYSTEM FIRMS

Amanuel Estifanos Asfha[1], Ph.D., Associate Professor Zakoldaev D.A.[1]
Ph.D., Assistant Professor Abhishek Vaish[2]
1 – ITMO University
2 – Indian Institute of Information Technology, Allahabad
baquesti2003@gmail.com

Abstract
Data is kept in complicated ways by information systems, which makes finding the data much simpler. It supports a company's decision-making process. It is easier to make better decisions when all relevant information is delivered through an information system. Therefore, because of their bulk, the significance of business processes, and their impact on the world and people's lives, industrial firms attract convicts. Luckily, cyberattacks on industrial firms with such catastrophic implications are rare, as they necessitate a higher degree of attacker qualification. In 2020, according to a survey by cybersecurity firm Positive Technologies, the industrial sector was the second most common target for hackers after the government sector.

Keywords
Industrial information system, information security, threat, vulnerability, and risk.

Previously, industrial automation systems were not linked to one another or to public networks such as the Internet. Today, the scenario is slightly different: because the market requires enterprises to make quick and cost-effective decisions, accurate and up-to-date information on the plant and process status must be available not just on the plant floor, but also to management and supply chain partners. As a result, there is improved interoperability between different automation systems as well as between automation and office systems.

Modern industrial automation systems are mostly based on commercial operating systems, protocol, and communication applications designed for the office IT environment [1] **.** As a result, most industries are currently influenced by the advancement or improvement of digital technology related to instrumentation and industrial automation (IA), Supervisory Control and Data Acquisition (SCADA) systems are used to manage and control equipment used in industrial facilities such as power plants, water treatment plants, or refineries, and it is also used to reduce waste, improve efficiency, and report any problems [2]. This SCADA, also known as an ICA (Industrial Control System), interfaces with sensors and systems in real time. Hackers may attack any number of key systems once inside an industrial environment. They may seek to manage the opening or closing of circuit breakers or change system logic to prevent alarms from tripping, depending on the facility [3]. An attacker might potentially impersonate a power meter or other sensors to enter bogus input values into a SCADA system. The SCADA system may respond to inaccurate values, resulting in unneeded shutdowns or equipment failures [4].

A cyber-attack can inflict financial loss as well as damage to equipment or infrastructure. Reduced productivity and unplanned downtime can have long-term consequences for brand reputation and trust when a facility's quality of service deteriorates [5]. Finally, humans are the most vulnerable cybersecurity risk in any industry. Malicious codes, for example, are typically distributed through human error. An email attachment is opened, memory sticks are inserted, cellphones are charged, laptops are connected to important networks, and so on. As a result, human errors are regarded as the greatest digital vulnerability in the ant sector [6].

As a result, the financial consequences of unwanted incidents based on cybersecurity vulnerabilities are that production must be stopped down, resulting in a loss of money for the sector. Indirect and indirect taxes will be reduced in society. Unwanted incidents will harm the

company's reputation, and if saboteurs and terrorist organizations gain control of critical manufacturing facilities, the company's reputation will suffer, in the worst case, environment destruction and fatalities may result [7].

Research Methodology

Process and eligibility criteria

The preferred reporting items for systematic review and meta-analysis framework (Page et al..,2021) served as the model for this systematic review's structure, and search was carried out in 2022. Only publications published in the recent eight years were taken into consideration due to the ongoing growth of information security vulnerability in industrial information systems and their responses. Additionally, only English-language articles that that were peer-reviewed were included. Only papers that use one or more information /cybersecurity risk datasets met the inclusion criteria as a last criterion.

Information sources

Several databases were searched to gather pertinent literature on the subject of information security risk in automated industrial information systems in order to cover the full range of available literature. The literature search was constrained to the following four databases for ease of use due to the dispersion of similar publications over numerous database: IEEE Xplore, Scopes, Springer Link, and Web of Science.

Search and study selection

All databases were searched using a single search string that was made. The following search string with a Boolean operator was used across all databases to ensure an effective and repeatable search: Database OR dataset AND cyber risk OR information security. Some changes have to be done for the individual search engine in order to guarantee consistency of the search across all databases.

Result and discussion

Most common cyberattacks in industrial information system

A purposeful and malicious attempt to compromise the information system of another person or organization is known as a cyberattack. Typically, the attacker intends to profit by disrupting the victim's network.

Malware

Malware is a catch-all phrase for all sorts of malicious software, which in the context of computer security means: software used to attempt to break a computer system's security policy in terms of Confidentiality, Integrity, or Availability**.** Malware is usually classified into several groups based on how it is introduced into the target system and the type of policy violation it is intended to cause, such as computer viruses, worms, Trojan horses, adware, and spyware...etc.

Phishing Attack

Phishing is a type of cybercrime in which targets are contacted via email, phone, or by mimicking a legitimate entity in order to dupe people into revealing sensitive information such as personally identifiable information, banking and credit card information, and passwords**.**

Password attack

A password attack is a technically sophisticated method of obtaining and misusing a genuine password in order to gain unwanted access to networks and user accounts. It is a relatively successful method for a perpetrator to obtain secret or sensitive information or to exploit it in a more sophisticated hack.

Distributed Denial of services (DDoS)

DDoS attacks are a type of denial of service (DoS) attack that attempts to exhaust network, server, or application resources so that they are no longer available to intended users, and it is a type of malicious cyber-attack used by hackers or cybercriminals to make an online service, network resource, or host machine unavailable to its intended users on the Internet.

Man-in-the-middle

One of the most popular types of network assaults is man-in-the-middle. This attack occurs when the attacker is able to get between two elements of a communication: the sender and the receiver.

Drive by download attack

By-pass download in recent years, assaults in which web browsers are hijacked by malicious material sent by web servers have become a common attack vector. When an Internet user accesses a rogue web page, a malicious web server sends a malicious HTML document to the user's computer system. The malicious content then exploits weaknesses in the visitor's computer system, such as web browser, plug-in, and operating system vulnerabilities.

Malvertising

According to the SANS institute, malvertising is "the installation of unwanted or outright malicious software through the use of internet advertising media networks, exchanges and other user supplied content publishing services common to the Social Networking space,"

According to the SANS organization, it is "the installation of undesirable malicious software through the use of internet advertising media networks, connections and other user provided content publishing services common to the Social Networking space,". The technique of putting malicious code into legitimate adverts is known as malvertising.

Rogue software attack

Rogue security software is becoming a more serious threat to computer security. It is a type of Internet scam in which malicious code is used to trick people into paying for a phony malware cleanup. In this regard, malicious software known as rogue security software is a form of ransomware or scareware (malicious programming that affects a computer system and requests money in order for the restriction to be lifted, in this case a fictitious security issue).

Most common consequences of cyberattacks in industrial sector

Interruption of operations

Honda had to suspend operations at numerous locations for an entire day due to a ransomware attack on the company's corporate network. After such an assault, it takes a significant amount of effort for the organization to resume normal operations, restore full functionality of technological and business systems, and prevent similar disasters in the future.

Disruption of technological processes

According to Tripwire integrity management, the Colonial Pipeline transports 2.5 million barrels per day, accounting for approximately 45% of the East Coast's gasoline, diesel, and jet fuel. The pipeline was down for six days following a ransomware attack on the organization. Normal operations did not resume fully for another three days. As a result, the average national price of gasoline increased to its highest level in six years, with prices climbing by six to 19 cents per gallon. The corporation paid the hackers a ransom of 75 bitcoins (about $4.4 million at the time).

- ✓ For example, in May 2021, Colonial Pipeline, the largest fuel provider in the United States, was hit by a ransomware attack. The company's computer systems went down for a week, resulting in the closure of half of the company's gas stations in numerous southeastern states, an increase in wholesale gasoline prices, and a surge in fuel demand.
- ✓ In 2020, terrorists attempted to compromise Israel's water supply and purification infrastructure. Finally, in February 2021, a hacker gained access to the water treatment facilities of a small American city and altered the chemical makeup of the water.

Disruption of business processes

Due to a hacking attack in February 2020, Croatian oil business INA was unable to send invoices, register loyalty card use, issue new mobile vouchers, or allow users to pay gas utility

bills. The Clop ransomware was responsible for the disruption of business activities by encrypting data on the company's internal systems.

Today, the primary goal of information security specialists is to assess the feasibility of various security risks in companies and identify the potential consequences of cyberattacks; and then build an efficient security system based on this knowledge. The issue is that a company's management will never agree to any infrastructural move that could harm the technical progress.

Analyzing the possibility of risks at the cyber domain

Cyberexercises are orchestrated attacks designed to test and improve the detection and response capabilities of information security professionals. Attacks are simulated in the Standoff Cyberexercises by attacker teams comprised of information security experts.

✓ In May 2021, attacker teams were tasked with triggering information security concerns at a gas distribution station at The Standoff cyber-range. The circumstances were as close to truth as feasible.

✓ According to the positive technology information, it took the assailants two days to disrupt the station's gas flow. They obtained access to the gas station management system, shut down the gas supply, and set off an explosion. At the company's office, the defenders recorded an average of 32 events per day.

✓ Human losses, management departure, lawsuits, erroneous actions or omissions by staff in the event of an emergency, equipment failure, restoration costs, and loss of income due to downtime can all result from such hazards being triggered at a gas distribution station.

✓ When examining the security of an industrial company's genuine infrastructure, information security specialists will most likely be able to demonstrate just gas transportation data spoofing in the monitoring system. They will almost definitely not be permitted to execute attacks that disrupt or impede technological or business activities, making a risk feasibility analysis impossible.

The most typical flaws in industrial information system

Most industrial enterprises have virtually little defense against cyber threats. Our research uncovered the following most common vulnerabilities:

✓ the external network perimeter is poorly protected and easily accessible from the Internet, Low level of protection against hackers infiltrating the industrial network, Device misconfiguration, In network segmentation and traffic filtering, Dictionary passwords, and Use of outdated software;

Based on the aforementioned vulnerabilities, security assessments done in 2020 found that an external attacker can infiltrate the corporate network in 91 percent of industrial organizations. Complete control of the infrastructure in 100% of situations, and they can take sensitive data in 69% of cases, including information about partners and corporate employees, email conversations, and internal documentation.

✓ most notably, 75 percent of industrial organizations based on the results of 12 internal and external security evaluations of corporate information systems done between 2017 and 2020, with the purpose of gaining access to industrial networks;

✓ information security professionals gained access to the network's technological part. In 56% of cases, this allowed criminals to gain access to industrial control systems (ICS). Such figures are concerning, and it is critical to comprehend their origins. PT NAD (Positive Technology Network Attack Discovery) test experiment completed by their experts reveals various suspicious events in each industrial company's internal network.

PT NAD is a network traffic analysis system that detects attacks on the network's perimeter and within it. PT NAD monitors the entire network, detects malicious behavior even in encrypted traffic, and aids in incident investigation (figure).

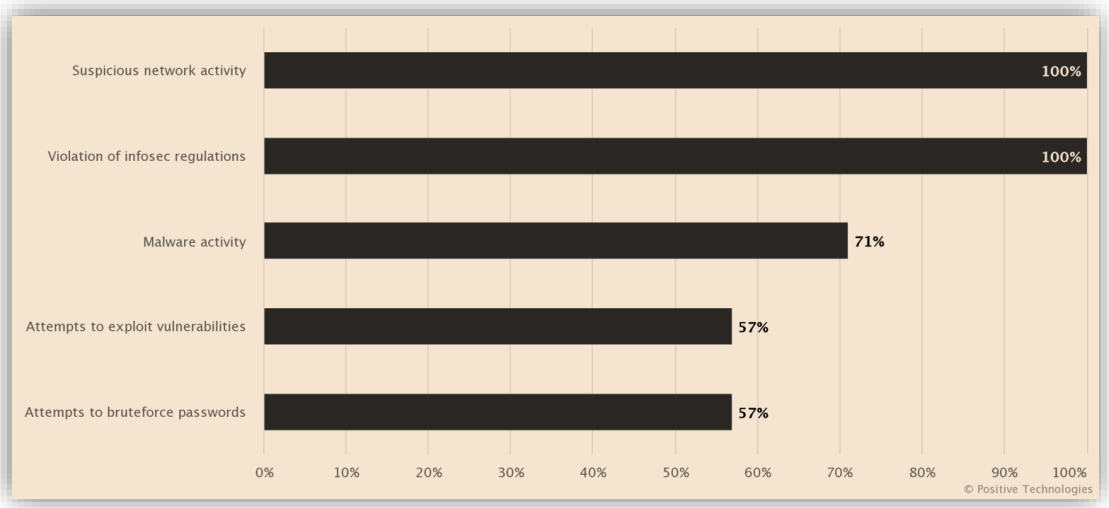


Figure. Information security threats actuated during the PT NAD pilot projects at industrial companies in 2020 (source: Positive Technology website)

For example, PT NAD registered an RDP connection to an external cloud storage at one industrial company. RDP and HTTPS were used to transfer a total of 23 GB of data to the address of this storage.

✓ furthermore, industrial organizations frequently deploy outdated software. According to Claroty statistics, the number of vulnerabilities in ICS components has gradually increased in recent years;

✓ security specialists discovered 25% more vulnerabilities in 2020 than in 2019, mostly in the energy sector, manufacturing, and water treatment plants;

✓ the issue is that updating industrial equipment necessitates a unique "maintenance window," which is only permitted for a few hours every week or even month.

On essential systems, the assault vector can be straightforward. Positive Technologies professionals, for example, accessed the corporate network and secured maximum domain privileges during a security audit of one industrial organization. They then gathered information about the industrial network hosts, received ICS equipment connection methods, and discovered that just one computer was linked to the ICS network. The specialists used this node to connect to the industrial network.

Saving connection parameters (username and password) in a remote access authentication form is a common mistake while administering such devices (for example, via RDP). An attacker can connect to the resources of an isolated segment without credentials by gaining control of such a computer. Furthermore, connection parameters, addresses, schemes, and passwords for industrial network systems are frequently maintained in clear text (for example, in Excel spreadsheets) on the computers of engineers and other personnel in charge. Attackers with access to ICS can cause the following risks:

✓ interruption of operations;
✓ industrial equipment failure;
✓ damage to products;
✓ accidents.

However, because of the unfavorable impact on technological processes, it is impossible to test these hazards in real infrastructure.

Industrial security to protect themselves against potential risks

✓ creating a strong information/cybersecurity policy and a secure network regardless of how many devices, applications, and servers the organization employs;

✓ securing critical data access from any place or device, especially in the WFH period;

✓ backing up data on a regular basis and constantly monitoring all activity related to vital resources;

✓ raising cybersecurity knowledge among employees through a people-centric strategy, as personnel can be an organization's greatest security risk or its most effective security defense;

✓ mandating the use of password managers, SSOs, and multi-factor authentication (MFA), as the usage of repetitious or weak passwords is still a frequent practice among employees of global corporations today;

✓ implementing an enterprise password manager is the most effective choice for improving an organization's cybersecurity and data privacy problems;

✓ implement a tight password policy;

✓ separate the company LAN and external networks from the ICS operating networks;

✓ physical access to ICS networks and components should be restricted;

✓ configure network equipment and firewall filtering rules correctly;

✓ safeguard privileged accounts;

✓ reduce the privileges of users and services;

✓ make use of antivirus software;

✓ install operating system and program updates on a regular basis.

Finally, if the system owner had applied the techniques outlined in the table below, 98% of the incidents reported would have been avoided (table).

Table

Risk management system

No	Respondent	%
1.	Implemented application whitelisting	38
2.	Configuration/patch management	29
3.	Reduce your attack surface	17
4.	Build a defendable environment	9
5.	Manage authentication	4
6.	Monitor and respond	2
7.	Implement secure remote access	1

Conclusion

According to a survey by cybersecurity firm Positive Technologies, the industrial sector was the second most common target for hackers after the government sector in 2020. As a results, the most vulnerabilities of industrial firms to information/cyberattacks were low level of protection of the external network perimeter accessible from the internet, low level of protection against hackers able to penetrate the industrial network, device misconfiguration, flaws in network segmentation and traffic filtering, dictionary passwords, and use of outdated software, and also the most common consequence in this industry being plant sabotage/shutdown, equipment damage, product quality(bad quality), safety violation (Death or injury) ,compliance violation, disruption of technological processes, and disruption of business processes.

Reference

1. Bajer M. “Dataflow in modern industrial automation systems. Theory and practice.” [Online]. Available: https:// www.researchgate.net/publication/320187633. (дата обращения: 02.03.2023).
2. Stouffer K., Pillitteri V., Lightman S., Abrams M. and Hahn A. “Guide to Industrial Control Systems (ICS) Security,” Gaithersburg, MD, Jun. 2015. doi: 10.6028/NIST.SP.800-82r2.
3. “Once inside an industrial environment”. [Электронный ренсурс]. https:// www. accuenergy.com/articles/cyber-security-in-industrial-environments/. (дата обращения: 02.03.2023).

4. Pliatsios D., Sarigiannidis P., Lagkas T. and Sarigiannidis A.G. “A Survey on SCADA Systems: Secure Protocols, Incidents, Threats and Tactics,” IEEE Communications Surveys and Tutorials. vol. 22. no. 3. pp. 1942–1976. Jul. 2020. doi: 10.1109/COMST.2020.2987688.
5. “Integrated infrastructure: cyber resiliency in society Cambridge Centre for Risk Studies Cambridge Risk Framework Cyber Risk Framework for Critical Infrastructure Threat Scenario.” [Online]. Available: http://www.risk.jbs.cam.ac.uk. (дата обращения: 02.03.2023).
6. Cremer F. et al. “Cyber risk and cybersecurity: a systematic review of data availability,” Geneva Papers on Risk and Insurance: Issues and Practice. vol. 47. no. 3. pp. 698–736. Jul. 2022. doi: 10.1057/s41288-022-00266-6.
7. Poljanšek K. et al. “Recommendations for National Risk Assessment for Disaster Risk Management in EU Where Science and Policy Meet Version 1.” doi: 10.2760/80545.

УДК 004.716

ОБЗОР МЕТОДОВ МОНИТОРИНГА ВСТРАИВАЕМЫХ СИСТЕМ НА ПРИМЕРЕ КОНТРОЛЛЕРА ЛИНИЙ СИСТЕМЫ АВАРИЙНОГО ОПОВЕЩЕНИЯ

Асьминкин Ф.А.[1] (аспирант)
Научный руководитель – к.пед.н., доцент Авксентьева Е.Ю.[1]
1 – Университет ИТМО
e-mail: faasminkin@itmo.ru

Аннотация
В данной статье рассматривается возможность использования методов машинного обучения для повышения эффективности системы аварийного оповещения. Контроллер линии аварийного оповещения представляет собой встраиваемую систему, в задачи которого входит эффективный контроль и мониторинг линии громкоговорителей. Авторы исследования предлагают повысить надежность системы аварийного оповещения за счет разработки программного обеспечения контроллера линии на основе методов машинного обучения для повышения точности и чувствительности результатов измерений импеданса.

Ключевые слова
Контроллер, импеданс, машинное обучение, измерения, надёжность/

Система аварийного оповещения используется для информирования людей при возникновении чрезвычайных ситуаций на производстве, поэтому предъявляются высокие требования к надежной работе такой системы. Система аварийного оповещения представляет собой аппаратно-программный комплекс, который включает контроллер, вызывающее устройство, абонентское устройство, усилитель мощности звука, линия громкоговорителей. Контроллер системы аварийного оповещения является встраиваемой системой, позволяющей осуществлять контроль и мониторинг линий и громкоговорителей. Контроллер системы аварийного оповещения также способен установить наличие короткого замыкания и внешнего постороннего напряжения на линии, вызванное неправильным подключением линии громкоговорителей или сторонним воздействием. Возможны ситуации, при которых кабели линии громкоговорителей могут быть оборваны, тогда оповещение пройдет по дополнительным линиям громкоговорителя. Устройство автоматически определяет, где и по какой причине произошла неисправность, сбой в работе линии или громкоговорителя. Информация о состоянии проверяемых объектов передается в систему мониторинга. Это позволяет сохранять работоспособность линии оповещения при возникающих критических ситуациях. Важно обеспечить, чтобы контроллер системы аварийного оповещения выполнял свои задачи максимально эффективно и обеспечивал надежную работу аварийного оповещения на производственных объектах.

В научной литературе выделяют следующие ключевые критерии, которые следует учитывать при оценке эффективности технической системы:

- надежность: определяет способность системы последовательно и без сбоев выполнять свои функции в течение определенного периода времени. Надежная система позволяет сократить время простоя, повысить эффективность работы;
- удобство пользования: это мера того, насколько легко и эффективно пользователи могут взаимодействовать с системой. Удобная, интуитивно понятная и доступная система может повысить производительность, а также сократить время обучения и повысить удовлетворенность пользователей;
- масштабируемость: это способность системы справляться с растущими рабочими нагрузками и приспосабливаться к росту с течением времени.

Масштабируемая система может адаптироваться к изменяющимся требованиям и обеспечивать будущий рост;

- производительность: этот критерий определяет быстродействие системы в целом. Хорошо работающая система может повысить производительность, минимизируя время на выполнение определенных задач;
- безопасность: это критический фактор при оценке эффективности технической системы. Безопасная система защищает конфиденциальную информацию, обеспечивает конфиденциальность данных и снижает риск кибератак и утечек данных;
- интероперабельность: этот критерий определяет возможность системы интегрироваться и беспрепятственно работать с другими системами и технологиями;
- ремонтопригодность: данный критерий подразумевает наличия у системы простоты в обслуживании. Ремонтопригодные системы легче поддерживать, а также сокращается время простоя, повышается доступность системы ввиду снижения на неё затрат.

В данной работе будут рассмотрены возможные варианты применения методов машинного обучения во встраиваемых системах для повышения эффективности работы контроллера линий громкоговорителей системы аварийного оповещения.

Встраиваемые системы стали неотъемлемой частью нашей повседневной жизни, начиная от смартфонов и заканчивая медицинскими устройствами и автомобилями. Они предназначены для выполнения конкретного функционала и интегрируются в более крупные системы, что позволяет решать комплексные задачи. За последние несколько лет в технологиях встраиваемых систем произошли многочисленные усовершенствования, которые сделали эти системы более надежными, эффективными и экономичными.

Одним из значительных улучшений встроенных систем является внедрение более передовых процессоров. Развитие интернета вещей привело к разработке более компактных и мощных процессоров, способных выполнять сложные задачи при меньшем энергопотреблении. Эти процессоры способны запускать несколько приложений и работать на высоких скоростях, что делает их идеальными для приложений встраиваемых систем.

Интеграция искусственного интеллекта и машинного обучения во встроенных системах привела к разработке передовых систем помощи водителю, умных домашних устройств и других приложений, которые могут улучшить нашу повседневную жизнь.

Еще одним улучшением данных систем является использование искусственного интеллекта и машинного обучения для измерений. Эти технологии позволяют встраиваемым системам принимать решения в режиме реального времени на основе анализа данных, что приводит к более эффективным и точным расчетам.

В статье “Artificial Intelligence for Efficient Thermal Comfort Systems: Requirements, Current Applications and Future Directions” [1] приводится модель встраиваемой системы, которая участвует в мониторинге условий помещения с целью их корректировки для благоприятного пребывания людей. Для достижения этой цели автор использует метод случайного леса, который является совокупностью нескольких деревьев принятия решений, работающих связано. Помимо измеряемых величин внутри помещения, могут измеряться также и внешние, косвенно помогающие установить время года и дня, что уже в свою очередь будет использовано, для расчетов условий самого помещения.

В другой статье автор использует нейронную сеть для корректировки измеренных координат с помощью GPS [2]. Однако более интересное применение искусственного интеллекта в данной разработке является определение местоположения возникающих во времена потерь связи со спутником. На основе ранее полученных данных с датчиков производится расчет координат.

В статье “Artificial intelligence driven hydrogen and battery technologies” [3] автор показывает, как искусственный интеллект способен контролировать и поддерживать в нормальных условиях процесс получения водорода.

В следующей статье “Hall Effect Sensor and Artificial Neural Networks Applied on Diagnosis of Broken Rotor Bars in Large Induction Motors” [4] искусственный интеллект используется для определения состояния электрического двигателя. С помощью датчика Холла измеряется магнитное поле рядом с катушками двигателя. Определяется плотность магнитного поля. С помощью нейронной сети производится анализ данных, в результате которого можно определить место поломки и количество сломанных элементов двигателя. Автор утверждает, что применение искусственного интеллекта позволило повысить точность и чувствительность результатов измерений. Используются нейронные сети прямого распространения.

В статье “Capacitive sensor-based fluid level measurement in adynamic environment using neural network” [5] с помощью нейронных сетей производится определение уровня жидкости, перевозимой в ёмкостях автотранспорта. Автор передает данные, полученные с датчиков уровня в нейронную сеть трёх видов. Лучше всех себя показала самая простая из трёх - нейронная сеть прямого распространения. Она имеет минимальный процент ошибок 0.11 %. Вторая нейросеть уже имеет 2.6 % ошибки. Нелинейная авторегрессивная сеть с внешними входными параметрами выдает результаты с максимальным процентом ошибок – 8.7 %. Автор объясняет данные результаты тем, что последние две нейросети требуют больше данных и времени обучения по сравнению с первой. Использование методов искусственного интеллекта в данной работе позволяет сократить ошибки в измерениях и повысить их точность.

Использование технологий подключения, таких как Wi-Fi, Bluetooth и Zigbee позволило встроенным системам взаимодействовать друг с другом, создавая сети устройств, которые могут работать вместе для выполнения сложных задач. Это привело к разработке умных домов, систем промышленной автоматизации и других приложений, которые работают с другими устройствами.

Наконец, стоимость встраиваемых систем с годами значительно снижается, что делает их более доступными для более широкого круга приложений. Это привело к разработке новых приложений в таких областях, как здравоохранение, транспорт и производство, где встраиваемые системы могут сыграть решающую роль в улучшении качества и эффективности выполняемых операций.

Следует отметить, что достижения в технологии встраиваемых систем значительно улучшили возможности и производительность этих систем. Благодаря постоянной интеграции передовых процессоров, искусственного интеллекта и машинного обучения, технологий подключения и экономической эффективности, возможности встраиваемых систем постоянно растут.

Для задачи контроля и мониторинга линии громкоговорителей в настоящее время существуют следующие способы решения:

- акустическое тестирование: проводя акустическое тестирование громкоговорителей в линии, можно определить, генерируют ли какие-либо громкоговорители звук, который значительно отличается от других громкоговорителей. Это может указывать неисправность громкоговорителя;
- визуальный осмотр: визуально осмотрев громкоговорители в линии, можно определить, есть ли у каких-либо громкоговорителей физические повреждения или другие проблемы, которые могут повлиять на их работу;
- анализ сигналов: анализируя сигналы, генерируемые громкоговорителями в линии, можно определить, создают ли какие-либо громкоговорители сигналы, которые существенно отличаются от сигналов других громкоговорителей;
- электромагнитное тестирование: проводя электромагнитное тестирование громкоговорителей в линии, вы можете определить, создают ли какие-либо

громкоговорители электромагнитные помехи, которые значительно отличаются от других громкоговорителей.

Используя комбинацию этих методов, можно выявить и диагностировать ошибки в одном или нескольких громкоговорителях в линии. После выявления ошибки ее устраняют для восстановления производительности линии.

Линии с громкоговорителями должны проверяться регулярно. Однако большинство из данных методов не позволяет автоматизировать контроль линий, рабочие вынуждены самостоятельно проверять громкоговорители. Ввиду обычно большого количества динамиков и большой протяженности линий, сделать это не представляется возможным.

Оптимальным вариантом является автоматическое периодическое измерение импеданса линий с громкоговорителями. По данному электрическому параметру можно определить состояние измеряемых объектов. Для этого измеряют заведомо исправную линию и полученное значение становится контрольным для последующих измерений. При соответствующих отклонениях от нормы определяются ошибки линии.

Измерение импеданса линий громкоговорителей в системах аварийного оповещения обычно выполняется с использованием специализированного тестировочного оборудования, например, такого как анализатор импеданса. Разрабатываемый контроллер выполняет эту функцию. Существуют несколько схем измерения импеданса: мостовая, резонансная, I-V схема. Все они обладают своими преимуществами и недостатками. Выбор схемы преимущественно зависит от измеряемого объекта и необходимой точности. В устройстве уже выбрана схема автобалансировочного моста.

Импеданс включает в себя сопротивление, индуктивность и емкость измеряемой цепи и дает полное представление об ее электрических свойствах (рис. 1).

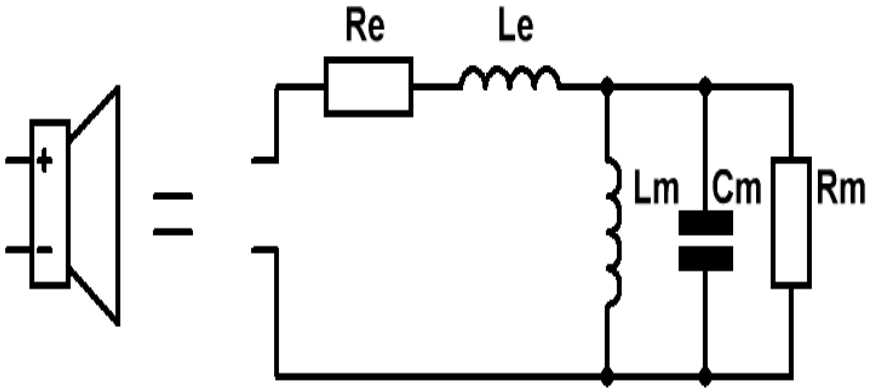


Рис. 1. Эквивалентная схема динамика

На рисунке 1, представленном выше изображена эквивалентная схема динамика. В его импеданс входят сопротивление и индуктивность звуковой катушки, а также дополнительные сопротивление, индуктивность и ёмкость, которые имитируют механические свойства громкоговорителя.

Также саму линию без громкоговорителей можно представить в виде распределенных по всей длине сопротивлению, ёмкости и индуктивности. Кабель для линии должен соответствовать подключаемой нагрузке и обеспечивать передачу сигнала с минимальными потерями. Распространенный вариант системы аварийного оповещения построен на двухпроводной линии, с мощностью до 340 Вт. Ему соответствуют параметры, представленные в таблице ниже.

Таблица

Соответствие длины кабеля с его диаметром для линии, не превышающей мощность более 340 Вт

Длина кабеля	Диаметр кабеля
менее 600 м	1 мм^2
от 600 до 1200 м	1,5 мм^2
более 1200 м	2,5 мм^2

Таким образом, при контроле линии громкоговорителей измеряется совокупность всех компонентов импеданса. Полное сопротивление линии с громкоговорителями определяется электрическими и механическими свойствами громкоговорителей, а также характеристиками линии передачи, соединяющей громкоговорители. Важно чтобы это значение было постоянным. В отличие от индуктивности и ёмкости, сопротивление не зависит от частоты. Однако все эти компоненты могут зависеть от других физических параметров, например температуры, влажности, давления. Это в свою очередь сказывается на результатах измерений.

Текущая версия устройства контроллера линии аварийного оповещения, подобно своим аналогам, в лабораторных условиях выдает стабильные результаты измерений линий громкоговорителей. Однако в реальных условиях из-за большой протяженности данные линии подвержены воздействию помех внешней среды. Прежде всего это перепады температуры, наличие объектов вблизи линий громкоговорителей, которые создают на них помехи. Это сказывается на точности измерений. На одну линию громкоговорителей может быть подключено несколько динамиков в параллель.

Измерение импеданса большого количества динамиков параллельно является сложной задачей, поскольку общий импеданс системы представляет собой суммой отдельных импедансов каждого из динамиков. В таком случае при одинаковых громкоговорителях общий импеданс на линии будет приближенно равен импедансу одного динамика, поделенного на количество динамиков в данной линии.

$$Z_{\text{линии}} \approx \frac{Z_{\text{динамика}}}{N_{\text{динамиков}}}$$

Импеданс большого количества громкоговорителей на одной линии является относительно маленькой величиной, которую трудно точно измерить. Кроме того, индивидуальные импедансы динамиков могут различаться, что затрудняет точное определение общего импеданса системы.

Данные проблемы можно отнести к критерию надежности системы, поскольку некорректные измерения влияют на работу не только системы аварийного оповещения, но и на безопасность предприятия, на котором она установлена.

Для преодоления этих проблем может потребоваться использование специализированного измерительного оборудования или другой метод измерения, например измерение импеданса каждого громкоговорителя по отдельности, а затем суммирование значений.

При измерении импеданса линии громкоговорителей в реальных условиях возникают ситуации, при которых непонятна причина, по которой могло произойти отклонение в измеренном значении импеданса. Причина в неисправности одного или нескольких громкоговорителей, либо же воздействие внешней среды на измеряемую линию. В таком случае возможны появления ложных ошибок. В существующих устройствах эту проблему решили ручным заданием порога отклонения импеданса от нормы. Также используют опцию отключения точного определения импеданса при

скачках температуры. Так ложные ошибки из-за перепада температур не будут появляться, но в то же время, контроллер не будет фиксировать ошибку на одном или нескольких громкоговорителях. Это в свою очередь не является разумным так, как производители оборудования не решают проблему полностью, а лишь избегают её.

В нашем исследовании рассматривались следующие реализации контроллера системы аварийного оповещения:

- LCU, Германия;
- SC-6224, Южная Корея;
- Sonar SSC-216M, Южная Корея;
- G&M ECU emergency control unit, Германия.

При измерении устройства позволяют детектировать нескольких типов ошибок: они представлены на шкале сопротивления на рисунках 2 и 3.

Рис. 2. Диапазоны импеданса линии громкоговорителей

При попадании одной линии на другую происходит короткое замыкание, что сигнализирует очень большое падение сопротивления, вплоть до нуля. Обратная ситуация происходит, когда сопротивление становится очень большим. Это может происходить из-за обрыва кабеля. При коротком замыкании линия для трансляции блокируется. При остальных ошибках трансляция возможна, поскольку главная задача – передать полезный сигнал. При быстро меняющихся погодных условиях возможны физические воздействия на линию громкоговорителей. Изменения температуры приводят к изменению сопротивления линии. В таких случаях можно отключать точное определение изменения импеданса и детектировать только короткое замыкание или обрыв линии. Это позволяет избежать определения ложных ошибок, но при этом снижается количество контролируемых громкоговорителей. На рисунке 3 ниже показаны диапазоны импеданса линии громкоговорителей с отключенным определением ошибки одного или нескольких громкоговорителей.

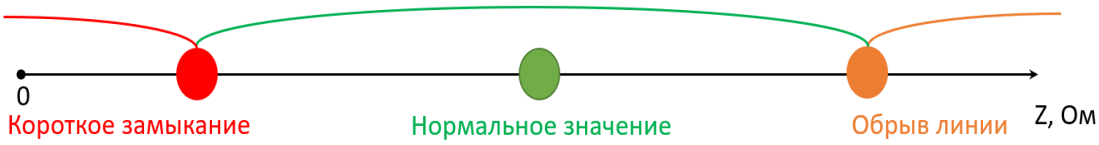


Рис. 3. Диапазоны импеданса линии громкоговорителей с отключенным определением ошибки одного или нескольких громкоговорителей

Были проведены измерения линии с тридцатью 6 Вт динамиками, включенными в параллель в текущей версии устройства. Результаты представлены на рисунках 4 и 5 ниже.

В линию каждый раз подключался дополнительный динамик и проводились измерения линии в новой конфигурации. После тринадцатого динамика для улучшения точности схема переводилась на шунтирующее сопротивление меньшего номинала, для получения большего разрешения и более точного результата.

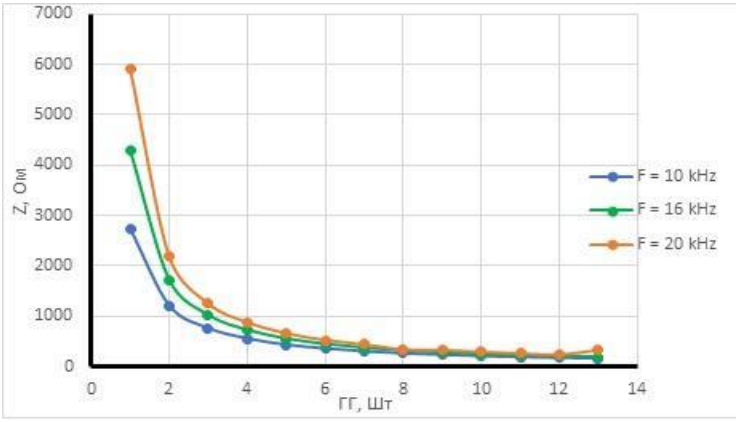


Рис. 4. Зависимость импеданса от количества динамиков на разных частотах (до 13 шт.)

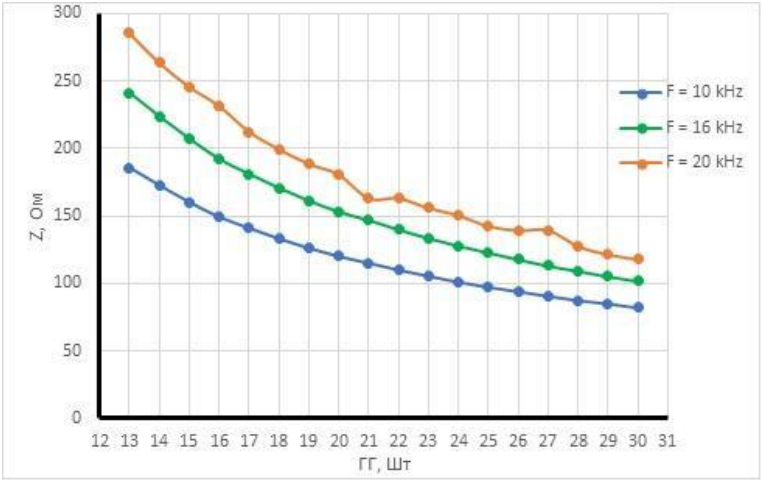


Рис. 5. Зависимость импеданса от количества динамиков на разных частотах (от 13 до 30 шт.)

По полученным данным видно, что система довольно точно определяет наличие 20 динамиков 6 Вт в лабораторных условиях. Однако в реальных условиях поскольку общий импеданс данной конфигурации является небольшим – 153 Ом, возникают трудности в измерении. (Сопротивление одного динамика 4300 Ом на 16 кГц)

Импеданс девятнадцати динамиков составляет – 161 Ом, а двадцати – 153 Ом. Разница между ними составляет – 8 Ом (5%). Чтобы детектировать аварию двадцатого динамика, необходимо, чтобы отклонение от нормы составило 5 % и более. Однако важно, чтобы погрешность измерений была меньше данной разницы. В лабораторных условиях это выполняется, в реальных – может нарушаться из-за влияния внешней среды. Возникает ситуация, при которой неясно чем было вызвано изменение в сопротивлении линии: аварией на одном из динамиков или внешним воздействием (помеха на линии или влияние температуры). В существующих устройствах на данный момент проблему ложного детектирования ошибки решают отключением точного детектирования ошибки на линии громкоговорителей, при этом система способна уже контролировать не двадцать, а десять динамиков. Это значительно снижает возможности контроля линии.

Научная проблема заключается в сложности автоматического мониторинга линии с более, чем двадцатью 6 Вт параллельно подключенными громкоговорителями при воздействии помех внешней среды. Мы планируем усовершенствовать встроенное программное обеспечение контроллера, чтобы осуществлять мониторинг, который позволит повысить максимальное количество измеряемых динамиков, что в свою очередь приведет к снижению поломок, в том числе, ложных срабатываний, повышению

надежности и отказоустойчивости, а также сокращению сроков сервисного обслуживания системы аварийного оповещения в целом.

В связи с вышесказанным, появляется потребность в улучшении способа измерений линий громкоговорителей, который позволит получать точные результаты большого количества динамиков в условиях, которые поддерживаются на предприятиях. Прежде всего проблема может быть решена программным путем. Для уменьшения влияния температуры необходимо ввести корректировочные коэффициенты, поскольку импеданс зависит от температуры. Это позволит корректировать меняющиеся результаты измерений совместно с изменяющимися условиями внешней среды и получать более достоверную информацию о линии с громкоговорителями. А также повысить точность измерений импеданса линии с громкоговорителями и снизить количество ложных ошибок может применение методов машинного обучения.

Машинное обучение позволит улучшить результаты измерения импеданса, на который влияют изменения температуры следующим образом:

- анализ данных: алгоритмы машинного обучения позволяют анализировать большие объемы данных об импедансе и температуре для выявления закономерностей и корреляций, которые могут быть неочевидны при обычной статистике;
- температурная компенсация: с помощью машинного обучения можно моделировать взаимосвязь между температурой и импедансом, а затем использовать для автоматической компенсации вызванных температурой изменений в измерениях импеданса в режиме реального времени;
- калибровка: алгоритмы машинного обучения могут быть использованы для автоматизации процесса калибровки, упрощая регулярную калибровку измерительной системы и учет любых изменений температурной чувствительности с течением времени;
- моделирование и прогнозирование: алгоритмы машинного обучения можно использовать для создания прогнозов того, как импеданс изменится в ответ на изменения температуры. Это позволит получать более точную информацию об измерениях импеданса.

Поскольку обработка данных происходит на микроконтроллере, являющегося частью встраиваемых систем важно учитывать ограниченную вычислительную мощность и объем памяти, а также необходимость обработки данных датчиков в режиме реального времени. Перечислим методы машинного обучения, которые хорошо подходят для встраиваемых систем опроса датчиков:

- дерево принятия решений. Данные алгоритмы хорошо подходят для встроенных систем из-за их низких вычислительных требований и простоты реализации. Они позволяют классифицировать данные с датчиков для их последующей обработки;
- метод случайного леса. Это продвинутый алгоритм дерева принятия решений, случайные леса могут обрабатывать более сложные отношения данных и обеспечивать более точные прогнозы;
- метод опорных векторов. Данные алгоритмы хорошо подходят для задач, в которых необходимо классифицировать по множеству признаков. Они способны обрабатывать большое количество функций и нелинейных отношений;
- обнаружение аномалий. Этот алгоритм может использоваться для обнаружения необычного поведения данных датчиков и хорошо подходит для нахождения аномальных отклонений.

Важно отметить, что для встраиваемых систем может потребоваться дополнительная оптимизация данных методов, такие как удаление избыточных частей сети и квантование, чтобы уменьшить вычислительную сложность алгоритмов.

В контроллере линий громкоговорителей методы машинного обучения могут быть использованы прежде всего для корректировки значения импеданса с учетом влияния факторов окружающей среды и обработки ложных ошибок, что позволит строить прогнозы и нарабатывать базу зависимостей различных факторов на результаты

измерения. Это позволит создать автоматическую адаптивную систему мониторинга линий громкоговорителей.

Заключение

В ходе данной работы были определены основные направления для повышения точности результатов измерения импеданса и надежности контроллера линий системы аварийного оповещения. Необходимо сделать автоматическую калибровку измерительной системы в зависимости от эксплуатационной температуры устройства, получать данные со вспомогательных датчиков. Полученную информацию анализировать и корректировать с помощью методов машинного обучения, тем самым приближая значения результатов измерений к реальным и прогнозируя возможные неисправности системы аварийного оповещения. Проведенные исследования, дают основание предположить, что для решения нашей задачи наиболее подходят метод случайного леса, либо нейросети прямого распространения.

Литература

1. Ghahramani A., Galicia P., Lehrer D., Varghese Z., Wang Z. and Pandit Y. Artificial Intelligence for Efficient Thermal Comfort Systems: Requirements, Current Applications and Future Directions // Frontiers in Built Environment. 2020.
2. N. Kathpalia, Gulati T. 3 Axis Gyro Accelerometer & Artificial Intelligence based Enhancement of GPS // Department of Electronics & Communication Engineering, Maharishi Markandeshwar Engineering College, MM (Deemed to be University), Mullana, Ambala, Haryana. India. 2022.
3. Ramesh A. Sai, Vigneshwar S., Vickram S. Artificial intelligence driven hydrogen and battery technologies // Fuel. 2023.
4. Hall Effect Sensor and Artificial Neural Networks Applied on Diagnosis of Broken Rotor Bars in Large Induction Motors, Dias C.G, Chabu I.E., Bussab M.A. // Computer Science, Engineering 2006 IEEE International Conference on Computational Intelligence for Measurement Systems and Applications. 2006.
5. Terzic E., Nagarajah C.R., Alamgir M. Capacitive sensor-based fluid level measurement in adynamic environment using neural network // Computer Science. 2010.

УДК 004.07

ПРОБЛЕМЫ БЕЗОПАСНОСТИ MESH-СЕТЕЙ БЕСПИЛОТНЫХ ЛЕТАТЕЛЬНЫХ АППАРАТОВ

Блинов А.В.[1] (аспирант)
Научный руководитель – д.т.н., доцент Беззатеев С.В.[1]
1 – Университет ИТМО
a.blin0v@mail.ru; bsv@aanet.ru

Аннотация
Беспилотные летательные аппараты (далее - БПЛА) все чаще появляются в роях или формациях для использования кооперативного поведения. Эти сети с поддержкой БПЛА могут удовлетворять нескольким сложным требованиям миссии и рассматриваются как потенциальный помощник для многих возникающих случаев использования в будущих сетях связи. Такие сети, однако, характеризуются высокой динамичностью и мобильностью без гарантии наличия центральной сетевой инфраструктуры, что может вызвать проблемы с подключением и безопасностью. Хотя беспроводные Mesh-сети рассматриваются как решение для таких сценариев, эти сети имеют свои собственные проблемы и уязвимости безопасности. В этой статье мы анализируем ключевые проблемы безопасности, возникающие при применении беспроводных Mesh-сетей в роях БПЛА.
В частности, мы выделяем основные проблемы применения современных Mesh технологий в области роев БПЛА и раскрываем существующие уязвимости во всем коммуникационном стеке. На основе этого анализа мы представляем архитектуру, ориентированную на безопасность для Mesh коммуникаций БПЛА.
Ключевые слова
БПЛА, безопасность, Mesh-сеть, шифрование, AAA.

Точное определение местоположения объектов, картирование в реальном времени неизвестных или труднодоступных мест, а также сокращение времени выполнения задач по сравнению с одиночными устройствами делают рой БПЛА привлекательным для коммерческого и военного использования. Кроме того, рои БПЛА могут способствовать удовлетворению требований сетей будущего, обеспечивая высокую доступность, большую плотность соединений и низкую сквозную задержку. Такие сценарии, однако, создают значительные проблемы в виде мобильности устройств, формирования динамических сетей и ограниченной энергии устройств, которые должны сочетаться с требованиями самоорганизации, надежности за счет избыточности и координации задач, предъявляемыми к роям БПЛА.

Существующие технологии беспроводных сетей имеют потенциал для решения этих проблем, и оба стандарта IEEE 802.11 [1] и стандарт IEEE 802.15.4 [2] популяризировали Mesh протоколы для бытового и промышленного использования.

Многие Mesh протоколы разработаны без обеспечения безопасности или предполагают, что вопросы безопасности решаются на более высоких уровнях. Это предположение, а также отсутствие центрального контроллера, делают протокольные уровни уязвимыми для нескольких угроз. Предположим, что цель злоумышленника - нарушить связь между БПЛА в рое. Таким образом, не имея доступа к сети, внешний злоумышленник может контролировать или нарушать связь внутри сети. С другой стороны, атака может быть внутренней - как только злоумышленник компрометирует надежные узлы или учетные данные сети посредством криптоанализа или физической атаки.

Широковещательная природа беспроводного канала делает передачи на физическом уровне восприимчивыми к различным атакам и злонамеренным действиям. В целом, угрозы на физическом уровне могут быть классифицированы как пассивные

или активные атаки. Пассивные атаки в основном классифицируются как подслушивание и анализ трафика, в то время как активные атаки в основном включают в себя глушение и помехи. Все атаки, представленные на физическом уровне, считаются внешними по отношению к рою БПЛА.

Канальный уровень отвечает за мультиплексирование потоков данных, контроль ошибок и управление доступом к среде. Основная уязвимость на этом уровне заключается в подмене MAC-адреса, когда одна станция выдает себя за другого члена сети. Другие уязвимости на уровне кадров включают в себя сниффинг и широковещательные штормы.

Сетевой уровень отвечает за инкапсуляцию IP-пакетов и маршрутизацию информации по сети. В частности, в беспроводных Mesh-сетях с помощью многоходовых алгоритмов маршрутизации. Атаки на этот уровень направлены на механизм маршрутизации механизм связи роя БПЛА и влияют на пересылку пакетов.

Транспортный уровень: уровень управляет потоком данных в коммуникационном стеке. Уязвимости этого уровня связаны с повреждением пакетов или использованием слабых мест в протоколах. Протокол транспортного уровня (TCP) (ориентированный на соединение), протокол пользовательских дейтаграмм (UDP) (без соединения) и протокол управляющих сообщений Интернета (ICMP) являются основными протоколами транспортного уровня.

Операционная система робота (ROS) – это фреймворк, широко используемый для программного обеспечения роботов, как пример прикладного уровня в роях БПЛА. Поскольку ROS была разработана в основном для исследовательских целей, при ее реализации не было добавлено никаких функций безопасности. Отсутствие таких функций безопасности, как шифрование данных и аутентификация, делает ROS критической поверхностью для различных векторов атак в сетях БПЛА. ROS2 [3] построен на базе службы распределения данных (DDS) с архитектурой Real-Time Publish Subscribe (RTPS) [4] и был разработан для того, чтобы устранить уязвимости, обнаруженные в его предыдущей версии - добавление аутентификацию, шифрование и функции профиля процесса, которые полагаются на инфраструктуру открытых ключей.

Чтобы защитить сеть, ключевым инструментом является шифрование. Шифрование позволяет защитить данные, передаваемые между БПЛА, от несанкционированного доступа и прослушивания.

Одним из распространенных методов шифрования является AES [5], который использует ключи длиной 128, 192 или 256 бит, чтобы зашифровать и расшифровать данные.

AES используется для шифрования информации в Mesh-сети БПЛА таким образом:

1. Генерация ключа AES: перед началом шифрования ключ AES должен быть сгенерирован и доставлен на все аппараты, которые будут использовать эту сеть.

2. Шифрование данных: когда данные готовы к передаче по Mesh-сети, они шифруются с использованием ключа AES.

3. Передача данных: шифрованные данные передаются по Mesh-сети между аппаратами.

4. Дешифрование данных.

Другой распространенный метод шифрования – RSA [6]. RSA является криптографическим алгоритмом ключа публикации. В Mesh-сети БПЛА, RSA может использоваться для шифрования информации во время передачи данных между аппаратами в сети.

В этом случае, каждый БПЛА имеет уникальную пару ключей RSA, состоящую из открытого и закрытого ключа. Открытый ключ используется для шифрования данных, которые передаются между аппаратами, а закрытый ключ используется для дешифрования полученных данных.

Защита закрытого ключа RSA на БПЛА может быть осуществлена с использованием различных методов, таких как хранение в зашифрованном виде, использование дополнительных протоколов аутентификации и авторизации, а также использование различных алгоритмов криптографической защиты. Важно обеспечить, чтобы закрытый ключ RSA находился в надежном месте, и никто, кроме авторизованных пользователей, не имел доступа к нему.

Кроме того, для дополнительной защиты можно использовать мультиплексирование шифрования. Мультиплексирование шифрование – это технология, которая позволяет одновременно передавать несколько зашифрованных потоков данных по одной канальной линии. В контексте Mesh-сетей для БПЛА, это означает, что несколько БПЛА могут одновременно передавать зашифрованные данные друг другу через общую Mesh-сеть. Это позволяет улучшить эффективность и безопасность передачи данных в сети.

В случае если необходимо добавить новый БПЛА в Mesh-сеть, сначала происходит идентификация и аутентификация устройства.

Для идентификации новых БПЛА в Mesh-сети используется уникальный идентификатор, который может быть назначен администратором сети.

Процесс идентификации новых БПЛА в Mesh-сети включает в себя следующие этапы:

1. Обнаружение: каждый БПЛА в Mesh-сети постоянно пытается обнаружить другие БПЛА в окружении.

2. Обмен идентификационными данными: когда два БПЛА обнаружили друг друга, они начинают обмен идентификационными данными, которые могут включать в себя имя, тип БПЛА, версию ПО и т.д.

3. Валидация: когда БПЛА получили идентификационные данные друг друга, они проверяют, являются ли данные действительными и достоверными.

4. Утверждение: если данные другого БПЛА прошли валидацию, они подтверждают его идентификацию и добавляют его в свой список известных устройств.

После идентификации БПЛА производится аутентификация. Аутентификация в Mesh-сетях БПЛА – это процесс, который проверяет, является ли БПЛА, который пытается подключиться к сети, допустимым участником. Это важный шаг, чтобы предотвратить неавторизованный доступ к сети.

Аутентификация может происходить с использованием пары ключей, таких как RSA, или посредством проверки подлинности идентификаторов, таких как имена пользователей и пароли. В зависимости от сложности сети и критичности информации, которую она передает, рекомендуется использовать надежные методы аутентификации.

Рои БПЛА используют беспроводную сеть Mesh для обеспечения инфраструктуры сетей Ad-hoc, как в простых приложениях, таких как географическое картирование, так и в сложных событиях, таких как климатические катастрофы или военные миссии. Однако недостатки, присущие протоколам и стандартам Mesh, привели к появлению новых поверхностей для атак и возникновению новых проблем безопасности. В этой статье мы выделили основные уязвимости во всем коммуникационном стеке сетки БПЛА - от физического до прикладного уровня. На основе этой информации мы предложили архитектуру безопасной Mesh-сети.

Литература

1. Педжман Р., Джонатан Л. Основы построения беспроводных локальных сетей стандарта 802.11: практическое руководство по изучению, разработки и использованию беспроводных ЛВС стандарта 802.11 // Вильямс. 2004. ISBN 5-8459-0701-2.
2. Стандарт IEEE 802.15.4. [Электронный ресурс]. Режим доступа: URL: https:// standards.ieee.org/ieee/802.15.4/7029/.

3. [Электронный реурс]. Режим доступа. URL: https://docs.ros.org/en/foxy/index.html.
4. Стандарт Real-Time Publish Subscribe. [Электронный ресурс]. Режим доступа. URL: https://www.omg.org/spec/DDSI-RTPS/2.3/PDF.
5. Стандарт AES. [Электроннйы ресурс]. Режим доступа. URL: https:// nvlpubs.nist. gov/nistpubs/FIPS/NIST.FIPS.197.pdf.
6. Хеллман М., Диффи У. Новые направления в криптографии. [Электронный ресурс]. Режим доступа. URL: https://www-ee. stanford.edu/~hellman/publications/24.pdf.

УДК 004.9

MULTI-TASK DEEP LEARNING-BASED INTRUSION DETECTION MODEL FOR IOT NETWORK

Huiyao Dong[1] (postgraduate)
Supervisor – PhD Kotenko I.V.[1]
1 – Университет ИТМО
hydong@itmo.ru; ivkote1@mail.ru

Abstract
With the growing importance of IoT networks in modern society, safeguarding the network and intra-network devices is critical. This study presents the design of a multi-task (MTL) deep learning-based intrusion detection system, and tests two MTL architectures, namely soft-parameter sharing and hard-parameter sharing. The experiment demonstrates that MTL approaches not only excel in identifying rare attacks when compared to normal STL, but also achieve a balance between detection and false alarms.
Keywords
Intrusion detection, cyberattacks, Internet of Things, machine learning, deep learning.

Introduction

The Internet of Things (IoT) connects physical and virtual entities through emerging information and communication technologies. This enables network properties to automatically analyse specific scenarios or surroundings, which benefits all network environments with the use of electronics, software, sensors, actuators, and internet connectivity [1]. As it is a current technological trend, IoT is applied in self-driving cars, smart healthcare, and smart cities. However, IoT vulnerabilities and intrusion threats can cause significant harm to devices and the system, with network intrusion being one of the most common threats. Although there is a plethora of research on Deep Learning-based intrusion detection systems (IDS) for the IoT, most existing studies used Single-task Learning (STL). An STL model predicts one task, binary classification or multiple-label classification, per model at a time. In cases of a shortage of label data or the need to predict multiple tasks, multi-task learning (MTL) becomes an option. MTL uses the same feature set to predict multiple tasks simultaneously with one model, resulting in faster convergence, improved training efficiency and a lower risk of overfitting due to shared representations between tasks. his paper proposes an MTL DNN-based intrusion detection method for IoT network data. This method trains each task through its own subnetwork and transfers knowledge between tasks based on their relationship. We compare this approach with a DNN-based STL model and a DNN-based hard-parameter sharing MTL, and the results show that it is more effective in detecting rare attacks.

Methodology

MTL model topologies can be classified as either hard parameter sharing or soft parameter sharing, as shown in Fig. 1. The hard parameter sharing model involves shared layers, and all tasks are learned simultaneously. On the other hand, the soft parameter sharing model constructs a sub-network for each task with an inter-model parameter sharing mechanism for knowledge transfer and task relationships. Although the hard parameter sharing model is easier to implement, the soft parameter sharing model facilitates complete and flexible intra-model information transfer, and is generally more effective when activities are unrelated. In the case of network intrusion detection, conflicting tasks occur because each sample can only be classified as one type of intrusion. Therefore, soft parameter sharing is more efficient than hard parameter sharing at minimizing loss.

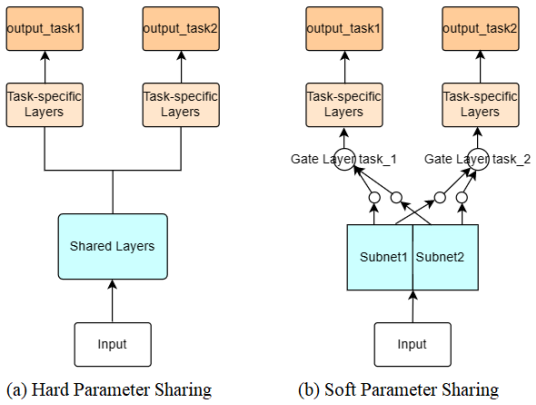


Fig. 1. Architectures of MTL

With shared-layer denoted as $f(x)$ and sub-network h^t for T $(t \in T)$ task-specific layers, the output of single task in hard-parameter sharing can be written as:

$$y_t = h^t(f(x)). \tag{1}$$

The soft parameter sharing architecture was inspired by the work of MOE [2], which was shown to have good results on even less related tasks. In soft-parameter sharing, there is a gate network g^t to compute the knowledge transfer. With gate networks implemented for each sub-network, the output can be written as:

$$y_t = h^t\left(\sum_{i=1}^{T} g^t(x)_i f_i(x)\right). \tag{2}$$

In which the gate layer performs SoftMax on the multiplication of input x and its corresponding trainable parameter matrix:

$$g^t(x) = Softmax(W_{gt}x) = \frac{W_{gt}x}{\sum_{n=1}^{N} W_{gt}x_n}. \tag{3}$$

This method implements deep fully connected layers with a consistent output unit of 128 in each sub-network. Regularizers are also applied to penalize the layer's parameters and the activation function during the optimisation process. The norm functions Ln used as the regularizer can be written as (4).

$$||W||_n = (|w_1|^n + |w_2|^n + \cdots + |w_k|^n)^{\frac{1}{n}}. \tag{4}$$

If use notation (5) to represent the output of a hidden layer, L1, L2-norm is applied to the weights of layer w, L1-norm is applied to the bias vector b, and L2-norm is applied to the output vector y.

$$y = f(wx + b). \tag{5}$$

For the three regularizers, the corresponding objections of optimization are as follows. λ are trainable parameters with positive value.

$$Obj_{l1-norm} = error(wx + b, y) + \lambda|w|, \lambda > 0, \tag{6}$$

$$Obj_{l2-norm} = error(wx + b, y) + \lambda w^2, \lambda > 0, \quad (7)$$

$$Objl_{l1,l2-norm} = error(wx + b, y) + \lambda_1\lambda_2\|w\| + \frac{\lambda_1(1-\lambda_2)}{2}\|w\|. \quad (8)$$

Once each subnetwork has completed the learning process, the knowledge sharing process begins, facilitated by the SoftMax gate layer and a dropout layer. The SoftMax gate layer computes the distinction between tasks and assigns sharing parameters accordingly. If another task has low similarity or sparse labels, the gate layer will prioritize the information provided by the subnetwork originally assigned to it. The transfer knowledge will be assigned a probability based on this process. Next, a dropout layer with a 10% rate is applied, followed by renormalization of the remaining data. This remaining data represents the utilization probabilities for task-specific layers to learn from.

Experiment

To compare STL and MTL and demonstrate the differences between the two basic architectures of MTL models, we constructed a DNN-based STL model and a hard parameter-sharing MTL model. All three models used Dense, the deep fully connected layer in Keras implementation, with an identical output unit of 128 and depth. They were trained with 200 epochs and batch sizes of 64.

All three models are tested on UNSW-NB15 [3], which is a public dataset for IoT network detection techniques. It covers normal behaviours and nine attacks. The corresponding share in training and testing datasets is listed below: Normal behaviour (44.9%|31.9%), Backdoor (4.2%|1.0%), Analysis (0.7%|1.1%), Fuzzers (5.0%|10.4%), Shellcode (13.5%|0.6%), Reconnaissance (0.8%|6.0%), Exploits (7.4%|9.0%), DoS (0.1%|7.0%), Worms (0.5%|0.1%), and Generic (22.9%|22.8%). It is worth noting that normal behaviour and generics represent the vast majority, while analysis, backdoors, shellcode, and worms occur significantly less frequently. The feature selection is performed through relatedness with the Elastic Net Regularization method. Resampling is done using the SVM-based Synthetic Minority Oversampling Technique (SMOTE) and Edited Nearest Neighbours (ENN). Following data pre-processing, we test all three models.

The overall accuracy rates for this experiment are as follows: STL-57%, hard parameter sharing MTL-93.58%, and soft parameter sharing MTL-93.62%. The detection performance for different traffic intrusions is listed in Table and attacks are ranked based on their occurrence frequency. As MTL models predict all intrusions with binary classification, false alarm rates are also comparable between these models. The results show that the STL model performs well on most attacks, but falls short of both MTL models in detecting rare attacks such as reconnaissance, analysis, backdoors, and worms. When comparing the two MTL models, which have identical DNN layer settings, the soft parameter sharing model has better detection performance overall. Still, the hard parameter sharing model performs better in detecting worms. Moreover, high detection rates are usually associated with high false alarm rates, but for IDS, low false alarm rates are crucial. Both MTL models presented acceptable false alarm rates of less than 10%. For example, the soft parameter sharing model had an outstanding high detection rate of 98.7% for generic, with only 0.1% of false alarms and the highest detection rate for backdoor with only 0.01% of false alarms. Overall, both MTL models outperformed the STL model, and their false alarms were acceptable. Furthermore, the soft parameter sharing MTL model performs better for anomaly detection when limited corresponding samples are available.

Table

Detection performance on UNSW-NB15 Dataset. Evaluation Metrics of different models

		Normal	Generic	Exploits	Fuzzers	DoS
STL	precision	0.970	0.990	0.370	0.250	0.330
	recall	0.620	0.980	0.130	0.360	0.680
	f1-score	0.756	0.985	0.192	0.295	0.444
H-MTL	precision	0.973	0.698	0.158	0.955	0.725
	recall	0.730	0.828	0.188	0.032	0.082
	f1-score	0.834	0.757	0.171	0.062	0.148
	FPR	0.000	0.083	0.000	0.000	0.003
S-MTL	precision	0.499	0.978	0.479	0.312	0.254
	recall	0.935	0.996	0.410	0.367	0.250
	f1-score	0.651	0.987	0.442	0.337	0.252
	FPR	0.128	0.001	0.012	0.004	0.005
		Reconnaissance	Analysis	Backdoor	Shellcode	Worms
STL	precision	0.210	0.040	0.080	0.070	0.000
	recall	0.530	0.010	0.020	0.360	0.060
	f1-score	0.301	0.016	0.032	0.117	0.000
H-MTL	precision	0.308	0.963	0.854	0.078	0.739
	recall	0.000	0.024	0.088	0.009	0.126
	f1-score	0.000	0.047	0.159	0.016	0.215
	FPR	0.0001	0.000	0.0002	0.025	0.005
S-MTL	precision	0.333	0.054	0.131	0.259	0.005
	recall	0.967	0.637	0.444	0.321	0.062
	f1-score	0.495	0.100	0.202	0.287	0.008
	FPR	0.0007	0.013	0.0001	0.008	0.010

Summary

We propose a hybrid intrusion detection technique based on Multi-Task Learning (MTL), which predicts traffic class labels as separate binary classification tasks. Both hard-parameter and soft-parameter sharing architectures of MTL were tested, and the experimental results demonstrate that they consistently achieve balanced performance on all traffic types, particularly when labelled data are limited. Moreover, as it balances the detection rate and false alarm traffic classification, our method would be highly practical in the field of real-world network traffic detection. To further investigate the proposed MTL-based approach, we plan to construct a detection model using different types of neural network structures such as Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), and Gated Recurrent Units (GRUs), and to design an MTL-based loss optimization to enhance the detection performance.

Reference

1. Denisenko Ruiz J.F., Harjani R., Maña A., Desnitsky V., Kotenko I., Chechulin A. A methodology for the analysis and modeling of security threats and attacks for systems of embedded components // The 20th Euromicro International Conference on Parallel, Distributed and Network-Based Processing. 2012. pp. 261–268.
2. Ma J., Zhao Z., Yi X., Chen J., Hong L., Chi E. Modeling Task Relationships in Multi-task Learning with Multi-gate Mixture-of-Experts // Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining (KDD '18). 2018. pp. 1930–1939.

3. Moustafa N., Slay J. UNSW-NB15: a comprehensive data set for network intrusion detection systems. // 2015 Military communications and information systems conference. 2015. pp. 1-6.

УДК 004.056

ОСОБЕННОСТИ ОБЕСПЕЧЕНИЯ БЕЗОПАСНОСТИ ПОЛЬЗОВАТЕЛЬСКИХ ДАННЫХ WEB-СЕРВИСОВ

Ишутина Е.[1] (студент**)**
Научный руководитель - к.т.н., доцент Чернов А.И.[1]
1 – Федеральное государственное бюджетное образовательное учреждение высшего образования «Санкт-Петербургский государственный морской технический университет»
e-mail: liz16z2001@mail.ru

Аннотация
Утечка пользовательских данных является одной из актуальных угроз безопасности веб-сервисов. В данной статье выполнен обзор таких видов уязвимостей веб-сервисов, как XXS и SQLi. Одним из эффективных методов защиты от данных угроз является применение параметризованных запросов в SQL. Этот способ позволяет различать код и данные независимо от того, что вводится пользователем.

Ключевые слова
Пользовательские данные, безопасность, уязвимость, SQL-инъекция, XSS-атака, запросы.

В настоящее время в целях безопасности пользовательских данных, значительное внимание уделяется обнаружению уязвимостей веб-сервисов. Наиболее распространенными эксплойтами являются XSS (Cross-Site Scripting – «межсайтовый скриптинг») и SQLi (Structured Query Language – «язык структурированных запросов» – инъекция), входящие в список 10 наиболее распространенных уязвимостей веб-приложений [1]. Эти атаки используют «доверие» между пользователем и сервером, а также недостаточные проверки ввода/вывода на уровне сервера. Уязвимости этих эксплойтов могут привести к удаленному управлению веб-приложениями или незаконной передаче пользовательских данных. В данной статье приведен обзор способов обеспечения безопасности пользовательских данных и надлежащей проверке уязвимостей.

XSS – это вид атак на web-приложения, при котором злоумышленник внедряет в страницу сайта вредоносный код, после чего может получить доступ к персональным данным пользователей, взломать учетные записи и т.д. XSS включает в себя фрагменты JavaScript для внедрения, позволяя злоумышленнику выполнять скрипт в веб-браузере целевого пользователя. Веб-приложения подвергаются риску через XSS, создавая зловредный фрагмент кода JavaScript, который вставляется внутри приложения. Внедрение скрипта происходит таким образом, чтобы выглядеть безопасным для «неподозрительного» пользователя. Наконец, доверенный домен веб-сайта может выполнять скрипт, содержащий вредоносный код. Предполагается, что в пост будут включены только тексты. Однако злоумышленник сможет внедрить скрипт типа “<скрипт>оповещение (‘XSS вредоносный’) </скрипт>” в качестве входа. Тем самым, целевой пользователь «нажимает» на зараженную ссылку, а внедренный JavaScript выполняется автоматически, что приводит к эксплойту XSS, влияющему на работу пользователя в Интернете [2].

Атаки XSS, как известно, бывают трех типов: сохраненные, отраженные и основанные на DOM. В хранимом XSS вредоносный скрипт внедряется на страницу или форму сайта, который хранит данный скрипт в своей базе данных. Рефлектированная XSS имеет место, когда веб-сервис передает недопустимые пользовательские данные. Злоумышленник отправляет вредоносный код в операторе URL-адреса под видом пользовательского ввода.

Эксплойты XSS на основе DOM считаются заменой отраженных или сохраненных кодов XSS. В таких атаках вредоносные фрагменты JavaScript не выполняются в целевом браузере без первоначального выполнения безвредной части кода JavaScript в браузере пользователя [2].

SQL – основной язык для работы с серверами баз данных, такими как Oracle, MSSQL и MySQL, и используется в программировании на многих языках, включая ASP, Java, .NET и PHP. Однако, если разработчики не проводят проверку входных операторов SQL, это может привести к ошибкам в кодировании и открытию уязвимостей для SQL-инъекций. SQL-инъекцию можно разделить на 3 основных класса: внутриполосная, внеполосная, инференциальная.

Внутриполосная SQLi – «классическая», эксплуатируется вследствие того, что используется тот же канал связи, который используется для введения кода SQL. Эксплойт внеполосной SQLi используется тогда, как злоумышленник не может извлечь данные из канала, который уже используется для введения кода SQL, он эксплуатирует уязвимость через другой канал связи.

При использовании инференциальной SQLi, данные не проходят через веб-приложение, злоумышленник не видит результатов атаки в группе, поэтому ее также называют «слепая» SQLi. Например, динамический SQL-запрос, как показано на рисунке 1, может стать уязвимым «классической» SQLi, если не были проведены необходимые проверки [2].

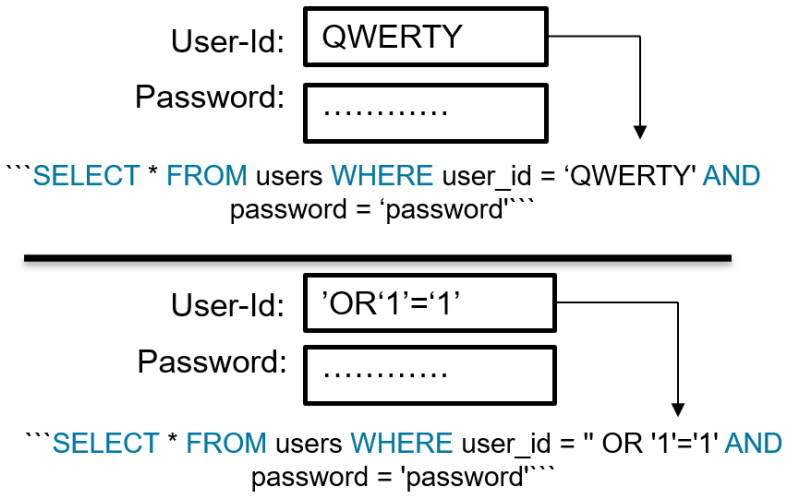


Рис. 1. Эксплуатация SQLi

Злоумышленник, скорее всего, воспользуется таким уязвимым запросом, предоставив значение «1» = «1», что является истинным для всех записей и позволит получить незаконный доступ к данным пользователей.

Анализ атак с нанесением ущерба на сохранность персональных данных позволяет утверждать, что наибольшую угрозу представляют вышеперечисленные атаки.

Существуют различные методы нейтрализации XSS-атак в зависимости от языка и фреймворка. Основным методом в PHP является метод htmlspecialchars, которому передается текст, который нужно обезопасить от возможных «проблем» [3].

Цель этого метода – найти в заданной строке символы, имеющие специальное значение в HTML, и заменить их безопасными версиями, которые преобразуются в обычный текст. К примеру, символ < заменяется в &1t; а символ > становится символом &qt;. Если в коде используется символы < и > , то они открывают и закрывают HTML-теги, но при использовании закодированных эквивалентов &1t и &qt;, передается, что нужно отобразить открывающую и закрывающую угловые скобки.

Использование подготовленных операторов со связкой переменных – параметризованных запросов в SQL [4]. Это запрос, в котором значения параметров задаются во время выполнения запроса, а не во время написания запроса. Это обеспечивает безопасность и гибкость запросов, так как параметры могут быть изменены без необходимости изменения самого запроса. На рисунке 2 рассмотрен пример параметризованного запроса.

```
1 – SELECT * FROM employees WHERE salary >?
2 – DECLARE @salary INT = 50000
3 – SELECT * FROM employees WHERE salary > @salary
```

Рис. 2. Применение параметризованного запроса

В первом запросе знак вопроса является параметром, который будет заменен на фактическое значение при выполнении запроса. К примеру, если следует выбрать всех сотрудников, у которых зарплата больше 50000, то выполняется запрос 2 и 3. В последнем (@salary) является параметром, который будет заменен на 50000 при выполнении запроса.

Данный способ позволяет базе данных различать код и данные независимо от того, что вводится пользователем. Для обхода аутентификации злоумышленник может использовать конструкцию, которая была рассмотрена на рисунке 1, но при использовании параметризованных запросов, конструкция сначала передается в виде запроса в базу данных, а после в него подставляются данные из переменных. Таким образом, за имя пользователя будет приниматься вся строка, введенная им, что не позволит злоумышленнику проэксплуатировать SQL-инъекцию. Данный метод считается наиболее эффективным методом защиты от такого рода атак и рекомендуемым подходом при работе с базами данных [4].

Экранирование входных данных. Суть данного метода – избежать пользовательский ввод, прежде чем поместить его в запрос. При содержании в параметре одинарной кавычки (‘), ее необходимо заменять двумя одинарными кавычками (‘’). В PHP параметры обычно экранируются с помощью функции mysqli_real_escape_string (); перед отправкой запроса SQL. Функция добавит косую черту к специальным символам в запросе, для предотвращения интерпретации как команды SQL.

Хранимые процедуры. Путем предварительной обработки входных данных и использования параметров, правильно выстроенные хранимые процедуры защищают от SQLi. Вместо вставки значений напрямую в SQL-запрос, параметры передаются в хранимую процедуру, которая затем обрабатывает эти параметры и выполняет соответствующий запрос. Если хранимые процедуры создаются с использованием параметров с типом данных, которые соответствуют типу данных входных параметров, то в случае содержания некорректных символов во входных данных запрос не будет выполнен, а пользователь получит сообщение об ошибке.

Использование WAF (Web Application Firewall). Работает путем анализа и фильтрации входящих запросов на наличие подозрительных символов и синтаксических конструкций, которые могут быть использованы для проведения атаки SQL-инъекций. Файрвол использует методы для обнаружения SQLi – проверка наличия запрещенных символов в запросе, таких как одинарные кавычки, двойные кавычки, точки с запятой и т.д., анализ синтаксиса запроса на наличие подозрительных конструкций, таких как UNION, SELECT, WHERE и т.д., использование списков блокированных слов и выражений для фильтрации запросов, мониторинг логов и обнаружение поведенческих аномалий, которые могут указывать на атаку SQL-инъекций.

В результате данного обзора выявлено, что уязвимости XSS и SQLi приводят к высоким рискам, связанных с кражей пользовательских данных. Рассмотренные методы защиты показывают важность фильтрации и проверки пользовательских данных.

Выявлено, что особенностью обеспечения безопасности пользовательских данных является метод параметризованных запросов и использование закодированных эквивалентов символов. На данном этапе работы были изучены основные методы обеспечения безопасности веб-сервисов. В дальнейшей работе планируется изучение обеспечения безопасности сетевых приложений на прикладном уровне. Спрос на WAF увеличивается и, соответственно, повышается необходимость его правильного функционирования, позволяющая защитить приложения наиболее эффективным образом [5].

Литература

1. OWASP Top Ten [Электронный ресурс]. Режим доступа: https://owasp.org/www-project-top-ten/ (дата обращения: 10.04.2023).
2. Abdalla W.M., Zarul F.Z., Khaled S. An Enhanced Static Taint Analysis Approach to Detect Input Validation Vulnerability // Journal of King Saud University - Computer and Information Sciences. 2023. V. 35. №5. P. 682–701.
3. Фленов М.Е. Web-сервер глазами хакера: в 3-е изд. СПБ.: БХВ-Петербург. 2022. 256 с.
4. Методы защиты от sql-инъекций [Электронный ресурс]. Режим доступа: https://htmlacademy.ru/blog/php/sql-injections (дата обращения: 11.04.2023).
5. Новости цифровой трансформации, телекоммуникаций, вещания и ИТ [Электронный ресурс]. Режим доступа: https://www.comnews.ru/content/224832/2023-03-17/2023-w11/spros-waf-rossii-kratno-vyros (дата обращения: 12.04.2023).

УДК 004.056.53

О ПОСТРОЕНИИ И ПРИМЕНЕНИИ КРИПТОГРАФИЧЕСКИХ ХЭШ-ФУНКЦИЙ, ОСНОВАННЫХ НА ГРАФАХ

Кирьянова А.П.[1] (магистрант), **Давыдов В.В.**[1] (преподаватель)
Научный руководитель – преподаватель Давыдов В.В.[1]
1 – Университет ИТМО
e-mail: anastaciakosanovskaya@gmail.com, vvdavydov@itmo.ru

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

Аннотация
В работе рассматриваются хэш-функции, основанные на расширяющих графах. Анализируется стойкость к коллизиям рассмотренных хэш-функций, делаются выводы о возможности и целесообразности использования таких хэш-функций в современных криптографических алгоритмах и протоколах. Изучается возможность применения таких хэш-функций для построения новых криптографических протоколов.
Ключевые слова
Хэш-функция, расширяющий граф, коллизия, криптографический алгоритм, дистанционно-ограниченные протоколы.

Криптографические хэш-функции – важный математический примитив, который на сегодняшний используется во многих современных системах. Условно их можно разделить на два типа: хэш-функции, основанные на специальных конструкциях (например, на конструкции Меркла-Дамгаара), и хэш-функции, основанные на некоторых трудных математических задачах. Конструкция функций первого типа близка к конструкции блочных шифров; такие хэш-функции хорошо изучены, однако, существует множество примеров, когда они оказывались неустойчивыми к коллизиям и подвергались криптоанализу, так как никаких формальных доказательств их стойкости нет. Например, функции MD5 и SHA-1 более не являются безопасными. Второй тип хэш-функций называется доказуемо безопасными и представляет определённый интерес, так как стойкость к коллизиям функций такого типа основывается на сложности решения различных трудных математических задач. Хэш-функции, основанные на расширяющих графах, относятся ко второму типу; стойкость к коллизиям зависит от свойств графа, например, от обхвата или константы расширения. Структура таких хэш-функций отличается от функций первого типа и выглядит компактно и элегантно, что позволяет рассмотреть варианты их применения при построении новых криптографических алгоритмов и протоколов. Данная работа носит обзорно-аналитический характер. В работе кратко рассматриваются важные идеи, лежащие в основе некоторых хэш-функций, построенных на графах, а также описываются возможные варианты их применения при конструировании современных алгоритмов и протоколов.

Хэш-функция – это математическая функция свёртки битовой строки произвольной длины в строку фиксированного размера n бит:

$$\mathrm{h}: \{0,1\}^* \to \{0,1\}^{\mathrm{n}}$$

Для того, чтобы хэш-функция считалась криптографической, она должна обладать следующими свойствами:

1. Необратимость. По полученному значению хэш-функции $\mathrm{H} = \mathrm{h}(m)$ вычислительно сложно определить m.

2. Стойкость к коллизиям. При заданном сообщении m_1 вычислительно сложно найти сообщение m_2 такое, что $\mathrm{h}(m_1) = \mathrm{h}(m_2)$.

3. Стойкость к атаке поиска второго прообраза. При заданном значении сообщения m_1 и $\mathrm{h}(m_1)$ вычислительно сложно найти другое сообщение m_2 такое, что $\mathrm{h}(m_1) = \mathrm{h}(m_2)$.

Одной из первых хэш-функций, построенных на расширяющих графах, является функция Zémor-Tillich (ZT), представленная в 1994 году на конференции CRYPTO [1]. Кратко опишем идеи, лежащие в основе данной хэш-функции:

- выбирается конечное поле $\mathbb{F}_{2^n}$ с неприводимым многочленом $\mathrm{P_n}(\alpha)$ степени n, α – примитивный элемент поля;
- задаются матрицы

$$\mathrm{A} = \begin{pmatrix} \alpha & 1 \\ 1 & 0 \end{pmatrix}, \mathrm{B} = \begin{pmatrix} \alpha & \alpha + 1 \\ 1 & 1 \end{pmatrix}$$

и отображение

$$\pi: \{0,1\} \to \{\mathrm{A}, \mathrm{B}\} : 0 \to \mathrm{A}, 1 \to \mathrm{B}.$$

- для двоичного сообщения $m = m_1 m_2 \dots m_k$ хэш-функция H имеет следующий вид:

$$\mathrm{H}(m) = \pi(m_1) \cdot \pi(m_2) \cdot \dots \cdot \pi(m_k)$$

Определитель получившейся матрицы всегда будет равен единице. Данной хэш-функции можно сопоставить граф Кэли $\mathrm{Cay}(G, S)$, где G – конечная группа $G = SL_2(\mathbb{F}_{2^n})$, S – набор генераторов $\{\mathrm{A}, \mathrm{B}\}$.

Безопасность схемы основывается на параметрах получившегося графа – его обхвата и расширяющих свойств, эффективность схемы обусловлена быстрым вычислением произведения матриц.

Рассмотрим свойства необратимости и стойкости к коллизиям для хэш-функции ZT. Необратимость вытекает из того, что полученное произведение матриц нельзя разложить на их последовательное произведение, а задача поиска коллизий сводится к поиску двух цепочек произведений генераторов $s_1, s_2, \dots, s_n = \sigma_1, \sigma_2, \dots, \sigma_m \in \mathrm{S}$ таких, что соответствующие произведения совпадают в G:

$$s_1 s_2 \dots s_n = \sigma_1 \sigma_2 \dots \sigma_m$$

В дальнейшем было опубликовано несколько работ, в которых описывался выбор параметров для ZT, например, для того, чтобы уровень стойкости к коллизиям был сопоставим с уровнем в хэш-функции SHA-256, предлагалось использовать неприводимый многочлен

$$\mathrm{P_n} = x^{251} + x^7 + x^4 + x^2 + 1$$

Помимо хэш-функции ZT, также были предложены хэш-функции LPS и Morgenstern, стойкость которых также основана на свойствах графов Кэли, однако, впоследствии был представлен криптоанализ всех трёх схем, поэтому использование хэш-функций, основанных на графах Кэли, на сегодняшний день не рекомендовано. Тем не менее, построение стойкой хэш-функции на графах Кэли остаётся открытой задачей и темой для исследований – взлом был представлен только для частных случаев, но не для общего случая [2].

Charles, Goren и Lauter в работе [3] была предложена хэш-функция, основанная на графе изогений суперсингулярных эллиптических кривых над конечным полем.

Основная идея хэш-функции:

- строится граф Рамануджана – большой граф изогений второй степени суперсингулярных эллиптических кривых над конечным полем $\mathbb{F}_{p^2}$ (p – простое);
- каждая вершина графа является j-инвариантом эллиптической кривой;
- фиксируется суперсингулярная кривая $E/\mathbb{F}_{p^2}$;
- для получения хэш-значения двоичного сообщения $m = m_1 m_2 \dots m_n$ строится путь в графе, начиная с кривой E, заранее определив, в какую сторону необходимо двигаться в случае $m_i = 0$ и $m_i = 1$.

На сегодняшний день хэш-функция CGL является стойкой, она обладает свойствами необратимости, так как вычислительно сложно найти путь в графе из конечной кривой в начальную кривую, а также стойкости к коллизиям: явное обнаружение коллизии в этой хэш-функции эквивалентно нахождению двух изогений одной и той же степени ℓ между парой суперсингулярных эллиптических кривых. Если граф $\mathrm{G}(p, \ell)$ не имеет малых циклов, то эта проблема очень сложна, поскольку построение изогений большой степени между кривыми является вычислительно сложной задачей.

Однако данная хэш-функция имеет ключевой недостаток – низкая скорость вычисления: «стоимость» за один бит сообщения в рамках количества операций умножений в конечном поле $\approx 2 \cdot \log(p)$ [3].

В дальнейшем CGL была усовершенствована (увеличена скорость хэширования) и получила широкое развитие [4]. Стоит отметить, что на сегодняшний день такая доказуемо безопасная хэш-функция является постквантовой, то есть основана на задаче, не решаемой на квантовом компьютере.

Доказуемо безопасные хэш-функции используются реже, чем хэш-функции на специальных конструкциях, что обусловлено более дешёвой конструкцией последних. Из этого вытекает вывод – открывается направление для будущих исследований. Рассмотренные хэш-функции выполняют функцию свёртки, а также являются доказуемо безопасными, то есть основываются на математически трудных задачах, в то время как хэш-функции на специальных конструкциях являются условно безопасными, пока не доказано обратное. Вследствие этого использование доказуемо безопасных функций более предпочтительно с точки зрения безопасности. При построении таких хэш-функций в их основе следует использовать задачи постквантовой криптографии для обеспечения безопасности данных на десятки лет вперёд.

Данные хэш-функции могут найти применение при создании схем электронных подписей, в дистанционно-ограниченных протоколах аутентификации, а также в тех схемах, где скорость соединения или обработки информации не является критичной.

Литература

1. Tillich J.P., Zémor G. Hashing with SL 2 //Advances in Cryptology – CRYPTO’94: 14th Annual International Cryptology Conference Santa Barbara, California, USA August 21–25, 1994 Proceedings 14. Springer Berlin Heidelberg. 1994. С. 40-49.
2. Petit C., Quisquater J.J. Rubik's for cryptographers //Cryptology ePrint Archive. 2011.
3. Charles D.X., Lauter K.E., Goren E.Z. Cryptographic hash functions from expander graphs //Journal of CRYPTOLOGY. 2009. Т. 22. №1. С. 93-113.
4. Doliskani J., Pereira G.C.C.F., Barreto P.S.L.M. Faster cryptographic hash function from supersingular isogeny graphs //Cryptology ePrint Archive. 2017.

УДК 004.75

ПРИМЕНЕНИЕ ДРЕВОВИДНЫХ СТРУКТУР ДЛЯ РАСПРЕДЕЛЕННОГО ХРАНЕНИЯ ДАННЫХ

Кича И.В.[1] (студент), **Тимкин А.К.**[1] (студент)
Научный руководитель – к.т.н., доцент ФБИТ Таранов С.В.
1 – Университет ИТМО
e-mail: kicha@niuitmo.ru

Аннотация
В ходе данной работы были рассмотрены различные методы организации систем распределенных реестров. Были рассмотрены примитивные структуры, древовидные и их модификации. Была проанализирована производительность системы хранения данных при использовании разных подходов. Были предложены эффективные алгоритмы подтверждения наличия блока в цепи.

Ключевые слова
Распределенные реестры, хранение данных, древовидные структуры данных, алгоритмы упрощенной проверки оплаты, блокчейн.

По данным различных исследования последнее десятилетие возрастает количество исследований в области систем распределенных реестров. Таким образом, число разработок увеличивается и их область применения расширяется. Одним из таких направлений является возможность хранения данных. Системы, основанные на технологии блокчейн, в свою очередь обеспечивают конфиденциальность информации и ее целостность, что является их важным преимуществом перед традиционными подходами. При этом существует проблема пропорционального роста размера цепочки блоков при добавлении новых данных [3].

В результате изучения одномерных методов организации блокчейна было установлено, что данный метод хранения данных способен осуществлять функционирование систем распределенных реестров на базовом уровне, но при этом важной проблемой будет пропорциональный рост размера цепочки блоков при добавлении новых данных. Более того, описанные методы позволяют хранить минимальное количество данных, что в свою очередь ставит под вопрос эффективность использования одномерных последовательных цепочек. Также при введении ограничения на размер блока возможно замедленное время обработки транзакции в будущем, так как при определенных алгоритмах консенсуса будет увеличиваться время генерации новых блоков. В связи с упомянутыми выводами возникает необходимость рассмотрения возможности применения других структур данных в системах распределенного реестра в контексте хранения информации.

Было установлено, что древовидные структуры данных позволяют ускорить процесс проверки валидности блоков по сравнению с хеш-цепочками. Применение деревьев Меркла позволяет получить прирост производительности системы за счет уменьшения числа необходимых действий для доказательства существования блока в структуре. Данная возможность породила появление упрощенной проверки оплаты (SPV), что позволяет, имея только корень дерева и аутентификационный путь, подтверждать данные без необходимости загружать все блоки [1]. На рисунке 1 изображен аутентификационный путь к блоку L1, построенный для подтверждения наличия этого блока в структуре.

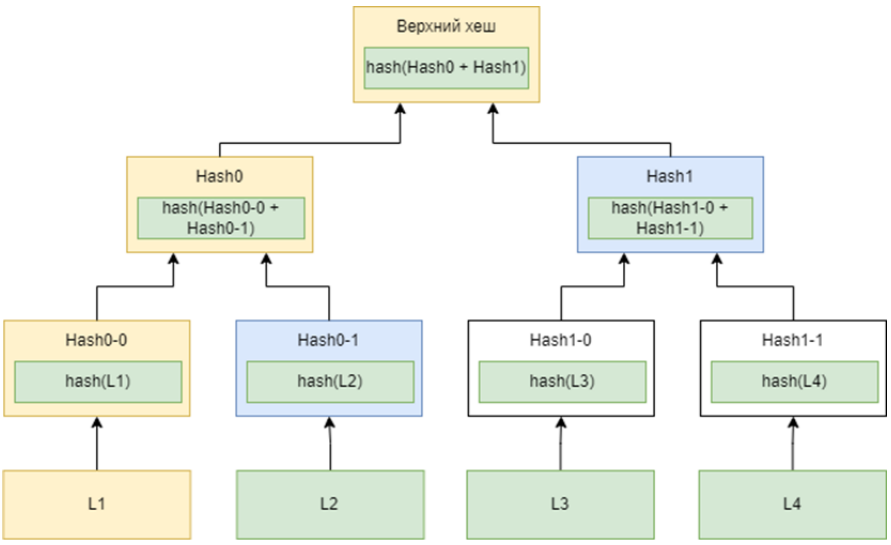


Рис. 1. Аутентификационный путь к блоку L1

При этом стоит упомянуть, что данная технология имеет и недостатки. Так, например, для вычисления верхнего хеша необходимо знать все промежуточные узлы дерева, для чего необходимо либо их хранить в системе, либо вычислять заново при каждом запросе. Первый вариант при больших размерах дерева потребует большего количества места в памяти. Второй в свою очередь повлечет снижение производительности системы. Рассмотренная структура данных имеет модификации, которые следует рассмотреть для более точного анализа вариантов хранения данных с помощью систем распределенных реестров.

Необходимо рассмотреть модификацию хеш-дерева – Merkle Patricia Trie, которая представляет собой дерево Меркла, которое хранит пары ключи значение, где ключи являются последовательностями байтов. Данная структура данных имеет определенную организацию блоков, которую удобно было бы рассмотреть на примере криптовалютной системы Ethereum. В данном контексте вводится понятие ниббла – половины байтов – последовательность которых выступает в дереве ребром. Каждый узел может иметь до 16 дочерних блоков.

Существует три вида блоков:

1. Branch node – блок, содержащий до 16 ссылок на дочерние узлы, но также может содержать значение.
2. Extension node – блок, используемый, используемый для хранения части пути, общей для какого-то множества узлов и содержащий ссылку на следующий branch node.
3. Leaf node – конечный блок, хранящий часть пути и значение.

Структура дерева Меркла-Патриция на примере криптовалюты Ethereum изображена на рисунке 2.

Очевидно, что описанная структура данных имеет ряд преимуществ над обычными хеш-деревьями. Прежде всего наличие префиксов позволяет упростить поиск значений, так как в отличие от дерева Меркла дерево Меркла-Патриция позволяет быстро найти путь для доказательства наличия блока. Более того, наличие трех видов узлов позволяет одинаково хорошо работать как с малыми объемами данных, так и с большими.

При всех положительных сторонах деревья Меркла-Патриция все же требуют для доказательства использование для каждого узла хешей всех дочерних блоков, что все так же остается неоптимальным.

Более современной и удобной структурой является дерево Веркла. Со структурной точки зрения оно очень схоже с деревом Меркла-Патриция, что изображено на рисунке 3. Так, например для нахождения блока с адресом 4cc, нам нужно найти дочерний узел корня по адресу 4, в его дочерних узлах найти блок с адресом c, в котором

опять необходимо обратиться к дочернему узлу по адресу c. В конечном счете мы получаем доступ к информации.

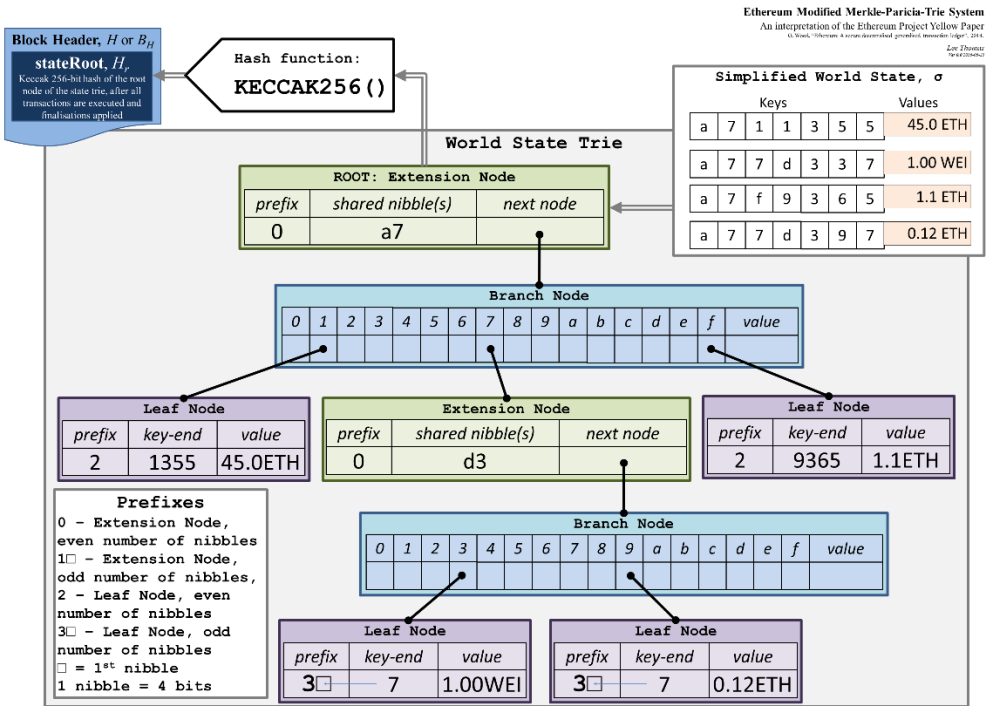


Рис. 2. Структура Merkle Patricia Trie на примере Ethereum [2]

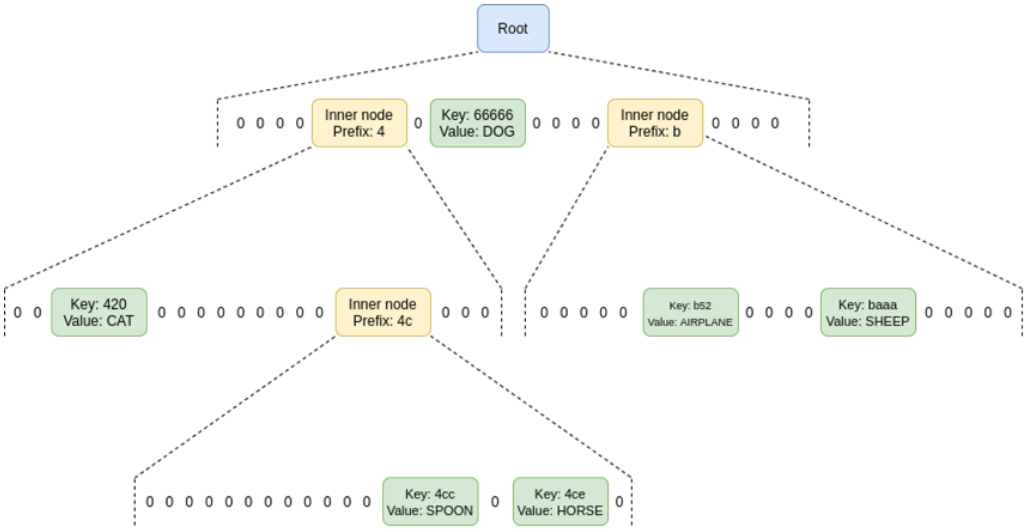


Рис. 3. Структура дерева Веркла

Так как дерево Меркла-Патриция требует учитывать хеши узлов-сестер, при, например, размере дерева в 256 узлов на верхнем уровне будет почти наверняка заполнены все 16 узлов, а на втором уровне заполненность будет в среднем на уровне 63.3%. Дерево Меркла не имеет такого недостатка. Вместо этого для доказательства нужно иметь только путь к блоку и небольшое дополнение.

Под дополнением подразумевается так называемое векторное обязательство. Под этим понятием подразумевается хеш-функция, которая обрабатывает список z1,z2…zn и получает результат C. Важной особенностью при этом является то, что векторное обязательство позволяет установить, что для определенного C i-ый элемент равен zi, при

котором нет необходимости задействовать другие дочерние узлы. Общая схема работы данного алгоритма доказательства принадлежности блока дереву Веркла изображена на рисунке 4.

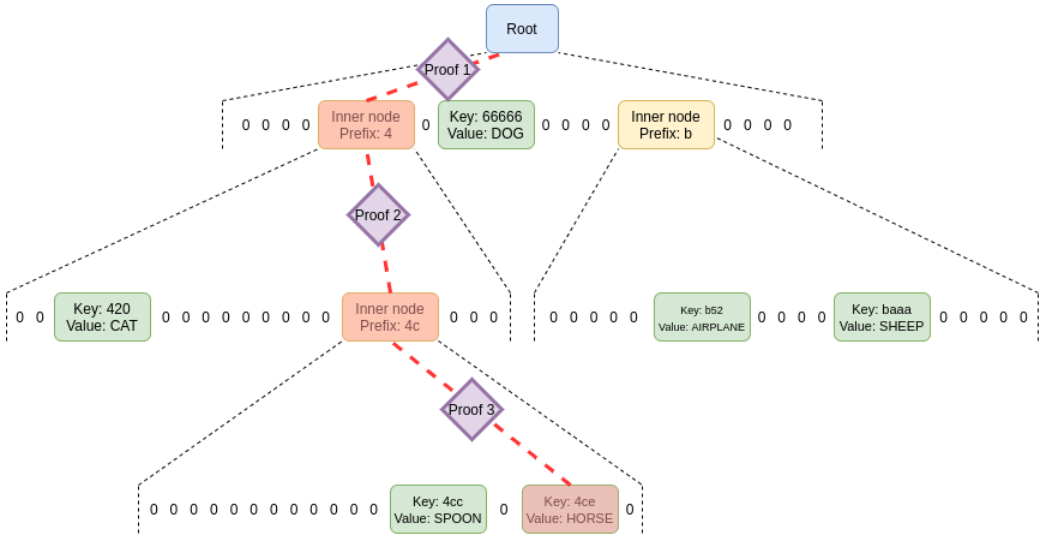


Рис. 4. Доказательство принадлежности блока в дереве Веркла

Более производительным обязательством, нежели векторное, является полиномиальное. Полиномиальное обязательство позволяет нам хешировать полином и доказывать существование на основе оценки полинома в заданной точке. Данный алгоритм можно использовать следующим образом: ввести набор c1,c2…ci и задать y1,y2…yi, так чтобы существовал полином P такой, что P(ci) = yi. Данный полином можно найти с помощью интерполяции методом Лагранжа. Одним из наиболее эффективных алгоритмов доказательств через полиномиальные обязательства является алгоритм KZG (Kate, Zaverucha and Goldberg), где обязательство представляет собой одну 32-48-байтную точку эллиптической кривой [4].

Таким образом, при ширине дерева равной 256 и использовании алгоритма KZG, доказательство будет занимать 96 байт, что в три раза меньше, нежели доказательство Меркла [5].

В результате исследования было установлено, что древовидные структуры являются наиболее эффективными среди рассмотренных для хранения данных. Были изучены достоинства и недостатки каждого решения, что позволило определить, какие из них оправданы для применения в области хранения данных. Таким образом, среди рассмотренных решений наиболее быстрым и удобным оказалось дерево Веркла. Результаты выполненной работы в перспективе планируется развить в рамках выполнения других научно-исследовательских работ, чтобы изучить тему исследования более подробно с перспективой включения результатов в магистерскую диссертацию.

Литература

1. Antonopoulos A.M. Mastering Bitcoin: unlocking digital cryptocurrencies. 2014 //isbn: 9781449374044. С. 272.
2. Ethereum StackExchange: ELI5 How does a Merkle-Patricia-trie tree work? [Электронный ресурс]. Режим доступа: https:// ethereum. stackexchange.com/ questions/6415/ eli5-how-does-a-merkle- patricia-trie-tree-work (дата обращения: 15.01.2023).
3. Vitalik Buterin's website: Verkle trees [Электронный ресурс]. Режим доступа: https://vitalik.ca/general/2021/06/18/verkle.html (дата обращения: 15.01.2023).

4. Kate A., Zaverucha G.M., Goldberg I. Constant-size commitments to polynomials and their applications //International conference on the theory and application of cryptology and information security. Springer. Berlin. Heidelberg. 2010. С. 177-194.
5. Dankrad Feist: KZG polynomial commitments [Электронный ресурс]. Режим доступа: https://dankradfeist.de/ethereum/2020/06/16/kate-polynomial-commitments.html (дата обращения: 15.01.2023).

УДК 004.056.53

ОСОБЕННОСТИ БИНАРНОЙ ЭКСПЛУАТАЦИИ МЕТОДА SIGRETURN ORIENTED PROGRAMMING (SROP)

Коновалов П.И.[1], Андрушкевич Д.В.[1]
Научный руководитель – к.т.н., полковник Глыбовский П.А.[1]
1 – Военно-космическая академия имени А.Ф. Можайского
e-mail: vka@mil.ru

Аннотация
В работе приводится обзор существующих мер защиты исполняемых файлов. Рассмотрены особенности одного из методов эксплуатации двоичных файлов SROP и применение его на практике. Приведены рекомендации, позволяющие избежать возможности его эксплуатации.
Ключевые слова
Программирование с возвратом сигнатуры, SROP, обработчик сигнала, гаджет, системный вызов.

На данный момент киберпреступность в мире процветает, исследователи прилагают все усилия для перекрытия обнаруживаемых уязвимостей. Мерами защиты являются предупреждение исполнения кода (NX, None Executable), запрещающее исполнение стека и рандомизация адресного пространства (ASLR, Address Space Layout Randomization), случайным образом размещающая код программы в памяти, но киберпреступники становятся всё более изобретательными, а вектор атак всё больше склоняется в сторону бинарной эксплуатации.

Целью данной статьи является рассмотрение одного из существующих методов атаки на исполняемые файлы, основанного на программировании с возвратом к функции восстановления контекста (SROP, Sigreturn Oriented Programming). Этот метод актуален по причине применения злоумышленниками самых нетрудно реализуемых и эффективных по отношению к современным мерам защиты атак. Данный метод был описан относительно недавно и активно применяется в настоящее время, его размещение на временной прямой видно на рисунке 1.

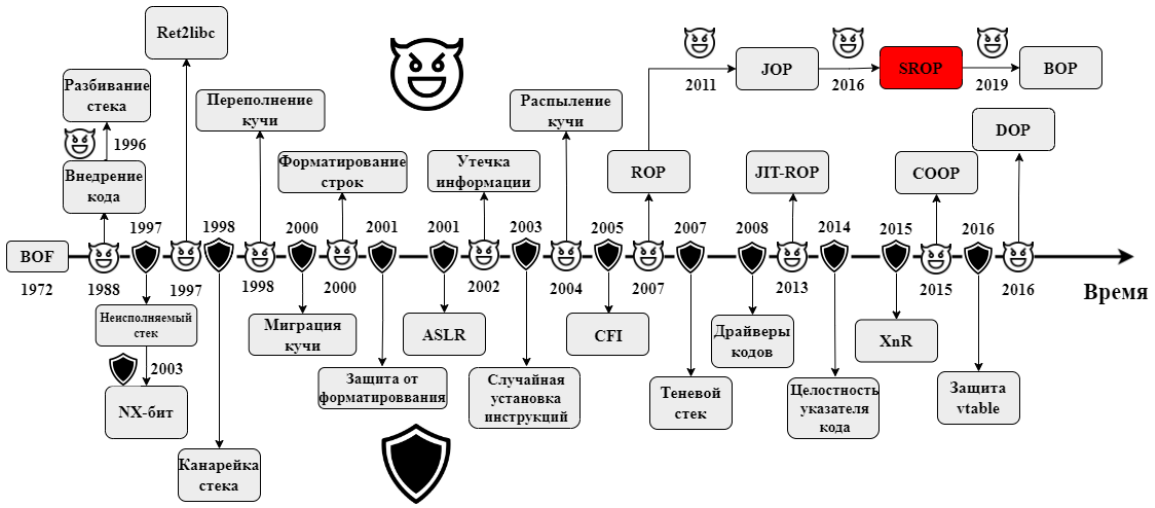


Рис. 1. Эволюция развития техник бинарной эксплуатации

Рассмотрим особенности техники SROP. SROP относится к ROP (Return-Oriented Programming) атакам. ROP – это техника эксплуатации, основанная на поиске блоков инструкций внутри бинарного файла, называемых гаджетами. Ее особенностью является

создание цепочек вызовов в стеке путем манипулирования адресами памяти, в которых содержатся необходимые инструкции в программном коде. Как правило, цепочки содержат множество гаджетов, но проблема, возникающая в данном случае, заключается в их ограниченности. Отличие SROP от стандартного ROP заключается в том, что для его применения требуется всего два гаджета: «pop rax», содержащий ID (identifier) системного вызова, который злоумышленник хочет выполнить и «syscall», необходимый для самого вызова. Но прежде, чем приступать к SROP необходимо сказать о том, как работают сигналы.

Приведем понятие сигнала. Сигнал – это форма межпроцессного взаимодействия, используемая системами Unix и соответствующая стандартам POSIX (Portable Operating System Interface). Они используются для уничтожения процессов, чтобы сообщить им об истечении времени ожидания или предупредить их об исключительном поведении [1].

Примеры сигналов:

1. Ctrl+C (уничтожение процесса);
2. Ctrl+Z (временная остановка процесса).

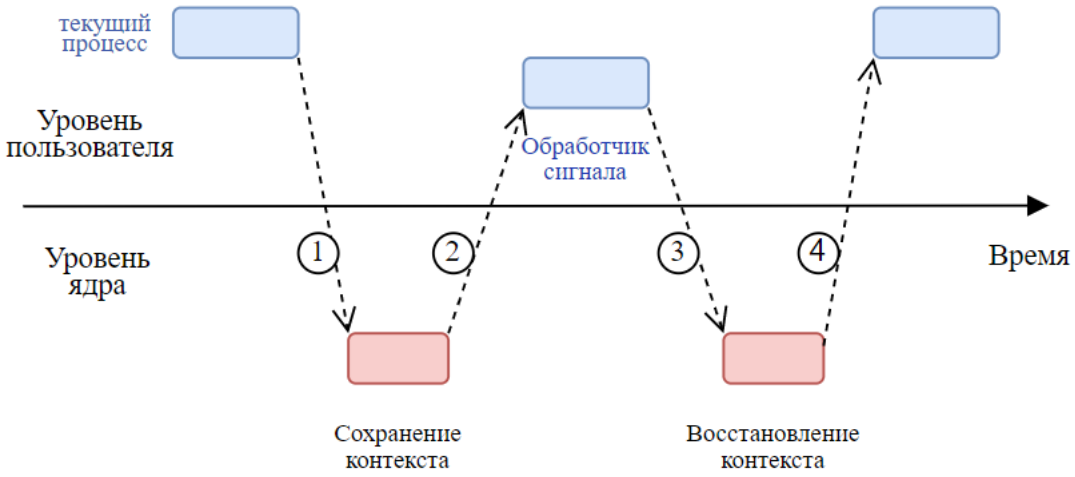


Рис. 2. Процесс обработки сигналов

Обработка сигналов осуществляется на двух уровнях и включает следующие этапы (рис. 2):

1. При появлении сигнала процесс временно приостанавливается и переходит в пространство ядра.
2. Ядро сохраняет регистры в соответствующем кадре стека для процесса и переходит к ранее зарегистрированному обработчику сигнала (в пространстве пользователя) для обработки соответствующего сигнала.
3. Ядро восстанавливает кадр стека, ранее сохраненный для процесса, благодаря sigreturn.
4. Процесс возвращается в пространство пользователя.

В этом процессе существует 2 недостатка:

1. Контекст сигнала можно редактировать при нахождении в пространстве пользователя.
2. Ядро не сравнивает сохраненный контекст сигнала с восстановленным.

Таким образом, процесс обработки сигналов заключается в следующем.

В момент, когда сигнал отправляется текущему процессу, и у процесса есть зарегистрированный обработчик сигнала, ядро установит его фрейм времени выполнения в пользовательском стеке. Этот фрейм содержит все детали, относящиеся к контексту процесса. По окончании, ядро передает управление обработчику сигнала. Контекст сигнала настроен таким образом, что адрес возврата обработчика сигнала

является функцией restore_rt, которая вызывает системный вызов rt_sigreturn. Этот системный вызов очищает контекст сигнала. После завершения очистки управление передается обратно. Ядро не только очищает стек, но также загружает детали контекста процесса, присутствующие в фрейме сигнала, а затем возвращает управление туда, куда указывал rip [2].

Подводя итог, настройка и очистка контекста сигнала довольно проста. Настройка происходит на уровне ядра и как только она кончается, ядро очищает все, загружая контекст процесса и передавая управление обратно при помощи функции restore_rt, вызываемой из кода пользовательского пространства. Так как это происходит в пространстве пользователя, злоумышленник имеет возможность подменить контекст процесса и выполнить любую функцию с доступными гаджетами. Рассмотрим пример использования SROP для более детального обсуждения.

Рассмотрим пример применения SROP. Исходный программный код написан на языке Ассемблер. Он считывает ввод и выводит его на экран. Здесь содержится стандартная уязвимость переполнения буфера, так как считывается 300 байт, а буфер имеет размер в 8 байт. NX включен, поэтому выполним атаку при помощи ROP. Программа слишком мала, в ней нет необходимых гаджетов, библиотека не подключена, а значит применяется SROP. Изменим разрешения на исполнение кода в стеке при помощи системного вызова mprotect, после чего запишем шелл-код и исполним его. В программе есть гаджет системного вызова, но нет системного вызова sigreturn. Системный вызов считывания ввода сохраняет количество прочитанных байт в регистре rax. Это означает, что доступно управление регистром rax со значениями в диапазоне от 0 до 300, включая ID системного вызова sigreturn, который равен 15. Теперь, когда есть все необходимое, можно приступить к созданию эксплойта.

Полезная нагрузка, отправляемая программе, будет состоять из следующих частей:

1) 8 байт до переполнения;
2) адрес вызова считывания ввода;
3) адрес гаджета системного вызова;
4) контекст функции mprotect.

На рисунке 3 приведена схема вызова sigreturn с визуализацией данных в стеке.

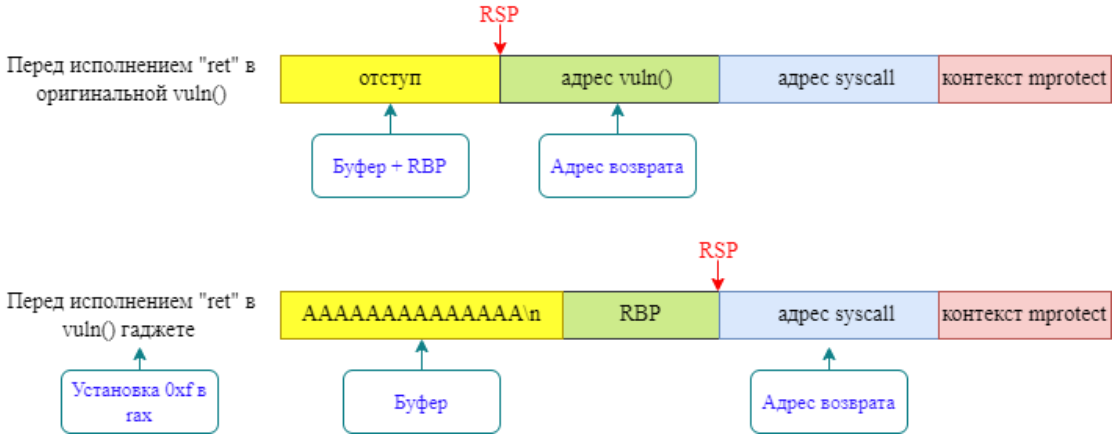


Рис. 3. Схема работы SROP

По достижении второго считывания ввода, отправляем 15 байт для управления rax и выполняем системный вызов sigreturn.

При помощи инструмента ROPgadget находятся необходимые адреса, при помощи библиотеки pwntools создается кадр стека для «sigreturn» и отправляется полезная нагрузка.

Это настроит системный вызов mprotect, чтобы пометить область памяти 0x400000 как исполняемую и доступную для записи, чтобы разрешить выполнение шелл-кода по известному адресу. Затем, перемещаем стек в эту область, чтобы иметь возможность записывать в него данные. Установив rsp в адрес, содержащий точку входа в программу, обеспечивается нормальный поток управления программой, поскольку она будет возвращаться в адрес, лежащий на вершине стека [3].

Теперь, когда в стеке разрешено исполнение кода, возможно использование переполнения буфера с полезной нагрузкой, например (рис. 4):

1) 8 байт до переполнения;
2) адрес шелл-кода;
3) шелл-код.

Рис. 4. Пример полезной нагрузки

Ответом на эту полезную нагрузку является командная оболочка операционной системы.

Таким образом, нетрудно заметить, что стандартных мер защиты зачастую оказывается недостаточно. Самая простая программа, содержащая в себе ошибку программиста, может привести к удаленному исполнению кода или повышению привилегий, в зависимости оттого, каковы права пользователя, запустившего программу.

К сожалению, последствия подобных ошибок могут быть крайне печальными для большого количества людей, поэтому важно, чтобы каждый IT-специалист разбирался в основах информационной безопасности, это поможет сократить количество возникающих уязвимостей и киберпреступников, эксплуатирующих их.

Для предотвращения возможности реализации SROP, необходимо готовить квалифицированных разработчиков, имеющих достаточные знания в области информационной безопасности, касающейся работы исполняемых файлов на низком уровне, работы с памятью, а также применение стандартных средств тестирования создаваемых ими программ на наличие уязвимостей.

Литература

1. SROP – An Introduction [Электронный ресурс]. Режим доступа: https://www.pwnthebox.net/reverse/engineering/and/binary/exploitation/series/2021/05/09/sigreturn-oriented-programming.html (дата обращения: 17.03.2023).
2. Hackthebox – sick rop write-up [Электронный ресурс]. Режим доступа: https://shakuganz.com/2022/07/11/hackthebox-sick-rop-write-up/ (дата обращения: 19.03.2023).
3. SROP – Sigreturn Oriented Programming [Электронный ресурс]. Режим доступа: https://lo0l.com/2020/01/01/srop.html (дата обращения: 20.03.2023).

УДК 681.584.782

АНАЛИЗ ПАТЕНТНЫХ ИССЛЕДОВАНИЙ И НАУЧНО-ТЕХНИЧЕСКОЙ ЛИТЕРАТУРЫ ПО ВОПРОСАМ ЭЛЕКТРОМАГНИТНЫХ ЗАМКОВ

Круглов Д.Д.[1] (студент), **Морозова Е.М.**[1] (студент),
Пирожникова О.И.[1] (к.т.н., доцент)
Научный руководитель – д.т.н., профессор Ткалич В.Л.[1]
1 – Университет ИТМО
e-mail: rd.zz7@yandex.ru

Исследовательская работа выполнена в рамках НИР МК 5323.2023 Развитие теоретических и экспериментальных методов защиты объектов транспортной инфраструктуры от техногенных и террористических угроз на базе компьютерных технологий проектирования компонент интегрированных систем безопасности.

Аннотация
В данной работе проведено патентное исследование по вопросам электромагнитных замков за 2012–2022 годы. Патенты были проанализированы на предмет достоинств полезных моделей и изобретений, а также вносимых в модель изменений, на основе которых был сделан вывод о дальнейших перспективах и тенденциях развития электромагнитных замков.
Ключевые слова
Патент, электромагнитный замок, запирающие устройства, физические средства, системы ограничения доступа.

Информационная безопасность – это защищенность информации и объектов инфраструктуры, в которых эта информация хранится и обрабатывается, от нежелательного искажения или утечки посредством действий злоумышленника или происшествий естественного характера. Ключевым из способов обеспечения целостности, доступности и конфиденциальности защищаемой информации является ее физическая защита. В целях защиты информации путем организационных мероприятий и совокупности технических средств создаются препятствия для несанкционированного физического доступа. При этом в физической защите используются различные средства – системы изоляции, системы контроля доступа и запирающие устройства.

Электромагнитный замок – устройство, относящееся к запирающей охранной технике, основным рабочим элементом которого является электромагнит. Принцип работы электромагнитного замка позволяет использовать его для запирания раздвижных дверей, с целью исключения возможности несанкционированного доступа, потому эти замки широко применяются в различных областях промышленности – сегодня электромагнитные замки ставят не только на проходных дверях и калитках предприятий, но и внедряют в запорные механизмы различных холодильников и сейфов (рисунок).

Рисунок. Приоритеты развития электромагнитных замков

Было проведено патентное исследование за 2012–2022 годы, в ходе которого было найдено 10 патентов. Детальное изучение найденных патентов выявило ряд проблем, на устранение которых и нацелены были изобретения. Все найденные патенты можно разбить на три большие группы в зависимости от приоритета, стоявшего перед автором патента: основными акцентами развития электромагнитных замков стали упрощение сборки конструкции, уменьшение потребления энергопитания, повышение надежности и отказоустойчивости конструкции. На рисунке выше представлено разделение патентов по группам, а также в скобках указаны года подачи данных патентов.

В каждой группе прослеживаются свои методы достижения желаемых результатов, кратко приведенные в таблице ниже. Так, упрощение сборки конструкции достигается путем замены усилий пружины взаимодействием магнитных полей, что позволяет избавиться от промежуточных механизмов. В патентах также форсируются идеи исполнения основного количества деталей штамповкой, что в свою очередь упрощает как техническую, так и финансовую реализацию замка. Встречаются и предложения по усовершенствованию формы ряда составных деталей, что позволяет не только уменьшить количество составных элементов путем комбинации их в одну деталь, но и снять ограничения на форму других компонентов (в частности убрать ограничения на форму головки ригеля) [1].

Таблица

Методы достижения желаемых результатов из патентов

Упрощение сборки конструкции	Уменьшение потребления электропитания	Повышение надежности рабочего механизма
Замена усилий пружины взаимодействием магнитных полей	Применение немагнитных материалов в фиксаторе	Введение датчика состояния двери
Упрощение составных деталей	Использование шунта для регулирования магнитного потока электромагнита	Применение магнитомягких ферромагнитных материалов
Уменьшение количества элементов	Уменьшение рассеивания магнитного потока катушкой электромагнита	Снижение величины тока в катушке электромагнита
	Внедрение актуатора над фиксатором (Сердечник актуатора - неодимовый магнит)	Уменьшение нагрузки на фиксатор
	Повышение тягового усилия электромагнита	

В целях уменьшения потребления электропитания предлагается применение немагнитных материалов при изготовлении элементов фиксатора. Это позволит исключить залипания конца ригеля при отпирании, что сейчас требует большего времени нахождения электромагнита под напряжением. Уменьшить потребление электроэнергии при отпирании замка можно и уменьшив рассеивание магнитного потока, используемого для нейтрализации поля постоянного магнита. Для этого сердечник получает дополнительные два плоских диска, между которыми и размещается катушка электромагнита. Идея внедрения актуатора, воздействующего на фиксатор подвижным штоком, позволит снизить потребление электроэнергии при открытии замка примерно в пять раз, а при изготовлении сердечника данного актуатора из неодимового магнита, находящегося под управлением магнитного поля обмотки, позволит перейти на питание от автономных источников электроэнергии, например на гальванические элементы. Достичь уменьшения потребления электропитания можно и повышением

тягового усилия электромагнита – тогда освобождение ригеля произойдет при прохождении меньшего тока через обмотку.

Уменьшая ток через обмотку электромагнита можно повысить и надежность запорного механизма: введение магнитных шунтов в фиксатор поможет снизить температуру нагревания катушки, что положительно скажется на ее физических свойствах. Повышение надежности рабочего механизма электромагнитного замка в полезных моделях может достигаться и путем введения датчика состояния двери, который позволит использовать замок в совокупности с системами дистанционного управления [3]. Нивелирование прилипания подвижных элементов к сердечнику электромагнита из-за остаточного магнетизма предлагается достигать за счет использования магнитомягких ферромагнитных материалов при изготовлении ригеля и шунта. Это позволит исключить несрабатывание отпирающего механизма замка, что благоприятно скажется на повышении надежности устройства. Во избежание же самопроизвольного отпирания запорного механизма предлагается ослабить нагрузку на фиксатор, что практически исключит вероятность перекоса ригеля. Это достигается изменением геометрических размеров отверстия ловителя, которое подгоняют по диаметру под размеры головки засова.

Переходя от анализа патентных решений к перспективным направления развития, можно обратить внимание на вторую группу патентов: как видно из дат подачи патентов, проблема высокого потребления электропитания запорными механизмами являлось актуальной на протяжении всего рассматриваемого временного интервала, причем в 2021 году продолжалась активная работа по данному направлению. В добавок к этому следует обратить внимание на то, что согласно тарифам на начало 2023 года цена электроэнергии приблизилась к 5 руб./кВтч, что за указанный десятилетний период увеличилась в 1,8–2,5 раза в зависимости от региона. Любая компания заинтересована в сокращении своих расходов для увеличения прибыли, потому данное направление с большой вероятностью не потеряет свою актуальность в ближайшее время, а значит работы в данном направлении будут продолжаться. Вопрос упрощения сборки тоже может стать одним из приоритетных векторов развития, так как в условиях развивающейся автоматизации производства упрощенные модели с меньшим количеством деталей станут доступны для сборки работами. Что касается повышения надежности, то в этом направлении, судя по патентному поиску, работа прекращена, но оно и понятно – известны методы взлома и отпирания подобных запирающих устройств, нивелировать которые невозможно, потому приходится комбинировать данные замки с другими средствами физической защиты [2]. Среди неуказанных ранее направлений можно также предположить, что упор будет сделан на внедрении систем дистанционного управления. Удобство дистанционного контроля состояния замков поможет своевременно среагировать на инциденты информационной безопасности. Также работы могут вестись и в уменьшении себестоимости производства электромагнитных замков – будет вестись поиск более дешевых материалов, использование которых не сильно скажется на надежности запорного механизма, но позволит за счет уменьшения стоимости заинтересовать компании в приобретении. Отчасти уменьшение себе стоимости является следствием упрощения сборки конструкции, так как поставленное на конвейер производство значительно уменьшает затраты, однако не всегда более простая модель становится более дешевой.

В заключении хочется отметить, что в современном мире крупные фирмы уделяют повышенное внимание информационной безопасности своих активов. Любые физические средства защиты информации постоянно совершенствуются, но чтобы понять, какие требования предъявляются к моделям тех или иных средств разграничения доступа, следует проводить патентный анализ, чтобы выпускаемая фирмой продукция была конкурентноспособной. В условиях реализации программ по импортозамещению

подобный подход позволит захватить преимущественное положение на рынке, заменив собой ушедшие иностранные компании.

Литература

1. Патент № 161964 Российская Федерация, МПК E05B 47/00 (2006.01). Замок: №2015155656/12: заявл. 24.12.2015: опубл. 20.05.2016/ Легостаев А.В., Полтаев А.Г.; заявитель ООО "Электронные запирающие механизмы". 28 с.: ил. Текст: непосредственный.
2. Ефременко Н.В., Цедрик С.Н. Современные электронные запирающие устройства: понятие, классификация, способы взлома и отпирания // Вестник Полоцкого государственного университета. Серия D. Экономические и юридические науки. 2015. № 14. С. 161–165.
3. Патент №2615718 Российская Федерация, МПК E05B 47/02 (2006.01). Электромагнитный замок: № 2015143911: заявл. 13.10.2015: опубл. 07.04.2017. Карюканов Г.С., Малиновский А.Е.; заявитель ООО “Инженерно-технический центр "ПРОМИКС"”. 11 с.: ил. Текст: непосредственный.

УДК 004.457

РАЗРАБОТКА ПРОГРАММНОГО СРЕДСТВА ДЛЯ ОБНАРУЖЕНИЯ АНОМАЛИЙ В ФУНКЦИОНИРОВАНИИ ПРОЦЕССОВ НА ОСНОВЕ ТРАССИРОВКИ СИСТЕМНЫХ ВЫЗОВОВ В ОПЕРАЦИОННОЙ СИСТЕМЕ LINUX

Ксенофонтов В.С.[1] (студент)
Научный руководитель – к.т.н., доцент Гирик А.В.[1]
1 – Университет ИТМО
e-mail: ksenofontovvs9a@mail.ru

Аннотация
В данной работе описан подход к обнаружению аномалий в функционировании процессов в операционных системах Linux во время их работы. Особенностью описанного подхода является минимизация ущерба производительности процессов за счет снижения числа переключений между режимами пользователя и ядра. Формирование реализованного подхода основано на изучении возможностей загружаемых модулей ядра, механизмов работы стандартных трассировщиков и функционала встроенных возможностей ядра.

Ключевые слова
Операционные системы Linux, пространство ядра, пространство пользователя, системные вызовы, трассировка, детектирование аномалий.

Несмотря на постепенное снижение активности вредоносного программного обеспечения, число фиксируемых инцидентов до сих пор велико. Вместе с этим убытки компаниям приносят непрофессиональные или злонамеренные действия взаимодействующих с информационными системами сотрудников.

Описанные условия являются предпосылкой для разработки и внедрения в информационные системы программных средств, детектирующих аномальную активность субъектов. Предотвращению или минимизации последствий действий, нарушающих политику информационной безопасности, способствует реализация анализа функционирования процессов во время их работы.

В операционных системах семейства Linux выполнение процессов осуществляется в двух режимах: в режиме ядра, обеспечивающем доступ к аппаратным ресурсам и управление памятью, и в режиме пользователя, обеспечивающем выполнение прикладных программ верхнего уровня. В режиме ядра процесс выполняется в адресном пространстве ядра, и процессор имеет возможность выполнять инструкции, описанные в коде, без ограничений. Процессор также имеет доступ к любому действующему адресу в памяти. Возможности задействования команд и аппаратного обеспечения в режиме пользователя существенно ограничены механизмами защиты [1].

Переключение и обмен информацией между режимами пользователя и ядра осуществляется посредством механизма системных вызовов, реализованных в виде интерфейсных библиотечных функций. В связи с тем, что каждое действие процесса, модифицирующее объекты системы, обеспечивается последовательностью системных вызовов, накопление и анализ совокупности этих последовательностей позволяет осуществить детектирование аномалий в поведении процесса.

Поскольку анализ данных о работе системы в реальном времени задействует часть доступных вычислительных ресурсов, важной задачей при разработке программного средства становится минимизация наносимого анализом ущерба производительности системы.

На сегодняшний день стандартным инструментом для отслеживания системных вызовов в Linux является утилита Strace. Принцип ее работы основан на возможностях системного вызова ptrace(), позволяющего одному процессу отслеживать и вмешиваться

в выполнение другого процесса путем проверки и изменения значений в памяти и регистрах [2]. Использование этого системного вызова является основным недостатком утилиты в контексте задачи быстрого анализа, так как приводит к дополнительным переключениям режимов и существенно замедляет подконтрольный процесс. Strace также не имеет собственных механизмов выявления аномалий, а эффективный анализ объемного текстового отчета о ее работе невозможен.

Другим механизмом, позволяющим взаимодействовать с инфраструктурой ядра, являются модули ядра (Loadable Kernel Modules). Механизм позволяет динамически загружать написанные пользователем модули, расширяющие функционал, в ядро и выгружать их из него. Для процесса разработки модулей ядра характерны некоторые трудности. Поскольку выполнение кода осуществляется в ядре, где отсутствует большое количество механизмов защиты и проверок, вероятность возникновения ошибок и ущерб системе от них возрастают. Разработку модулей ядра осложняет также значительная зависимость диапазона поддерживаемых возможностей от версии ядра.

Актуальной на сегодняшний день технологией, предоставляющей в числе прочего средства для производительной трассировки, является Berkeley Packet Filter (BPF). BPF позволяет пользователю создавать программы, загружаемые в ядро и подключаемые к тому или иному событию. Возникновение интересующего события провоцирует запуск подключенной программы. Компоненты BPF, такие как таблицы, представляющие собой настраиваемые структуры данных для хранения данных по принципу ключ / значение; вспомогательные функции для взаимодействия с ядром, представляющие собой решения базовых задач для взаимодействия с архитектурой BPF; механизмы обеспечения безопасности; ссылочная файловая система для сохранения объектов – таблиц и программ; механизмы загрузки и выгрузки программ BPF создают единую архитектуру, способствующую упрощению и повышению безопасности процесса разработки программного кода для ядра [3]. Управление объектами BPF из пространства пользователя осуществляется путем передачи команд и параметров через системный вызов bpf()**,** позволяющий, среди прочего, создавать в памяти ядра таблицы, искать, обновлять и удалять в них элементы, загружать в ядро программы BPF [4]. Таблицы BPF представляют собой память ядра, организованную в виде последовательных массивов, хэш-таблиц и других структур данных и допускающую совместный доступ к ней нескольких программ.

Наличие у технологии описанных преимуществ и отсутствие недостатков, свойственных альтернативным рассмотренным инструментам, позволили опираться на нее в процессе разработки анализатора.

Обозначенное требование быстродействия анализатора обусловило реализацию постоянно выполняемых частей функционала по сбору и анализу данных на стороне ядра, что дает следующие преимущества с точки зрения производительности:

– наличие прямого доступа к общей памяти ядра;
– возможность привязки программ к конкретным событиям;
– отсутствие в процессе работы генерации дополнительных системных вызовов.

В пространство пользователя были вынесены первоначальная настройка параметров и режима работы анализатора и журналирование, осуществляющееся в случае обнаружения аномалии.

Анализатор запускается из пространства пользователя. Алгоритм его работы состоит из следующих шагов:

– при настройке пользователем выбираются важные в контексте информационной безопасности системные вызовы;
– в ядре создаются таблицы для накопления информации о выбранных вызовах;
– в ядро загружаются программы для перехвата вызовов, реализующие запись информации в таблицы и детектирование аномалий;

– далее анализатор запускается в режиме обучения и его ядерные перехватчики сохраняют текущую работу пользователя с наблюдаемой программой (браузером, файловым менеджером) как ее эталонную, нормальную активность;

– в последствии, при запуске анализатора в режиме контроля, ядерные перехватчики будут сравнивать текущую активность программ с сохраненным для них эталоном нормальности и журналировать обнаруженные аномалии.

Схематично алгоритм работы анализатора описан на рисунке.

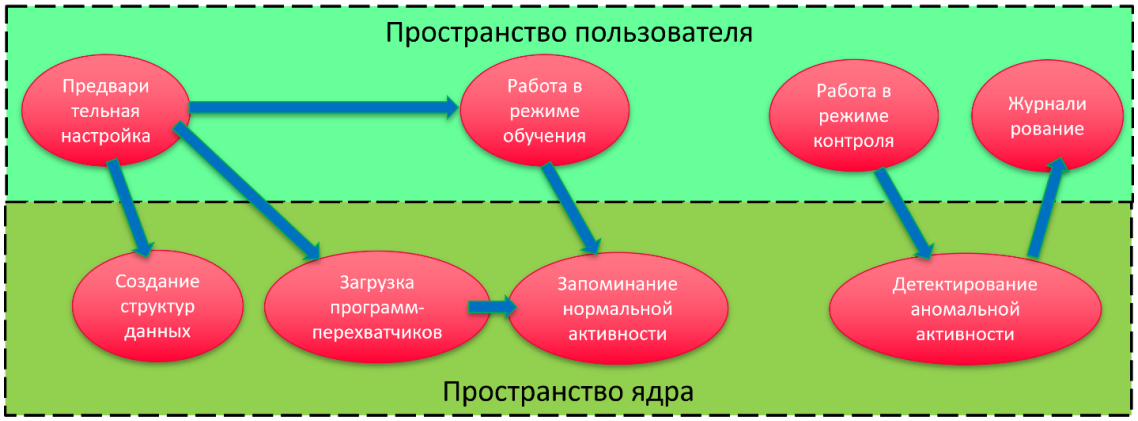


Рисунок. Принцип действия созданного программного средства

Настройка параметров и режима работы анализатора осуществляется в пользовательском режиме путем передачи программе аргументов командной строки:

– опция l запускает анализатор в режиме обучения для отслеживания нормальной работы программы, название которой передано в следующем параметре;

– опция s дает анализатору указание удалять все закрепленные в файловой системе BPF структуры данных по окончании своей работы. Без указания данной опции закрепленные структуры данных не будут удалены после выполнения программы, оставаясь пригодными для отложенного более глубокого анализа;

– опции p и u дают анализатору указание отслеживать системные вызовы, сгенерированные, соответственно, конкретным процессом или пользователем, чей идентификатор равен переданному в следующем параметре.

Получив аргументы, пользовательская часть анализатора перезаписывает значения глобальных переменных в различных секторах объекта ядерной части программы, таким образом передавая ей информацию о требуемом режиме работы ядерных перехватчиков.

Эффективности анализа, нацеленного на детектирование аномалий, способствует сбор данных об ограниченном подмножестве системных вызовов, потенциально способных нанести ущерб информационной безопасности. Перехват конкретного поднабора системных вызовов уменьшает объем обрабатываемой информации и позволяет быстрее ее анализировать. В качестве перехватываемых системных вызовов были выбраны связанные с изменением файлов и прав доступа к ним, установлением сетевых соединений, планированием задач и с запуском процессов.

Для хранения информации о перехватываемых системных вызовах была реализована система структур данных, актуальность которых поддерживается в процессе работы анализатора.

Основным хранилищем данных является общая для всех (или для конкретных выбранных) запущенных в системе процессов главная таблица, содержащая глобальный порядковый номер текущего перехваченного системного вызова с начала процесса трассировки, название вызова, временные метки входа и выхода, идентификаторы процесса, группы процессов, пользователя и группы пользователя, а также возвращенное

им в процессе работы значение. Пример фрагмента заполненного хранилища приведен в таблице 1.

Таблица 1

Основное хранилище информации о системных вызовах

Глобальный номер записи	Системный вызов	Время входа	Время выхода	Идентификатор процесса	Идентификатор пользователя	Возвращенное значение
1	write	08:19:23.384	08:19:23.39	1094	1	26
2	execve	08:21:12.489	08:21:12.512	789	100	0
3	write	08:22:56.321	08:22:56.339	789	100	2048
4	write	09:03:37.113	09:03:37.115	789	100	2048
5	execve	09:12:02.976	09:12:02.976	1094	1	0
6	read	10:39:13.843	10:39:13.845	913	100	128

Программно основное хранилище реализовано в виде таблицы, представляющей собой последовательный массив структур, описывающих записи таблицы. В случае переполнения хранилища новые записи будут заноситься в начало, перекрывая находящиеся на их месте старые.

Технология BPF предоставляет возможность трассировки системных вызовов с помощью подключения программ BPF к таким событиям, как вход в системный вызов и выход из него. В подключаемые к событиям программы в качестве аргумента передается указатель на структуру, хранящую информацию о контексте выполнения. При входе в системный вызов доступна информация о названии системного вызова, флагах выполнения и переданных аргументах. В момент выхода из системного вызова доступна информация о названии вызова, флагах выполнения и о возвращенном значении.

Во время входа в системный вызов возможна фиксация в основном хранилище только порядкового номера текущего перехваченного системного вызова, названия вызова, временной метки входа, идентификатора процесса, группы процессов, пользователя и группы пользователя. Таким образом, заполнение временной метки выхода из системного вызова и возвращаемого вызовом значения должно быть реализовано при выходе из системного вызова. При следовании этой логике программа BPF, подключенная к выходу из системного вызова, должна иметь информацию о том, в какую именно запись из основного хранилища осуществлять сохранение информации. Для решения данной задачи была реализована вспомогательная структура данных, проиллюстрированная в таблице 2.

Таблица 2

Соответствие порядкового номера текущего системного вызова в главной таблице конкретному идентификатору процесса

Идентификатор процесса	1094	789	913	…
Глобальный номер записи	1	3421	6	…

При генерации системного вызова во вспомогательной таблице создается запись, ключом которой является идентификатор процесса, а значением – порядковый номер записи о системном вызове из главной таблицы. В момент завершения работы вызова из контекста выполнения считывается номер текущего процесса. После этого временная метка выхода и возвращаемое значение заносятся в соответствующие поля записи

основного хранилища, имеющей номер, сохраненный во вспомогательной таблице для текущего процесса при входе в системный вызов. После фиксации данных в основном хранилище запись об отработавшем системном вызове из вспомогательной таблицы удаляется. Подходящей для этой задачи структурой данных является хэш-таблица, так как индексация данных осуществляется не по последовательности подряд идущих чисел, а по номерам процессов, имеющих произвольные значения.

В процессе работы программного средства также поддерживается таблица, содержащая статистические данные о работе системных вызовов в вычислительной системе. Статистические показатели могут использоваться для выявления аномалий при обнаружении излишне интенсивной работы некоторых вызовов, не характерной для конкретного процесса или пользователя.

Поскольку цель проводимого создаваемым в данной работе программным средством анализа состоит в обнаружении аномалий, из того, что под аномалией понимается отклонение от некоторой нормы, общей закономерности, следует, что норма должна быть изначально определена. Для обучения анализатора эталонному функционированию целевой программы пользователь инициирует запуск анализатора в соответствующем режиме.

В режиме обучения в специальную таблицу сохраняются системные вызовы, сгенерированные процессом во время работы. Данные хранятся в виде хэш-таблицы, каждая запись которой описывает нормальное поведение соответствующей программы. Название программы является ключом, значением является список характерных системных вызовов. Информация о характерных для программы системных вызовах хранится в виде последовательности битов, где каждый n-й бит соответствует системному вызову, имеющему в архитектуре вычислительной системы номер n. Если вызов с некоторым номером является характерным для программы, соответствующий по номеру бит в последовательности устанавливается в единицу. Для предотвращения удаления накопленных данных о нормальном поведении программ после выключения системы, при завершении работы анализатора данные выводятся в текстовый файл, из которого считываются при первом запуске программного средства после включения системы.

Анализатор, запущенный в режиме контроля, проверяет, является ли программа, сгенерировавшая системный вызов, отслеживаемой (содержатся ли данные о ней в соответствующей таблице) и, если является, проверяет наличие текущего системного вызова в списке допустимых для программы. В случае обнаружения аномалии осуществляется журналирование. Все операции чтения и записи в структуры данных, а также анализ осуществляются за константное время. В случае же релокации структур данных, амортизированная сложность по-прежнему остается константной. Таким образом, суммарная вычислительная сложность системы анализа не содержит узких мест.

Практическое применение созданного программного средства осуществлялось на примере программы «Браузер Firefox», файлового менеджера, и самостоятельно написанной программы, имеющей разнородный функционал.

При работе анализатора в режиме обучения нормальному поведению программы Firefox пользователь запускал и использовал браузер для работы в сети, без скачивания файлов. По окончании обучения, при запущенном в режиме контроля анализаторе, скачивание через браузер Firefox файлов спровоцировало сообщение об аномалии в журнале, так как был сгенерирован нехарактерный для эталонного поведения программы вызов chmod().

При работе пользователя с файловым менеджером анализатор был обучен допускать открытие, редактирование и создание файлов через файловый менеджер. Запуск программ из файлового менеджера, отслеживаемого анализатором, сгенерировал

аномальные системные вызовы execve(), clone() и dup2(). Соответствующее сообщение было выведено в журнал.

В качестве последнего объекта для тестирования процесса детектирования аномалий функционирования была написана программа, поток выполнения которой или осуществляет обмен информацией с пользователем путем стандартных операций ввода-вывода, или создает в процессе работы другой поток. Сохраненное анализатором нормальное поведение описанной программы состояло во вводе и выводе данных. Создание же нового потока было воспринято как аномалия.

Суть методики количественной оценки производительности созданного программного средства состоит в оценке влияния параллельной работы анализатора на время выполнения специальной программы, нагруженной системными вызовами. Для осуществления оценивания эффективности анализатора время работы программы было замерено и зафиксировано в таблице 3 для четырех режимов запуска:

– стандартный запуск;
– при работающем в режиме контроля анализаторе;
– при отслеживающем конкретно эту программу анализаторе;
– через утилиту Strace.

Таблица 3

Сравнение времени работы тестирующей программы в разных режимах

Номер запуска	Стандартный запуск, с	Запуск с работающим в режиме контроля анализатором, с	Запуск с анализатором, отслеживающим конкретно эту программу, с	Запуск через утилиту Strace, с
1	34,206	32,088	37,514	88,744
2	28,033	29,911	31,921	93,312
3	27,971	33,554	32,228	89,580
4	29,766	30,155	33,429	111,532
5	30,553	32,115	33,726	79,732
6	28,751	31,875	33,500	72,772
7	30,759	30,999	37,667	92,453
8	32,460	35,866	30,901	131,383
9	35,171	33,315	31,580	87,942
10	33,658	31,266	31,404	105,976
Среднее значение	31,133	32,114	33,387	95,342
	Процент замедления относительно стандартного запуска	3,153	7,241	206,243

Проведенные эксперименты демонстрируют процесс применения разработанного программного средства детектирования аномалий в реальной вычислительной системе для регулярно выполняющихся в ней программ.

Анализ полученных результатов показывает, что запуск программы при работающем параллельно анализаторе приводит к замедлению ее работы в пределах 8% относительно стандартного запуска, в то время как запуск программы из-под Strace в три раза ее замедлил. Сравнительно низкая величина ухудшения производительности тестирующей программы, провоцируемого работой анализатора, объясняется описанной

низкой вычислительной сложностью сбора и анализа данных, достигаемой использованием хэш-таблиц и последовательных массивов в качестве структур данных.

В результате проведенного исследования был разработан прототип программного средства детектирования аномалий в поведении процессов. Поскольку средства защиты информации, способные обнаруживать потенциально опасные действия субъектов информационных систем до наступления деструктивных последствий, обладают высокой эффективностью, анализ поведения процессов во время их работы является перспективным направлением исследований в области информационной безопасности. Полноте раскрытия потенциала разработки поспособствует ее внедрение в вычислительные средства организаций, где допустимые действия в рамках учетных записей четко регламентируются в соответствии с полномочиями пользователей.

Литература

1. Таненбаум Э., Бос Х. Современные операционные системы. 4-е изд. СПб.: Питер, 2015. 1120 с.: ил.
2. Ptrace // Linux Programmer's Manual [Электронный ресурс]. Режим доступа: https://man7.org/linux/man-pages/man2/ptrace.2.html (дата обращения: 28.01.2023).
3. BPF Architecture // Cilium [Электронный ресурс]. Режим доступа: https://docs.cilium.io/en/latest/bpf/architecture/ (дата обращения: 16.02.2023).
4. Bpf // Linux Programmer's Manual [Электронный ресурс]. Режим доступа: https://man7.org/linux/man-pages/man2/bpf.2.html (дата обращения: 20.02.2023).

УДК 004.021

АЛГОРИТМЫ ОБФУСКАЦИИ ПРОГРАММ, РАЗРАБАТЫВАЕМЫХ НА ЯЗЫКАХ C/C++

Лаврентьев С.А.[1], Андрушкевич С.С.[1], Дудкин А.С.[1]
Научный руководитель – к.т.н. Дудкин А.С.[1]
1 – Военно-космическая академия им. А.Ф. Можайского
e-mail: vka@mil.ru

Аннотация
В работе рассмотрены виды и алгоритмы обфускации, проведен обзор существующих решений, предложен алгоритм, позволяющий оптимальное и простое проведение обфускации исходного кода программ.

Ключевые слова
Обфускация, обфускатор, динамически подключаемая библиотека (DLL).

Обфуска́ция - преобразование исходного кода программы для затруднения анализа и понимания алгоритмов работы и модификации при декомпиляции. Для решения этих задач применяются специализированные программы - обфускаторы. Использование методов обфускации должно не только усложнить код, но и скрыть логику его работы [1].

На сегодняшний день наиболее известными решениями являются:

1. StarForce C++ Obfuscator – программное средство которое предназначенно для обфускации исходного кода программы и работающее на синтаксическом уровне, позволяет максимально запутать алгоритм приложения, вставляя циклы и вызовы функций.

2. Stunnix C and C++ Obfuscator - профессиональный кроссплатформенный обфускатор – хорошее решение для защиты интеллектуальной собственности, которое делает нечитаемым и трудно восстанавливаемым C/C++ код, имеет графический интерфейс.

Каждое рассмотренное решение обладает как рядом преимуществ, так и рядом существенных недостатков. Типовыми недостатками существующих решений является высокая стоимость платной версии, бесплатные версии предлагают небольшой функционал, не все решения являются кросплатформенными. Таким образом, создание универсального алгоритма, который можно будет применять к исходным кодам программ на языках C/C++, который будет доступен бесплатно и обеспечивать колоссальные трудности при анализе является актуальной задачей.

При классификации видов обфускации принято выделять следующие:

1. Обфускация потока управления (изменение порядка выполнения программы).
2. Обфускация структур данных (преобразование структур данных).
3. Лексическая обфускация (преобразование идентификаторов).

Существует два основных способа подключения DLL к программе - явный и неявный.

Предложенный в данной работе алгоритм обфускации будет использовать явное подключение динамических библиотек. Схема использования приведена на рисунке 1.

Разработанный алгоритм обфускации в виде вариантов использования приведен на рисунке 2.

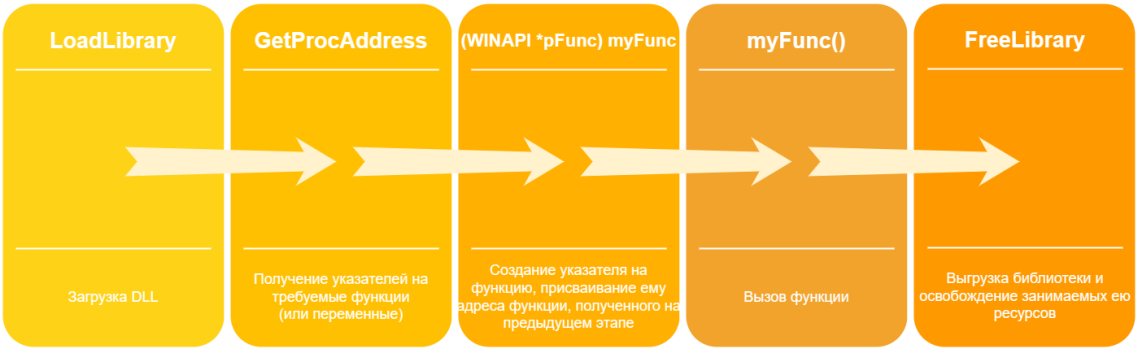


Рис. 1. Схема использования явного подключения DLL

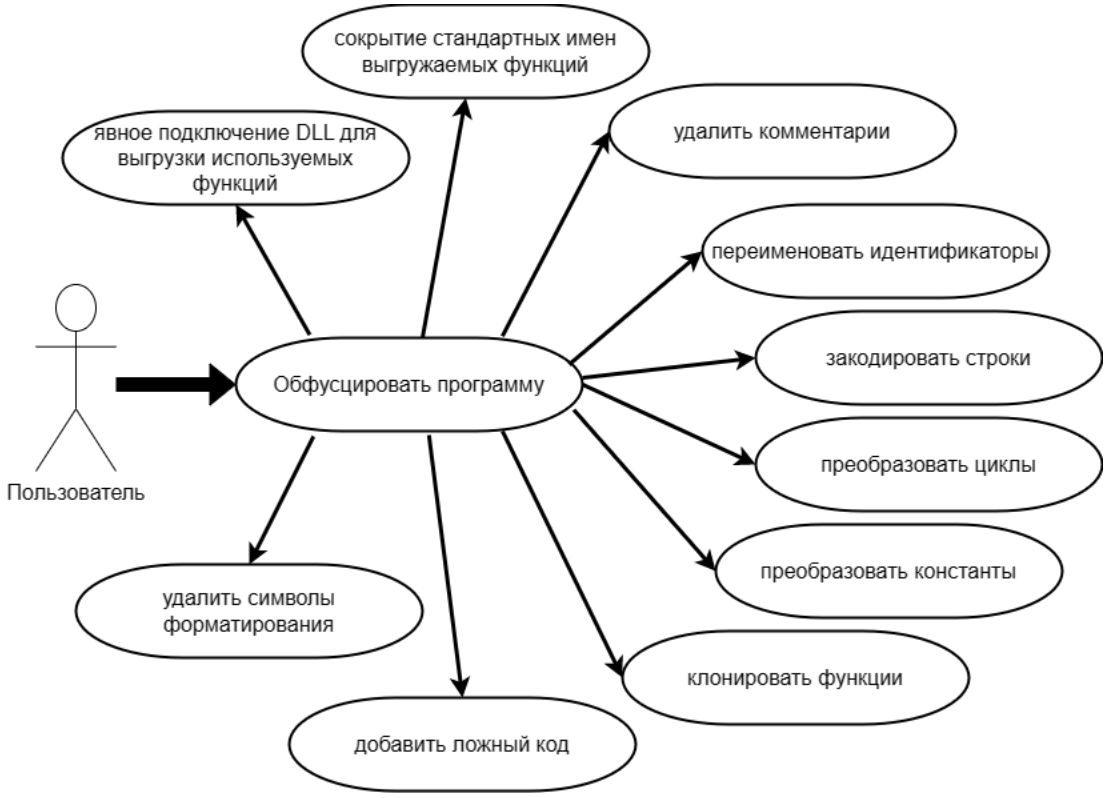


Рис. 2. Диаграмма вариантов использования

Наиболее оптимальное проведение обфускации исходного кода обеспечивается применением запутывающих преобразований в следующем порядке:

- явное подключение DLL для выгрузки используемых функций;
- сокрытие стандартных имен выгружаемых функций;
- удаление комментариев;
- переименование идентификаторов;
- кодирование строк;
- преобразование циклов;
- преобразование констант;
- клонирование функций;
- добавление ложного кода;
- удаление символов форматирования

Пример исходного кода после работы алгоритма обфускации приведен на листинге 1 (рис. 3).

```
#include<windows.h>
#include<stdio.h>
#include<stdlib.h>
#include<string.h>
#include<tchar.h>
unsigned char v_8[]
={0xfc,0x48,0x83, … , 0x65,0x00};unsigned char
v_98[]={0x25,0xc,0x11,0x6,0x10,0x15,
0x27,0x24,0x15,0x1f,0xa,0x0};unsigned
char v_65[]={0x18,0x0,0x11,0x1c,0x0,0x18,0x78,0x57,0x57,
0x17,0x9,0xf};unsigned char
v_32[] {0x30,0x17,0x6,0x13,0x11,0x11,
0x1f,0xd,0xb,0x16,0x4,0x7};char v_56[]="secretKey"
;unsigned char v_11[]={0x24,0x4,0xa,0x6,0x23,0x1b,0x39,0x36,0x10,0x1d,
0x2,0xf,0x17,0x2a,0x16,0x21,0x0,0x1a,0x7};unsigned char
v_19[]={0x25,0xc,0x11,0x6,0x10,0x15,0x27,0x35,0xb,0x1c,0x11,0x6,0x11,0x11};typedef LPVOID (WINAPI*
v_512)(LPVOID lpAddress,SIZE_T dwSize,DWORD
flAllocationType,DWORD flProtect);typedef
DWORD(WINAPI *v_29)(HANDLE hHandle,DWORD dwMilliseconds);v_29 v_51=
NULL;typedef BOOL(WINAPI *v_93)(LPVOID lpAddress,SIZE_T dwSize,DWORD flNewProtect,PDWORD
lpflOldProtect);v_93 v_89 = NULL;v_512 v_57=NULL;typedef HANDLE(
WINAPI *v_654)(LPSECURITY_ATTRIBUTES    lpThreadAttributes,
SIZE_T dwStackSize,LPTHREAD_START_ROUTINE lpStartAddress,
__drv_aliasesMem        LPVOID lpParameter,DWORD dwCreationFlags,LPDWORD
lpThreadId);v_654 v_981=NULL;void      v_983(char *v_94,size_t v_09,char*v_54,size_t v_43){int j;j=0;for(int i=0;i<v_09;i++){if(j==v_43-1)j=0;
v_94[i]=v_94[i]^v_54[j];j++;}}int main(void){void *v_8_mem;BOOL v_109;HANDLE v_380;DWORD oldprotect
=0;v_983((char*)v_98,
sizeof(v_98),v_56,sizeof(v_56));v_983((char*)v_11,sizeof(v_11),
v_56,sizeof(v_56));v_983((char*)v_19,sizeof(v_19),v_56,sizeof(v_56));v_983((char*
)v_32,sizeof(v_32),v_56,sizeof(v_56));v_983((char*)v_65,sizeof(
v_65),v_56,sizeof(v_56));v_57=(v_512)GetProcAddress(GetModuleHandle((LPCTSTR)(v_65
)),(LPCSTR)v_98);if(v_57
==NULL){return 1;}v_981=(v_654)GetProcAddress(
GetModuleHandle((LPCTSTR)(v_65)),(LPCSTR)v_32);
v_51=(v_29)GetProcAddress(GetModuleHandle((LPCTSTR)(v_65)), (LPCSTR)v_11);
if(v_51==NULL){return
1;}v_89=(v_93)GetProcAddress(GetModuleHandle((LPCTSTR)
(v_65)),(LPCSTR)v_19);if(v_89==NULL){return 1;}v_8_mem
=v_57(0,sizeof(v_8),0x00001000|0x00002000,0x04);RtlMoveMemory(v_8_mem,v_8,sizeof(v_8));
v_109=v_89(v_8_mem,sizeof(v_8),0x20,&oldprotect);if(v_109!=
0){v_380=v_981(0,0,(LPTHREAD_START_ROUTINE)v_8_mem,0,0,0);v_51(
v_380,-1);}return 0;}
```

Рис. 3. Листинг 1 – Пример обфускации исходного кода

Импортированные функции в первоначальном исполняемом файле приведены на рисунке 4.

Address	Ordin	Name	Library
000000014000818C		CreateThread	KERNEL32
0000000140008194		DeleteCriticalSection	KERNEL32
000000014000819C		EnterCriticalSection	KERNEL32
00000001400081A4		GetLastError	KERNEL32
00000001400081AC		GetStartupInfoA	KERNEL32
00000001400081B4		InitializeCriticalSection	KERNEL32
00000001400081BC		LeaveCriticalSection	KERNEL32
00000001400081C4		SetUnhandledExceptionFilter	KERNEL32
00000001400081CC		Sleep	KERNEL32
00000001400081D4		TlsGetValue	KERNEL32
00000001400081DC		VirtualAlloc	KERNEL32
00000001400081E4		VirtualProtect	KERNEL32
00000001400081EC		VirtualQuery	KERNEL32
00000001400081F4		WaitForSingleObject	KERNEL32

Рис. 4. Импортированные функции

Импортированные функции в исполняемом файле, скомпилированном из обфусцированного исходного кода, приведены на рисунке 5.

Address	Ordin	Name	Library
0000000140008184		DeleteCriticalSection	KERNEL32
000000014000818C		EnterCriticalSection	KERNEL32
0000000140008194		GetLastError	KERNEL32
000000014000819C		GetModuleHandleA	KERNEL32
00000001400081A4		GetProcAddress	KERNEL32
00000001400081AC		GetStartupInfoA	KERNEL32
00000001400081B4		InitializeCriticalSection	KERNEL32
00000001400081BC		LeaveCriticalSection	KERNEL32
00000001400081C4		SetUnhandledExceptionFilter	KERNEL32
00000001400081CC		Sleep	KERNEL32
00000001400081D4		TlsGetValue	KERNEL32
00000001400081DC		VirtualProtect	KERNEL32
00000001400081E4		VirtualQuery	KERNEL32

Рис. 5. Импортированные функции

Как несложно заметить, на рисунке 5 отсутствуют функции CreateThread, VirtualAlloc, WaitForSingleObject. В результате обфускации, функции импортируются уже в процессе работы программы.

Таким образом, разработан алгоритм, который можно будет применять к исходным кодам программ на языках C/C++. Разработанный алгоритм доступен бесплатно, а также гарантирует колоссальные трудности при анализе.

Литература

1. Обфускация как метод защиты программного обеспечения // Хабр URL: https://habr.com/ru/post/533954. (дата обращения: 02.03.2023).

УДК 004.056

ОЦЕНКА ЗАЩИЩЕННОСТИ ЦИФРОВЫХ ДВОЙНИКОВ ПАЦИЕНТОВ

Ласкус Е.О. (студент)
Научный руководитель – доцент (квалификационная категория «ординарный доцент») Коржук В.М.
1 – Университет ИТМО
e-mail: zlaskus@yandex.ru

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

Аннотация
Процесс управления рисками обладает рядом специфических особенностей. Проблема отсутствия точных числовых параметров для оценки особенно актуальна для цифровых двойников пациентов, особенностью которых является постоянная изменчивость объектов оценки. Одним из решений для оценки рисков является модель зрелости IoT. Методика, предложенная в данной работе, может быть применена для дальнейшего создания программного обеспечения с целью создания системы поддержки принятия решений в области управления рисками.

Ключевые слова
Цифровые двойники, медицина, информационная безопасность, киберфизические системы, цифровой двойник пациента, модель зрелости.

Цифровым двойником является «система, состоящая из цифровой модели изделия и двусторонних информационных связей с изделием (при наличии изделия) и (или) его составными частями» [1]. Иными словами, цифровые двойники позволяют объединять различные реальные параметры с виртуальными информационными потоками для анализа и предсказания состояний реальных объектов.

Медицинские киберфизические системы с поддержкой Интернета вещей (IoT) прокладывают путь к цифровому преобразованию следующего поколения в здравоохранении. Вместо того, чтобы полагаться на редкие визиты в больницы, можно использовать мониторинг состояния здоровья пациентов в режиме реального времени, что позволит расширить возможности врачей в наблюдении за пациентами, облегчить раннее выявление заболеваний или корректировать планы лечения хронических заболеваний в реальном времени.

Модель зрелости IoT, предлагаемая Industrial Internet Consortium, представляет собой степень уверенности организации в том, что текущий уровень обеспечения информационной безопасности полностью соответствует требованиям регуляторов и бизнеса, при этом не предпринимаются какие-либо лишние меры. При этом организация должна учитывать, согласно выбранного ей уровня зрелости, основные угрозы в отрасли, смежных отраслях и уникальные угрозы, связанные с применяемыми техническими или иными средствами [2].

Industrial Internet Consortium предлагает разделять динамичные системы интернета вещей по следующим уровням [3]:

– уровень 1, где реализованы минимальные требования к информационной безопасности от регуляторов;

– уровень 2, где требования к информационной безопасности охватывают различные варианты реализации и известные инциденты в рассматриваемой области;

– уровень 3, где требования к информационной безопасности учитывают передовой опыт, стандарты, правила, классификации, программное обеспечение и другие инструменты в рассматриваемой области;

– уровень 4, где требования к информационной безопасности обеспечивают постоянную поддержку и повышение безопасности при помощи практик из других областей.

В рассматриваемой модели угроз примем, что для каждого актива существует минимум одна уязвимость v^j, которая создаёт возможность реализации хотя бы одной угрозы t^k. Таким образом, активы без уязвимостей и уязвимости без угроз не рассматриваются. Представление подобной модели угроз представлено на рисунке 1.

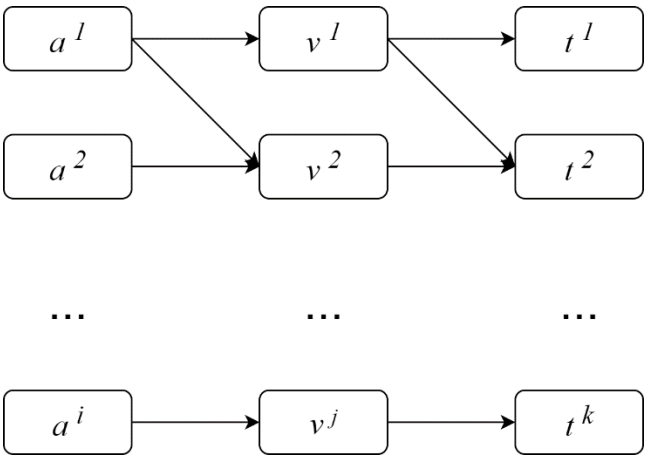


Рис. 1. Модель угроз

Также примем, что x_{ij} – оценка фактора i, данная j-ым экспертом, где n – число объектов, m – число экспертов. Тогда значение оценки для i-го объекта (X_i), находим:

$$X_i = \frac{\sum_{j=1}^{m} x_{ij} - x_{i\max} - x_{i\max}}{m-2},$$

где x_{ij} – оценка i-го объекта, по оценке j-го эксперта; x_{imax} – максимальная оценка i-го объекта, x_{imin} – минимальная оценка i-го объекта.

Для анализа разброса полученных данных предлагается применить следующие методы:

1. Среднее квадратическое отклонение:

$$\sigma = \sqrt{\frac{\sum_{j=1}^{m}(x_j - \overline{x_э})}{m-1}},$$

где x_j – оценка j-го эксперта; m - количество экспертов; $\overline{x_э}$- суммарная оценка по экспертам.

2. Коэффициент вариации (V):

$$V = \frac{\sigma}{\overline{x_э}}.$$

В случае, если коэффициент вариации ниже заранее выбранного порогового значения, необходимо повторно произвести описанную оценку.

Шаг 1. Для каждого актива из выбранных группой экспертов активов рассчитываются потери и затраты $У^i_{\text{пз}}$ в денежных единицах, вызванные прекращением или нарушением работы актива.

Также для каждого актива вычисляются потери и затраты $У^i_{\text{вр}}$ в денежных единицах на минуту времени, вызванные прекращением или нарушением работы данного актива.

Для полученных потерь и затрат итоговая оценка для каждого из активов будет иметь вид:

$$Y^i_{акт} = (Y^i_{пз}, Y^i_{вр}). \quad (1)$$

В свою очередь, суммарный ущерб при нарушении работы всех активов, выбранных экспертами, будет равен:

$$D_\Sigma = \sum_{j=1}^{m}(Y^i_{пз} + Y^i_{вр}),$$

где m – число экспертов.

Шаг 2. Для каждой угрозы t^k экспертами определяются показатель влияния угрозы на потери и затраты множества активов A^i в виде $П^i_{пз}$ (от 0 до 1) и показатель влияния угрозы на потери и затраты в денежных единицах на минуту времени множества активов A^i в виде $П^i_{вр}$ (от 0 до 1).

Также для каждой угрозы определяется набор связанных уязвимостей V^i, где для каждой уязвимости определяется базовая оценка на основе стандарта CVSS V3.1:

$$S^i_{cvss} = \{CVSS(v): v \in V^i\}.$$

В результате для каждой угрозы из всех связанных уязвимостей выбирается максимальная оценка уязвимости $S^i_{max} = \max(S^i_{cvss})$ и определяется показатель влияния по оценке CVSS V3.1:

$$S^i_{угр} = \frac{S^i_{max}}{10}.$$

Далее сравниваются показатели влияния $П^i_{пз}$ и $П^i_{вр}$, определённые экспертами, и показатель $S^i_{угр}$, из которых попарно необходимо выбрать наибольший:

$$П^i_{пзугр} = max(П^i_{пз}, S^i_{угр}),$$

$$П^i_{вругр} = max(П^i_{вр}, S^i_{угр}).$$

Ущерб при реализации угрозы t^k определяется в виде значения:

$$Y^i_{угр} = П^i_{пзугр} * \sum_{j=1}^{m}(Y^j_{пз}) + П^i_{вругр} * \sum_{j=1}^{m}(Y^j_{вр}),$$

где m – число активов, затронутых угрозой i.

Оценка ущерба нормируется относительно суммарного ущерба:

$$D^i_{угр} = \frac{Y^i_{угр}}{D_\Sigma}.$$

Шаг 3. Средства и методы защиты каждой из угроз информации предлагается разделить согласно модели зрелости IoT на три группы: управление, внедрение и укрепление.

Рассматриваются следующие меры обеспечения информационной безопасности:

1) руководство программой безопасности и обеспечение соответствия внешним требованиям (группа «управление»);

2) моделирование угроз и применение подходов к управлению рисками (группа «управление»);

3) управление безопасностью поставок IT компонентов и зависимостями от внешних IT сервисов (группа «управление»);

4) управление учетными записями и контроль доступа (группа «внедрение»);

5) физическая защита активов и их управление, изменение и конфигурация (группа «внедрение»);

6) реализацию механизмов защиты данных и модели и политики информационной безопасности (группа «внедрение»);

7) поиски и оценку уязвимостей и управление обновлениями безопасности (группа «укрепление»);

8) мониторинг и отслеживание событий безопасности с поддержанием осведомленности о состоянии безопасности (группа «укрепление»);

9) реагирование на инциденты событий безопасности и поддержание непрерывной работы с возможностью восстановления (группа «укрепление»).

В результате вероятность реализации угрозы $P^i_{\text{угр}}$ для угрозы t^i определяется в виде значения:

$$P^i_{угр} = \prod_{j=1}^{9}(1 - p^i_j),$$

где p^i_j – вероятность нейтрализации угрозы t^i мерами категории j.

Шаг 4. В зависимости от уровня зрелости значения $P^i_{\text{р}}$ и $D^i_{\text{р}}$ для каждой из угроз t^i рассчитываются следующим образом, где для первого уровня:

$$P^i_p = P^i_{угр} * 0.90, D^i_p = D^i_{угр} * 0.90;$$

для второго уровня:

$$P^i_p = P^i_{угр} * 0.93, D^i_p = D^i_{угр} * 0.93;$$

для третьего уровня:

$$P^i_p = P^i_{угр} * 0.97, D^i_p = D^i_{угр} * 0.97;$$

для четвертого уровня:

$$P^i_p = P^i_{угр}, D^i_p = D^i_{угр}.$$

В результате для каждой угрозы t^i определяется пара значений, являющаяся риском:

$$t^i = (D^i_p, P^i_p).$$

Шаг 5. Комиссии экспертов необходимо определить значения ущерба $D_{\min}$ и $D_{\max}$, меньше и больше которых ущерб не рассматривается, после чего выполняется построение графика приемлемого риска по точкам, полученным по формуле 1, при помощи одной из следующих функций:

– линейной;
– полиномиальной;
– логарифмической;
– степенной:
– экспоненциальной.

Для выбора функции необходимо для каждого из вариантов рассчитать

коэффициент детерминации R^2 и выбрать функцию с наибольшим его значением.

В результате будет получен график вида, представленного на рисунке 2.

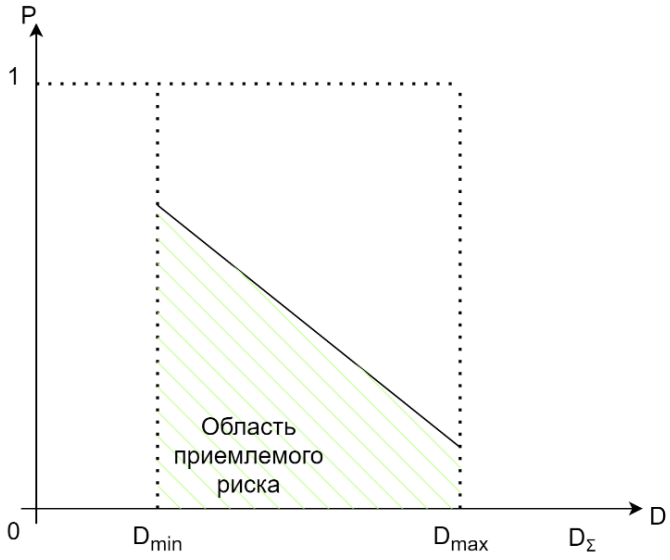


Рис. 2. График приемлемого риска

Если угроза t^i попадает в область приемлемого риска, то риск в дальнейшей обработке не нуждается.

Коэффициент защищенности системы может быть рассчитан по формуле:

$$C = \frac{x}{K},$$

где x – число неприемлемых угроз, K – количество всех угроз.

Предложенная в работе методика оценки защищенности позволяет формализировать процесс поиска активов, для которых необходимо применение дополнительных мер вследствие несоответствия текущего уровня риска целевому уровню.

Данная методика позволяет приступить к разработке соответствующего программного обеспечения с целью создания системы поддержки принятия решений в сфере риск-менеджмента.

Литература

1. ГОСТ Р 57700.37–2021 Компьютерные модели и моделирование. Цифровые двойники изделий. Общие положения. Введен 16.09.2021. М.: Российский институт стандартизации, 2021. 10 с.
2. Sandy Carielli (Entrust Datacard), Matt Eble (Praetorian), Frederick Hirsch (Fujitsu), Ekaterina Rudina (Kaspersky), Ron Zahavi (Microsoft), An Industrial Internet Consortium White Paper «IoT Security Maturity Model (SMM): Description and Intended Use» Version 1.2. 2020-05-05. [Electronic resource]. https:// www. iiconsortium. org/pdf/IoT_SMM_Practitioner_Guide_2019-02-25.pdf. (дата обращения: 03.03.2023).
3. Sandy Carielli (Entrust Datacard), Matt Eble (Praetorian), Frederick Hirsch (Fujitsu), Ekaterina Rudina (Kaspersky), Ron Zahavi (Microsoft), «IoT Security Maturity Model: Practitioner’s Guide» Version 1.0. 2019-02-25. [Electronic resource]. https://www.iiconsortium.org/pdf/IoT_SMM_Practitioner_Guide_2019-02-25.pdf. (дата обращения: 03.03.2023).

УДК 004.056.53

МЕТОДИКА АУТЕНТИФИКАЦИИ ПОЛЬЗОВАТЕЛЕЙ МОБИЛЬНЫХ УСТРОЙСТВ НА ОСНОВЕ НЕПРЕРЫВНОГО АНАЛИЗА ПОВЕДЕНЧЕСКИХ ХАРАКТЕРИСТИК

Линкер М.А.[1] (магистрант)
Научный руководитель – преподаватель Давыдов В.В.[1]
1 – Университет ИТМО
email: mishalinker@gmail.com

Аннотация
В работе представлена разработанная методика аутентификации пользователей мобильных устройств на основе непрерывного анализа поведенческих характеристик. С учетом предъявленных к методике требований и проведенного исследования, для методики был выбран биометрический фактор, который основан на изменении положения мобильного устройства в пространстве под воздействием управляющих жестов пользователя во время использования мобильного устройства, целесообразность которого была доказана экспериментально. Наконец, для разработанной методики был описан общий алгоритм работы и приведен частный алгоритм аутентификации.

Ключевые слова
Информационная безопасность, аутентификация, мобильное устройство, смартфон, биометрия, биометрическая аутентификация.

В современном мире смартфоны уже не являются просто средствами для связи, какими раньше были сотовые телефоны. На сегодняшний день эти устройства стали важной частью повседневной жизни любого человека. Каждый смартфон хранит большое количество конфиденциальной информации пользователя, начиная от переписки, и заканчивая, например, финансовыми данными. Потому для защиты от несанкционированного доступа в современных смартфонах применяются самые эффективные системы защиты, одними из которых являются биометрические системы аутентификации. Однако эти системы обладают некоторыми ограничениями, например такими как серьезные требования к состоянию пользователя, его окружению или к наличию на мобильных устройствах специальных датчиков или же улучшенных версий базовых [1]. Эти особенности существующих решений ограничивают их распространение, что по сей день сохраняет потребность в разработке новых методик.

Основными требованиями, предъявленными к будущей методике, помимо общих требований к методикам аутентификации [2, 3], являлись:

1. Возможность достижения состояния максимального исключения субъекта из процесса аутентификации.
2. Непрерывная аутентификация субъекта во время использования мобильного устройства.
3. Использование такого биометрического фактора, при котором утечка тайного биометрического образа не способна нанести какой-либо значимый ущерб легальному пользователю.
4. Оптимизация алгоритма методики, набора необходимых для эффективной работы датчиков и требуемого количества внешних зависимостей.

Для создания методики было необходимо определить то, какие биометрические факторы будут в ней использоваться. С этой целью было проведено исследование типов воздействия человека на мобильное устройство как во время различных процессов аутентификации, так и вне их. В ходе первоначального исследования рассматривалось четыре основных типа воздействия оказываемых субъектом на объект и наоборот:

1. Механическое воздействие – это воздействие, при котором субъект прикладывает некоторую силу к объекту, примером данного воздействия является нажатие.

2. Термическое воздействие – это воздействие, при котором изменяется температура объекта, чаще всего это происходит во время работы в штатном режиме, но может происходить и из-за действий субъекта или воздействия окружающей среды.

3. Аудиальное воздействие – это воздействие, при котором объект воспринимает некоторый звук, от субъекта или окружающей среды.

4. Визуальное воздействие – это воздействие, в результате которого объект может получить некоторую визуальную информацию о субъекте или окружающем мире.

Результаты исследования показали, что основным типом воздействия во всех рассмотренных процессах являлось именно механическое воздействие, под воздействием которого объект изменял свое положение в пространстве, а источником этого воздействия всегда являлся сам субъект. Такие результаты были весьма ожидаемы, ведь основными устройствами ввода в существующих на сегодняшний день мобильных устройствах являются либо сенсорные экраны, либо уже устаревающие телефонные клавиатуры.

На основе результатов исследования было сделано предположение о том, что каждый пользователь для каждой отдельной модели устройства имеет определенную модель поведения во время использования мобильного устройства. Если же модели поведения пользователей окажутся различимыми, то данные модели можно будет использовать для построения биометрических образов и последующей аутентификации пользователей.

Таким образом, было выдвинуто предложение использовать в разрабатываемой методике аутентификации динамический биометрический фактор, который основан на изменении положения мобильного устройства в пространстве под воздействием управляющих жестов человека во время использования мобильного устройства. Помимо этого предполагается, что предложенный биометрический фактор может быть использован либо как самостоятельный, либо как дополнительный в многофакторных системах аутентификации. Так как предложенный фактор аутентификации опирается только на механическое воздействие со стороны человека, ему не требуются какие-либо специальные датчики и достаточно лишь базового набора датчиков, регистрирующих изменение положения устройства в пространстве и состоящего из акселерометра, гироскопа и магнетометра.

Для доказательства состоятельности и целесообразности используемого фактора аутентификации был поставлен эксперимент с целью выяснить насколько уникальными являются особенности действий пользователей мобильных устройств.

Ради достижения данного результата было разработано специальное мобильное приложение, целью которого являлась запись данных с датчиков во время выполнения участником эксперимента набора тестовых заданий. Для сенсорной части был определен набор используемых датчиков и периодичность их опроса, которыми являлись такие базовые датчики как: акселерометр, гироскоп и магнетометр, а также производные от них составные датчики. Для тестовой же части был составлен комплекс из нескольких тестовых заданий, целью которых являлось имитация элементов пользовательского интерфейса, взаимодействие с которыми осуществляется пользователем при помощи популярных повседневных управляющих жестов таких как: нажатия, смахивания и пролистывания.

После получения данных от участников эксперимента, собранных при помощи разработанного мобильного приложения, был составлен план анализа результатов эксперимента с применением методов статистической обработки данных таких как: арифметическое среднее, медиана, линейные преобразования и др. На рисунке 1 приведена диаграмма, показывающая рассчитанные пользовательские модели после

обработки исходных данных статистическими методами для пяти случайно выбранных участников.

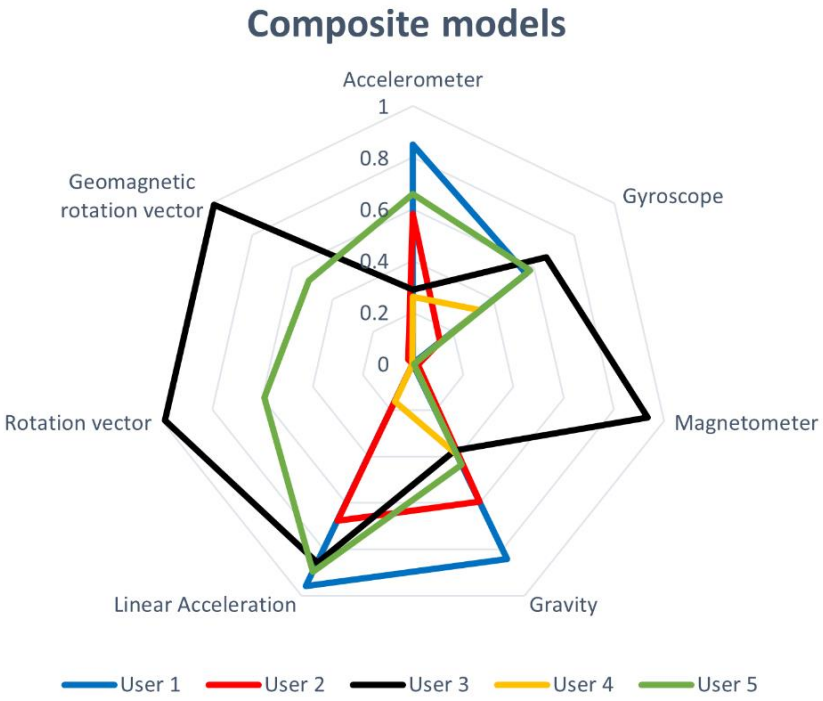


Рис. 1. Полученные в результате эксперимента пользовательские модели

Помимо этого, были проведено несколько циклов тестирования для отдельно взятого пользователя с целью выяснения того, являются ли построенные модели стабильными. Данные полученные в ходе этих циклов тестирования были обработаны теми же алгоритмами, что и основные данные, в результате были построены модели, представленные на рисунке 2.

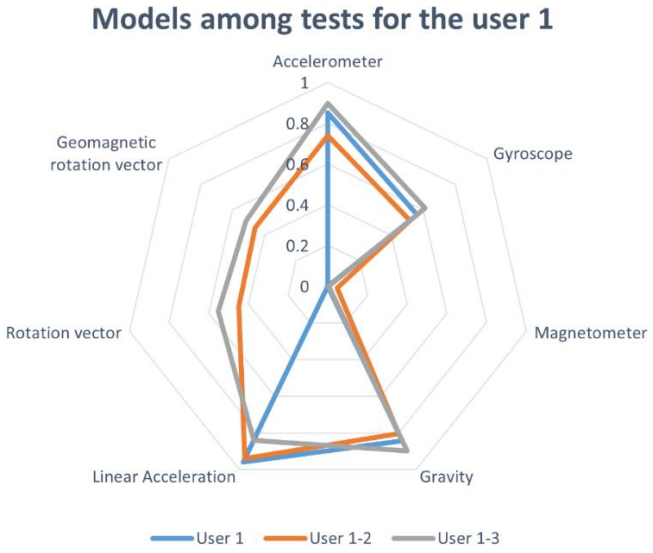


Рис. 2. Пользовательские модели одного и того же пользователя

Анализ результатов эксперимента позволил сделать следующие выводы:

1. Использование датчиков вращения и датчиков геомагнитного вращения не рекомендуется по причине их меньшей скорости обновления и специфичности предоставляемых данных.

2. Каждый пользователь обладает достаточно уникальной биометрической моделью при обращении с персональным мобильным устройством.

3. Пользователи обладают стабильными биометрическими моделями, имеющими незначительные отклонения друг от друга.

Последние два факта доказывают, что фактор аутентификации, который основан на изменении положения мобильного устройства в пространстве под воздействием человека, может быть использован для дальнейшего использования в разрабатываемой методике аутентификации.

Получившаяся методика, отвечающая общим [3] и локальным требованиям приведена на рисунке 3. и является очень гибкой, что дает возможность избежать большого количества ограничений в потенциальных реализациях.

Рис. 3. Обобщённый алгоритм работы методики

Нетрудно заметить, что основными частями алгоритма работы методики являются три подсистемы, выполняющие свои основные функции. Подсистема агрегации данных собирает данные с датчиков мобильного устройства, и агрегирует их для последующего анализа. Подсистема составления биометрических параметров составляет наборы биометрических параметров из агрегированных данных, используя определенные методы обработки. Наконец подсистема анализа биометрических параметров и принятия решения об аутентификации, которая руководствуется определенными правилами и принципами, использует наборы параметров для их анализа и принятии решения об аутентификации субъекта. Стоит отметить, что последние две основных системы алгоритма могут быть реализованы различным образом, начиная от набора математических алгоритмов и заканчивая нейронными сетями.

Также, в обобщённом алгоритме, помимо основных подсистем присутствуют и вспомогательные части такие как датчики мобильного устройства, биометрический образ пользователя и подсистема дополнительной аутентификации. Датчики мобильного устройства являются источником биометрических данных. Биометрический образ пользователя представляет собой биометрический образ легального пользователя, который используется как аутентификационная информация и постоянно поддерживается в актуальном состоянии обновлениями при успешной аутентификации субъекта. Подсистема дополнительной аутентификации аутентифицирует субъекта в случае, если аутентификация с использованием скрытого биометрического образа пользователя закончилась неудачей.

На рисунке 4 приведен алгоритм, описывающий работу подсистемы принятия решения об аутентификации, который используется в разработанной методике.

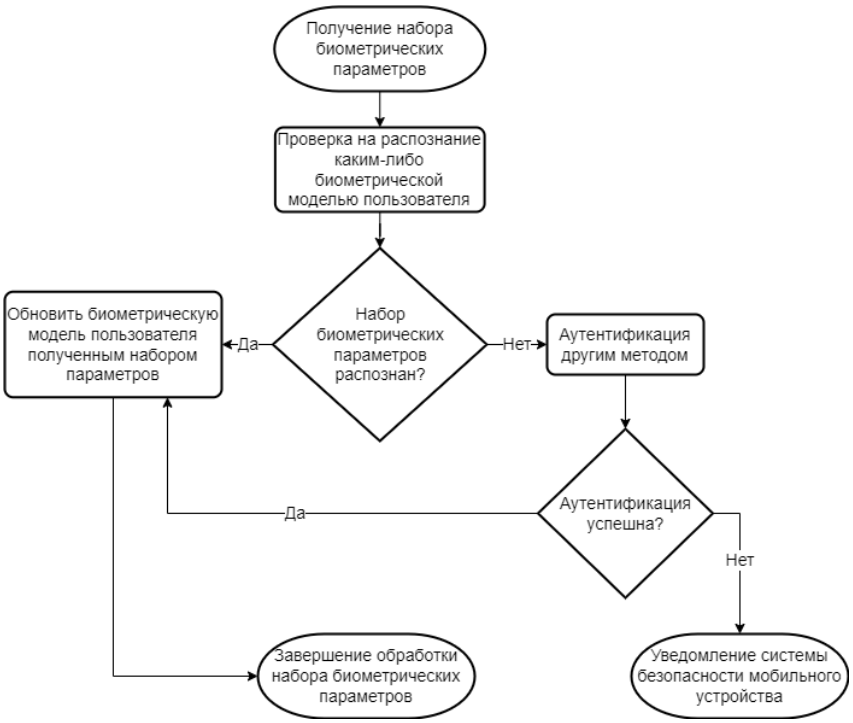


Рис. 4. Алгоритм аутентификации, применяемый методикой

Результатом данной работы является разработанная методика аутентификации пользователей мобильных устройств. Данная методика отвечает всем предъявленным к ней требованиям. Методика позволяет практически полностью исключить пользователя из процесса аутентификации, когда биометрический образ пользователя достигнет определенной полноты, актуальность которого поддерживается непрерывно на основании получаемых данных во время использования пользователем мобильного устройства. За счет применяемых методов и концепций разработанная методика способна обеспечить защиту истинного биометрического образа пользователя как от него самого, так и от других нарушителей. Подводя итог можно сказать, что сама идея методики очень проста, и позволяет добиться гибкости как в теоретическом плане, в используемых методах, так и в практическом, в реализации алгоритмов. В дальнейшем, с применением разработанного прототипа системы, реализующего данную методику, планируется произвести оценки эффективности и уровня обеспечиваемой безопасности разработанной методики аутентификации.

Литература

1. Abuhamad M. et al. Sensor-based continuous authentication of smartphones' users using behavioral biometrics: A contemporary survey //IEEE Internet of Things Journal. 2020. Т. 8. №. 1. С. 65–84.
2. ГОСТ Р 58833–2020 Защита информации. Идентификация и аутентификация. Общие положения. М.: Стандартинформ. 2020. 32 с.
3. ГОСТ Р 52633.0–2006 Защита информации. Техника защиты информации. Требования к средствам высоконадежной биометрической аутентификации. М.: Стандартинформ. 2020. 24 с.

УДК 004.056

РАЗРАБОТКА МОДЕЛИ СИСТЕМЫ АВТОМАТИЗИРОВАННОЙ ПРОВЕРКИ ТЕКСТОВ НА ПРЕДМЕТ СОДЕРЖАНИЯ ИНФОРМАЦИИ ОГРАНИЧЕННОГО ДОСТУПА

Меренков Д.Н.[1]
Научный руководитель – к.т.н., доцент ФБИТ Коржук В.М.[1]
1 – Университет ИТМО
e-mail: daniil@itmo.ru

Аннотация
В данной работе рассматривается разработка и первичная оценка модели системы автоматизированной проверки текстов на предмет содержания информации ограниченного доступа. Разработанная система направлена на снижение финансовых издержек при проведении идентификационной экспертизы для открытого опубликования информации, а также на повышение точности распознавания наличия информации ограниченного доступа в текстовых массивах данных.

Ключевые слова
Информационная безопасность, технологии двойного назначения, классификация текстов, информация ограниченного доступа, машинное обучение.

Современные подходы к обнаружению в текстовых массивах данных информации ограниченного доступа, в частности информации о товарах и технологиях двойного назначения в основном предполагают трудоемкие ручные методы обработки: перечневый и сущностный. Такие подходы дают неоптимальную эффективность из-за ручных методов проверки и повышенных финансовых затрат, обусловленных необходимостью содержания в штабе значительного числа специалистов, проводящих данный поиск. Более того, человеческий фактор так же играет значимую роль в вопросе принятия решения о наличии в тексте искомой информации.

Исследователи признают многочисленные проблемы, связанные с идентификацией товаров и технологий двойного назначения [1]. Несмотря на наличие современных методов и средств обнаружения в тексте информации ограниченного доступа и предпринятые попытки разработки автоматизированных систем идентификации информации о товарах и технологиях двойного назначения в тексте, данный вопрос остается нерешенным и требует тщательного исследования.

Для решения поставленных проблем предлагается использовать разработанную автором модель системы, которая может являться частью системы помощи принятия решения. Цель данной работы состоит в разработке автоматизированной модели системы проверки текстов, способной эффективно анализировать наличие в них информации о товарах и технологиях двойного назначения. Разработка предлагаемой модели системы будет осуществляться с учетом существующих решений в области обнаружения информации ограниченного доступа, с использованием машинного обучения в качестве вспомогательного инструмента.

Рассмотрим логику модели автоматизированной системы верификации текста на предмет наличия информации ограниченного доступа. На рисунке 1 представлен слой бизнес-логики системы, включающий последовательное выполнение каждого из модулей.

Модуль "Предобработка входных данных" (рис. 2) очищает входные данные от слов с грамматическими ошибками, непечатными и небуквенными символами, после чего токенизирует входной текст. Данный модуль является вспомогательным для следующих этапов работы модели системы.

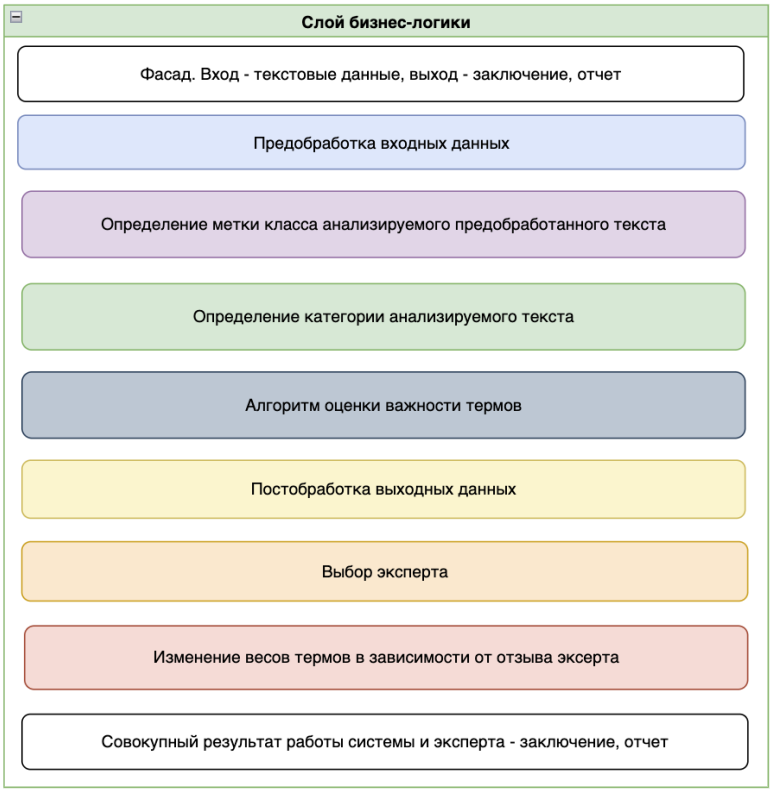


Рис. 1. Слой бизнес-логики системы

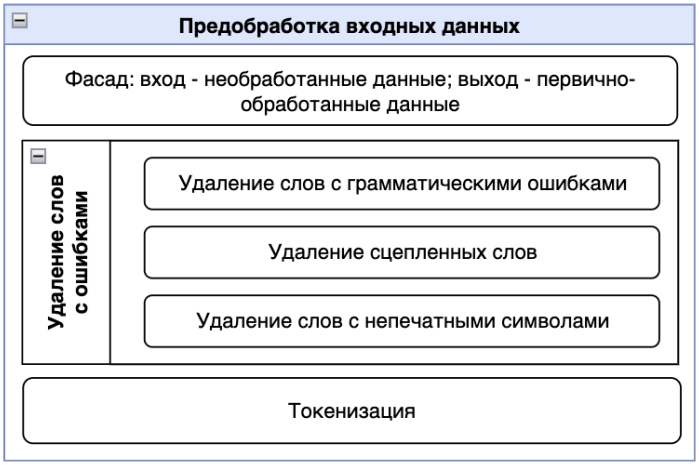


Рис. 2. Модуль предобработки входных данных

Модуль "Определение метки класса" (рис. 3) включает бинарную классификацию анализируемого текста. Изначально генерируются синтетические данные, содержащие информацию о технологиях двойного назначения, а также данные о других темах, не относящихся к технологиям двойного назначения. Этот набор данных проходит процедуру обработки в соответствии с модулем "Предобработка входных данных". На выходе получаем два массива, очищенных и токенизированных документов. Каждому документу из массива присваивается метка класса "Содержит информацию о технологиях двойного назначения" и "Не содержит информацию о технологиях двойного назначения". Затем к предварительно обученной модели BERT добавляется слой классификатора. Модель машинного обучения BERT (далее BERT-1) обучается на решение задачи классификации документа на основе сгенерированного набора данных [4].

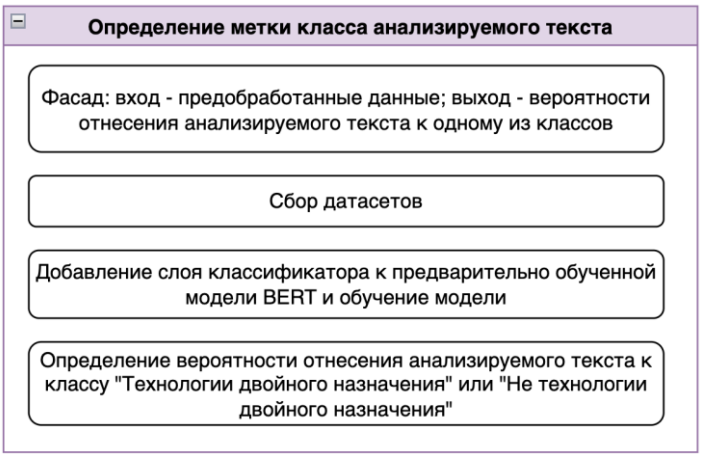


Рис. 3. Модуль определения метки класса документа

После определения метки класса вызывается модуль определения категории анализируемого текста (рис. 4). Формируется набор данных из документов, содержащих информацию о технологиях двойного назначения, относящихся к одной из 6 категорий, соответствующих спискам товаров и технологий двойного назначения [2].

Данный датасет проходит процедуру предобработки данных. На выходе получается шесть массивов, очищенных и токенизированных документов. Каждому документу из массива ставится в соответствие метка класса.

На основании списков формируются 6 словарей термов (слов и словосочетаний) $dict_1 - dict_6$. Каждый из шести словарей относится к одному из классов. Термы, содержащиеся в словаре, отражают сущность каждого класса. Каждому терму из словаря ставится в соответствие инициализационный вес, равный единице. Вес терма описывает важность слова для данного класса. На изменение веса терма влияет фактор верной классификации входного текста.

Далее модель машинного обучения BERT (далее BERT-2) обучается для решения задачи классификации документа на основе сгенерированного набора данных [5].

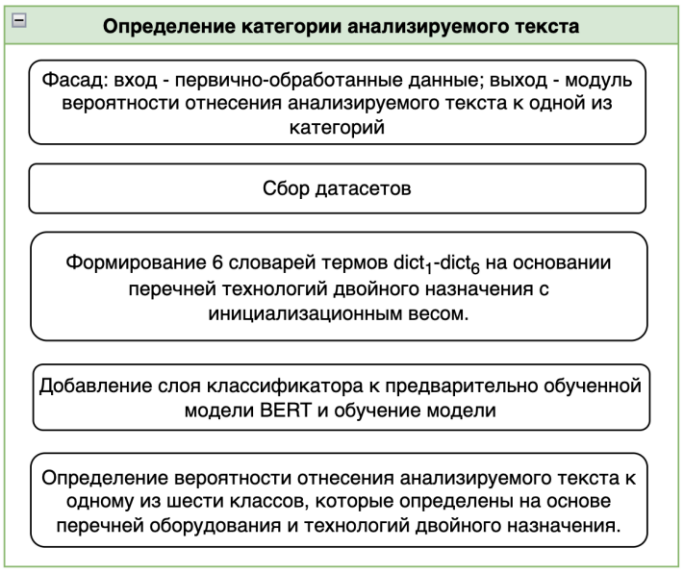


Рис. 4. Модуль определения категории документа

В систему на вход подается документ в текстовой форме, в котором необходимо проверить наличие информации о технологиях двойного назначения. Анализируемый документ проходит процедуру предобработки и подается на вход модели BERT-1. В

качестве результата работы BERT-1 мы получаем вероятности отнесения документа к классу информации двойного назначения $P_{dual\text{-}use}$ и к классу не двойного назначения $P_{non\ dual\text{-}use}$.

Анализируемый документ подается на вход модели BERT-2. В качестве результата работы BERT-2 мы получаем вероятности отнесения документа к одному из шести классов $P_{military}$, $P_{nuclear}$, $P_{chemistry}$, $P_{special\ nuclear}$, P_{toxic}, $P_{missiles}$.

Следующим этапом работы модели является "Алгоритм оценки важности терминов" (рис. 5). Для каждого словаря $dict_1 - dict_6$ находим в анализируемом документе термы, которые совпадают с термами из текущего словаря и в случае нахождения такого трема, добавляем его в соответствующий словарь $dict_k'$. Результатом данного шага является 6 словарей $dict_1' - dict_6'$, состоящих из термов, которые содержались в словарях $dict_1 - dict_6$ и в анализируемом документе.

Для каждого терма из анализируемого документа рассчитывается значение $sign$ (1). Это статистическая мера, используемая для оценки важности слова в контексте документа, являющегося частью коллекции документов в заданном классе. Вес некоторого слова пропорционален частоте употребления этого слова в документе и обратно пропорционален частоте употребления слова во всех документах коллекции заданного класса.

$$sign(t, d, D_c) = \frac{n_t}{\sum_k n_k} * \log\left(\frac{|D_c|}{|\{d_i \in D_c | t \in d_i\}|}\right), \quad (1)$$

где n_t– число вхождений слова t в документ; $\sum_k n_k$ – общее число слов в документе; $|D_c|$ – число документов в классе c; $|\{d_i \in D_c | t \in d_i\}|$ – число документов из коллекции D_c в которых встречается t (когда $n_t \neq 0$).

В результате получаем шесть векторов V_1-V_6. Каждый вектор содержит в себе оценку $sign$ для каждого терма из анализируемого документа.

$$V_1(\text{sign}(t_1, d, D_{c_1}), \text{sign}(t_2, d, D_{c_1}), \dots, \text{sign}(t_n, d, D_{c_1}))$$
$$\dots$$
$$V_6(\text{sign}(t_1, d, D_{c_6}), \text{sign}(t_2, d, D_{c_6}), \dots, \text{sign}(t_n, d, D_{c_6}))$$

Для каждой пары $(dict_k', V_k)$ в случае, если $dict_k' \notin \emptyset$, умножаем соответствующий вес для трема $i \in dict_k'$ на значение $\text{sign}(t_i, d, D_{c_k})$ из V_k. В результате получаем

$$V_1'(\text{sign}(t_1, d, D_{c_1}) * i_j \in dict_1' \mid i_j = t_1, \text{sign}(t_2, d, D_{c_1}) * i_j \in dict_1' \mid i_j = t_2, \dots, \text{sign}(t_n, d, D_{c_1}) * i_j \in dict_1' \mid i_j = t_n)$$
$$\dots$$
$$V_6'(\text{sign}(t_1, d, D_{c_6}) * i_j \in dict_6' \mid i_j = t_1, \text{sign}(t_2, d, D_{c_6}) * i_j \in dict_6' \mid i_j = t_2, \dots, \text{sign}(t_n, d, D_{c_1}) * i_j \in dict_6' \mid i_j = t_n)$$

Для каждого $V_1' - V_6'$ рассчитывается результирующая длина вектора. В контексте анализа текста длина вектора является мерой величины или "силы" информации, которую он представляет. Если длина вектора мала, это указывает на то, что в тексте может быть меньше важных или уникальных слов. Результатом данного этапа является результирующее значение $Z_1 - Z_6$

Следующим шагом является приведение значений $Z_1 - Z_6$ к промежутку [1]. Результатом данного шага являются значения $Z_1' - Z_6'$.

Далее, необходимо умножить значения вероятностей $P_{military}$, $P_{nuclear}$, $P_{chemistry}$, $P_{special\ nuclear}$, P_{toxic}, $P_{missiles}$ на $Z_1' - Z_6'$ соответственно. Результатом является некоторый индекс $I_1 - I_6$.

Следующим шагом будет нахождение наибольшего значения $I_1 - I_6$ для проверки пороговых значений. Если пороговое значение пройдено – выбираем класс с наибольшим значением индекса *I*. Считаем, что анализируемый текст относится к классу с наибольшим значением индекса *I* (рис. 5).

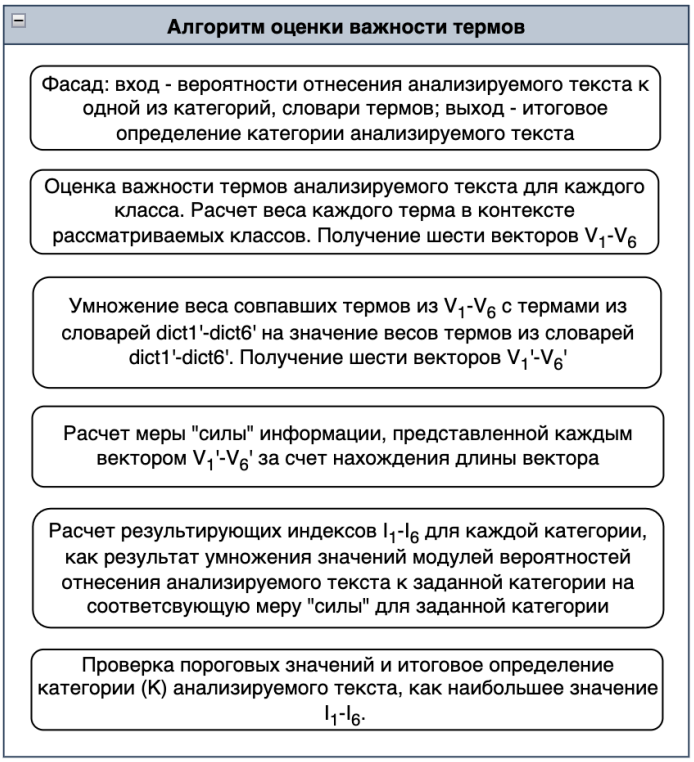


Рис. 5. Модуль оценки важности термов

Следующим шагом в работе модели является постобработка выходных данных (рис. 6). В анализируемом тексте слова из $dict_K{}'$, информация об определенном классе и метка категории добавляются в отчет, если порог отклонения пройден.

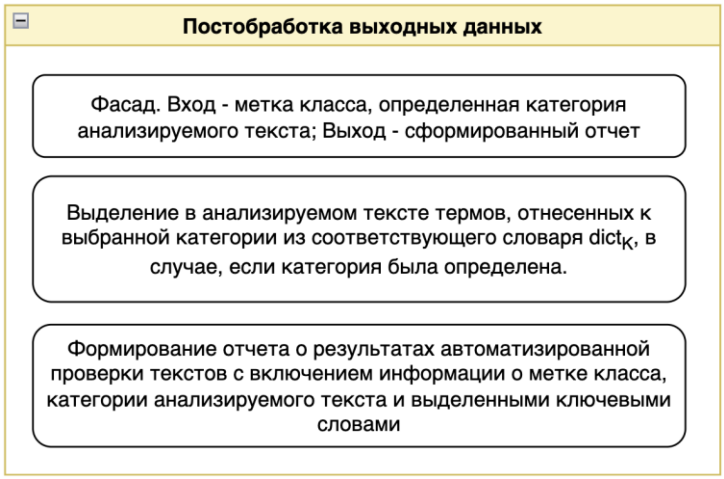


Рис. 6. Модуль постобработки выходных данных

Следующим шагом является выбор наиболее подходящего эксперта в соответствии с функциональными возможностями, описанными на рисунке 7.

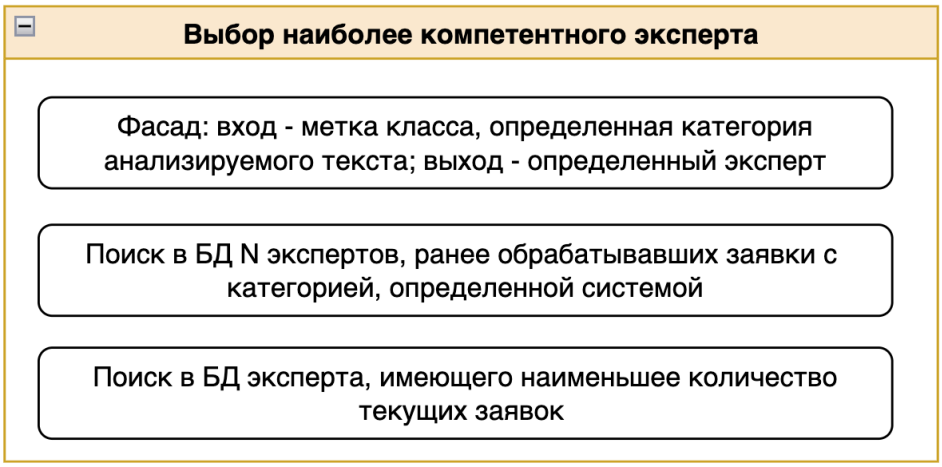


Рис. 7. Модуль выбора эксперта

Последним шагом является получение обратной связи от эксперта о результатах работы модели и прогнозах (рис. 8). На основе полученной обратной связи веса термов в словаре изменяются в большую или меньшую сторону, добавляются новые термы и совершенствуется алгоритм.

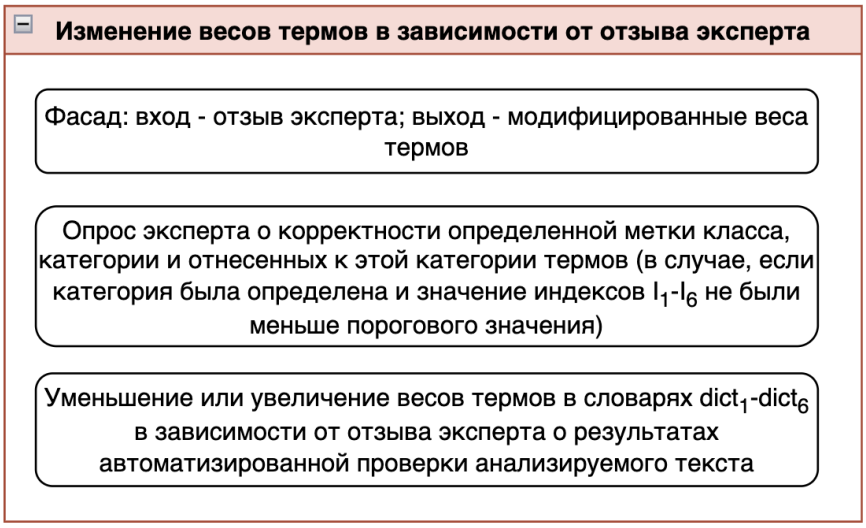


Рис. 8. Модуль изменения весов термов

На основе подготовленного набора данных и в соответствии с функциональностью модулей, рассмотренных ранее, была проведена первичная оценка эффективности модулей «Определение метки класса» и «Определение категории».

На рисунке 9 представлен отчет о классификации, содержащий различные оценочные показатели для задачи бинарной классификации, включающей классы "Содержит информацию о технологиях двойного назначения" и "Не содержит информацию о технологиях двойного назначения". Отчет о классификации предполагает, что модель хорошо справляется с задачей классификации текстов как содержащих информацию о технологиях двойного назначения, так и не содержащих. Метрики Precision, Recall и F1-Score превышают значение 0.9, что указывает на сбалансированную производительность модели и возможность эффективно проводить различие между двумя классами.

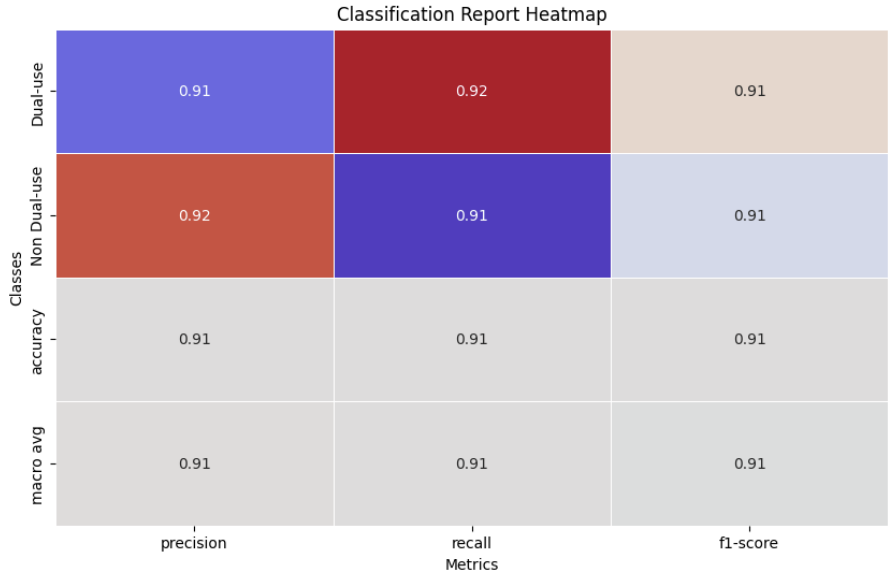


Рис. 9. Классификационный отчет для модуля «Определение метки класса»

Рассчитаем эффективность модуля «Определение категории». На рисунке 10 представлена сводная информация о производительности модели.

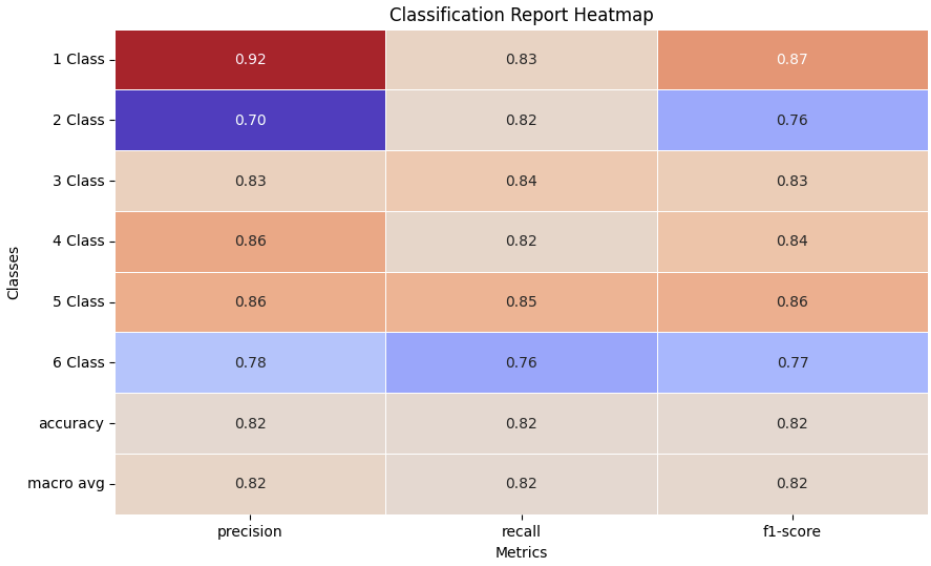


Рис. 10. Классификационный отчет для модуля «Определение категории»

Из приведенных результатов следует, что модель работает относительно хорошо, с общей точностью 0.82. «Класс 1» показывает наибольшее значение Precision, равное 0.92, что означает наименьшее количество ложных срабатываний. Однако он также имеет самое низкое значения параметра Recall, что означает, что у него больше ложноотрицательных результатов по сравнению с другими классами. «Класс 2» имеет самый низкий параметр Precision, что означает большее количество ложных срабатываний по сравнению с другими классами.

В целом, характеристики модели являются сравнительно хорошими, поскольку значения F1-Score достаточно высоки для всех классов. Тем не менее, все еще есть возможности для улучшения, особенно для «Класс 2», который имеет самое низкое значение метрики F1-Score.

В заключение мы разработали модель системы автоматизированной проверки текстов на предмет содержания информации ограниченного доступа, и провели первичную оценку ее эффективности. Первичная оценка эффективности модели была проведена с использованием таких стандартных метрик, как Accuracy, Precision, Recall, F1-Score. Эти метрики позволили получить представление о способности модели правильно определять категорию входного анализируемого текста, а также о соотношении между ложноположительными и ложноотрицательными предсказаниями [3].

Анализ показал, что модель продемонстрировала высокую эффективность при определении метки класса документов как содержащих информацию о технологиях двойного назначения, так и не содержащих. Метрики Accuracy, Precision, Recall, F1-Score были выше значения 0.9, что говорит о том, что модель способна эффективно различать два класса.

В дополнение к двоичной классификации модель также была оценена на предмет ее эффективности при определении категории документа. Самый высокий показатель F1-score составил 0.87, что является относительно высоким значением для этой задачи.

Разработанная модель может выступать в роли системы помощи принятия решения при проведении идентификационной экспертизы. В дальнейшей работе планируется добиться повышения эффективности работы модели за счет сбора более обширного набора данных и дообучения модели. Также планируется проведение апробации модели за счет применения ее на практике и оценка результатов работы в реальных условиях.

Литература

1. Борисова Е.М., Михайлова И.Л. Актуальные проблемы и направления развития экспортного контроля в россии // БИТ. 2021. №5. С. 58–63.
2. Типовая методическая инструкция по проведению экспертизы научно-технических материалов, подготовленных к открытому опубликованию и обладающих признаками контролируемых технологий: утв. Комиссией по экспортному контролю РФ от 03.04.14.
3. Yacouby R., Axman D. Probabilistic extension of precision, recall, and f1 score for more thorough evaluation of classification models //Proceedings of the first workshop on evaluation and comparison of NLP systems. 2020. С. 79–91.
4. Lagutina K. Topical text classification of Russian news: a comparison of BERT and standard models //2022 31st Conference of Open Innovations Association (FRUCT). IEEE. 2022. С. 160-166.
5. Qasim R. и др. A Fine-Tuned BERT-Based Transfer Learning Approach for Text Classification // Journal of Healthcare Engineering. 2022. Т. 2022. С. 1–17.

УДК 004.056.53

ПОСТ-КВАНТОВЫЕ ПРОТОКОЛЫ ВЫРАБОТКИ ОБЩЕГО СЕКРЕТНОГО КЛЮЧА

Ниткин И.С.[1] (магистрант), **Кирьянова А.П.**[1] (магистрант)
Научный руководитель – Давыдов В.В.[1]
1 – Университет ИТМО
e-mail: exebopen@gmail.com, anastaciakosanovskaya@gmail.com

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

Аннотация
В работе рассматриваются пост-квантовые протоколы выработки общего секретного ключа. Исследованы базовые принципы построения таких протоколов на основе изогений суперсингулярных эллиптических кривых и алгебраических решеток. Рассмотрены проблемы использования других пост-квантовых криптографических примитивов для построения протоколов выработки общего секретного ключа. Сделан вывод о наиболее перспективных направлениях исследований по разработке пост-квантовых протоколов выработки общего секретного ключа.

Ключевые слова
Криптография на основе изогений супесингулярных эллиптических кривых, криптография на основе алгебраических решеток, обмен информацией по открытому, защита информации, ключ симметричного шифрования.

Одной из ключевых задач, решаемых методами современной криптографии, является организация защищенного обмена информацией по открытым каналам в компьютерных сетях. Наиболее рациональным решением является использование методов симметричной криптографии. Однако использование симметричных шифров требует, чтобы участники обмена информацией обладали общим значением ключа шифрования. Для решения указанной проблемы применяются методы асимметричной криптографии.

В рамках асимметричной криптографии существуют два основных подхода к решению описанной задачи:

– применение алгоритмов инкапсуляции общего ключа (Public Key Encapsulation, PKE);

– применение алгоритмов выработки общего секретного ключа (Key Exchange, KEX).

Инкапсуляция общего ключа предполагает использование асимметричного шифрования. В рамках протоколов, реализующих данный метод, одна из сторон обмена информации вырабатывает открытый и закрытый ключи асимметричного шифрования и передает второй стороне значение открытого ключа. Вторая сторона вырабатывает значение секретного ключа симметричного шифрования, шифрует полученное значение при помощи открытого ключа, полученного от другой стороны, и пересылает обратно зашифрованное значение. Используя секретный ключ асимметричного шифрования, первый участник обмена информации может расшифровать значение общего ключа симметричного шифрования, которое в дальнейшем используется при осуществлении обмена информацией.

Протоколы выработки общего секретного ключа предполагают, что каждая из сторон выполняет некоторое, отличное от другой стороны, преобразование над общим известным значением. После этого стороны обмениваются результатами преобразований, и каждая из сторон повторно выполняет свое преобразование над переданным ей результатом. Если выбранные преобразования коммутируют, стороны в

конце получат одинаковое значение, которое используется в качестве общего секретного ключа.

В рамках настоящего исследования рассматриваются протоколы второго типа.

Важно отметить особенности протоколов выработки общего секретного ключа. Для обмена информацией используется открытый канал. Кроме того, в отличие от протоколов инкапсулящии общего ключа, вычисления могут выполняться участниками протокола параллельно, что потенциально позволяется достигать более высокой скорости выработки общего секретного ключа.

Примерами наиболее используемых в современных системах защиты информации протоколов KEX являются протокол Диффи-Хеллмана и протокол Диффи-Хеллмана на эллиптических кривых. Стойкость этих протоколов основана на сложности решения задачи дискретного логарифмирования (DLP), которая относится к классу вычислительных задач BQP и может быть эффективно решена при помощи алгоритма Шора с использованием квантового компьютера. Чтобы закрыть данную уязвимость, необходимо использовать пост-квантовые протоколы выработки общего секретного ключа.

Актуальность исследования усиливается тем фактом, что передаваемая информация может быть похищена злоумышленником в момент передачи, а расшифрована позже, когда станет доступен квантовый компьютер необходимого размера.

Алгоритмы пост-квантовой криптографии строятся на основе специальных пост-квантовых примитивов, к которым относятся изогении суперсингулярных эллиптических кривых, алгебраические решетки, хэш-функции, корректирующие коды и системы многомерных уравнений. В рамках настоящего исследования рассматриваются существующие протоколы выработки общего секретного ключа на основе пост-квантовых криптографических примитивов.

Базовым протоколом выработки общего секретного ключа на основе изогений суперсингулярных эллиптических кривых является протокол SIDH [1], схема которого представлена на рисунке 1.

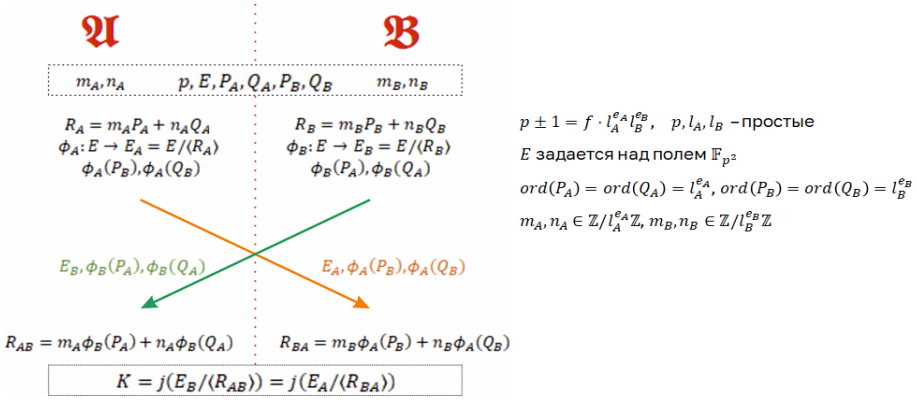


Рис. 1. Схема работы протокола SIDH

Порядок известной эллиптической кривой, которая является публичным параметром системы, выбирается таким образом, чтобы группа точек имела две подгруппы кручения. Стороны выбирают произвольные коэффициенты, при помощи которых вычисляется точка соответствующей подгруппы кручения, на основе которой строится изогения исходной кривой. Повторно выполнив вычисления над переданными другой стороной значениями, участники выработки общего ключа получают значения изоморфных эллиптических кривых, которые имеют общий j-инвариант, значение которого выбирается в качестве общего ключа.

Стойкость протокола SIDH основана на сложности задачи поиска изогении между суперсингулярными эллиптическими кривыми. Данный протокол обладает стандартным для протоколов на основе изогений набором достоинств и недостатков. Сравнительно небольшой объем передаваемых в процессе реализации протокола данных компенсируется значительным объемом вычислений, что снижает скорость работы.

Главной проблемой является тот факт, что в 2022 году протокол SIDH был взломан Кастриком и Декрю [2].

Несмотря на это существует ряд стойких протоколов выработки общего секретного ключа на основе изогений суперсингулярных эллиптических кривых.

В качестве основного может рассматриваться протокол CSIDH [3], схема которого представлена на рисунке 2. Стойкость данного протокола основана на сложности решения обратной задачи группового действия в графе изогений эллиптической кривой. Ядро изогении задается при помощи класса идеалов кольца эндоморфизмов эллиптической кривой. Класс идеалов для построения изогении каждая из сторон вычисляет на основе набора случайных экспонент. Характеристика кривой выбирается таким образом, чтобы кольцо эндоморфизмов было коммутативным. С учетом того, что существует взаимно-однозначное соответствие между изогениями суперсингулярной эллиптической кривой и левыми идеалами кольца эндоморфизмов такой кривой, в результате повторного группового действия над изогенией исходной кривой полученного от другого участника процесса выработки общего ключа, обе стороны получат одинаковую итоговую эллиптическую кривую. Общим ключом выбирается коэффициент при x^2 итоговой кривой, представленной в форме Монтгомери.

Атака Кастрика-Декрю не может быть реализована в отношении протокола CSIDH, так как она специфически ориентирована на набор открытых параметров протокола SIDH.

На основе изогений суперсингулярных эллиптических кривых заслуживают внимания модифицированные варианты протокола SIDH, которые маскируют открытые параметры схемы, а также протокол OSIDH, который построен схожим с протоколом CSIDH образом, однако предусматривает более сложные перемещения по графу изогений.

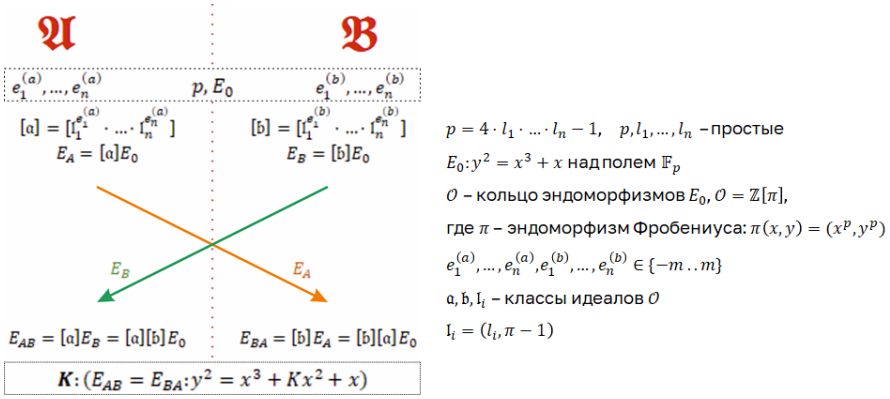


Рис. 2. Схема работы протокола CSIDH

В основе протоколов выработки общего секретного ключа на основе алгебраических решеток лежат NP-полные задачи теории решеток – задача поиска кратчайшего вектора (Shortest Vector Problem, SVP) и задача поиска ближайшего вектора (Closest Vector Problem, CVP).

Существуют две основные схемы выработки общего секретного ключа на основе решеток – схема Пикерта [4] и схема Динга [5]. Обе схемы построены на основе задачи обучения с ошибками (Learning with errors, LWE). Основная разница заключается в том, что схема Пикерта построена на основе обучения с ошибками в кольце многочленов

(Ring-LWE), а схема Динга – на основе обучения с ошибками на множестве векторов, задаваемых при помощи коэффициентов из поля вычетов по модулю простого числа.

Алгоритм выработки общего секретного ключа на основе схемы Пикерта представлен на рисунке 3. На рисунке 4 представлен соответствующий алгоритм на основе схемы Динга.

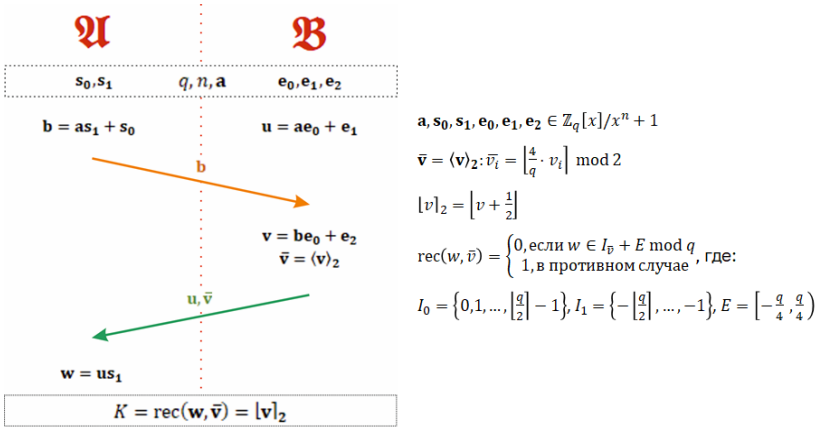


Рис. 3. Выработка общего секретного ключа по схеме Пикерта

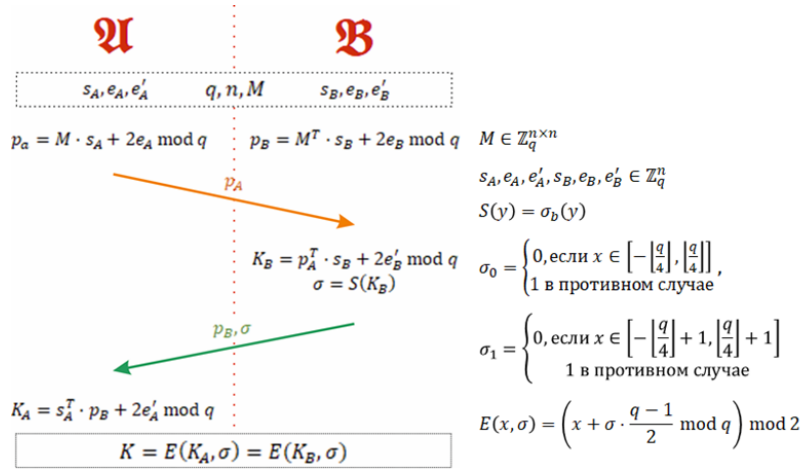


Рис. 4. Выработка общего секретного ключа по схеме Динга

Принцип работы алгоритмов описан ниже.

Первый участник выработки общего ключа выполняет преобразование публичного параметра (многочлена или вектора значений) и передает результат второму участнику. Второй участник выполняет схожее преобразование над полученным результатом. Принципиально преобразование состоит из умножения на один малый вектор, который условно называется секретным, и сложения полученного результата с другим малым вектором, который условно называется вектором ошибок. В случае работы в кольце многочленом данные вектора являются набором коэффициентов многочлена.

Общим ключом для схем на основе алгебраических решеток является битовый поток, формируемый на основе некоторого значения, которое может быть получено путем исправления векторов ошибок при помощи функций маскирования и сглаживания.

Протоколы выработки общего секретного ключа на основе алгебраических решеток обладают набором достоинств и недостатков, свойственным всем криптографическим схемам на решетках. Дополнительно в качестве недостатка важно отметить неполный параллелизм вычислений. При этом итоговое значение ключа

зависит от параметров, генерируемых обоими участниками, следовательно эти схемы должны рассматриваться как схемы выработки общего секретного ключа.

Среди наиболее актуальных решений следует отметить протокол New Hope на основе Ring-LWE, и протокол «Swoosh» на основе Module-LWE.

Для хэш-функций коммутативным является многократное последовательное вычисление значения хэш-функции. На основе данного свойства могла бы быть реализована схема выработки общего секретного ключа на подобии классического протокола Диффи-Хеллмана. При этом, в отличие от возведения в степень по модулю простого числа, для хэш-функций не может быть реализован аналог алгоритма быстрого возведения в степень. Следовательно, схема выработки общего секретного ключа на основе хэш-функций не имеет практического смысла, потому что вычислительная сложность выработки общего ключа эквивалентна сложности атаки перебором на основе значений, передаваемых по открытому каналу.

Использование различных хэш-функций участниками протокола также невозможно, потому что в общем случае такие преобразования некоммутативны.

Для корректирующих кодов и многомерных уравнений реализация классической схемы выработки общего секретного ключа невозможна, потому что длина блока информации на входе преобразования не соответствует длине блока на выходе. Следовательно, не могут быть реализованы последовательные коммутативные преобразования.

Таким образом, для построения пост-квантового протокола выработки общего секретного ключа необходимо использование коммутативного последовательного необратимого преобразования, стойкого к атакам при помощи квантового компьютера.

Среди существующих решений наибольший интерес представляют протоколы на основе алгебраических решеток и на основе изогений эллиптических кривых, которые имеют принципиально различную структуру. Структура других пост-квантовых примитивов не позволяет реализовать классические схемы выработки общего секретного ключа на их основе.

В качестве наиболее перспективного криптографического примитива в рамках исследований по разработке протоколов выработки общего секретного ключа необходимо отметить изогении суперсингулярных эллиптических кривых. Их главный недостаток –сравнительно низкая вычислительная скорость – частично компенсируется полным параллелизмом вычислений.

Литература

1. Feo L. De. Towards quantum-resistant cryptosystems from supersingular elliptic curve isogenies. / L. De Feo, D. Jao, J. Plût // Journal of Mathematical Cryptology: электронный журнал. URL: https://doi.org/10.1515/jmc-2012-0015. (дата обращения: 15.03.2023). Дата публикации: 2014.
2. Castryck W. An efficient key recovery attack on SIDH / W. Castryck, T. Decru // Cryptology ePrint Archive: электронный журнал. URL: https://eprint.iacr.org/2022/975. (дата обращения: 15.03.2023). Дата публикации: 2022.
3. CSIDH: An Efficient Post-Quantum Commutative Group Action / Castryck W., Lange T., Martindale C. [и др.] // Advances in Cryptology – ASIACRYPT 2018: электронный журнал. URL: https://doi.org/10.1007/978-3-030-03332-3_15. (дата обращения: 15.03.2023). Дата публикации: 2018.
4. Peikert C. Lattice Cryptography for the Internet / Peikert C. // Cryptology ePrint Archive: электронный журнал. URL: https://eprint.iacr.org/2014/070. Дата публикации: 2014.
5. Ding J. A Simple Provably Secure Key Exchange Scheme Based on the Learning with Errors Problem / J. Ding // IACR Cryptol. EPrint Arch.: электронный журнал. URL: https://eprint.iacr.org/2012/688. Дата публикации: 2012.

УДК 004.056

СИСТЕМЫ ЗАЩИТЫ ВЕБ-ПРИЛОЖЕНИЙ

Осипов И.Н.[1] (магистрант)
Научный руководитель – к.т.н. Югансон А.Н.[1]
1 – Университет ИТМО
e-mail: a_yougunson@itmo.ru

Аннотация
Данная работа рассматривает современные системы защиты веб-приложений, такие как брандмауэр и система обнаружения вторжений. Рассмотрены различные варианты реализации, выделены их достоинства и недостатки.

Ключевые слова
Веб-приложение, уязвимости, WAF, система обнаружения вторжений, наборы данных.

В наши дни многие люди не могут представить жизнь без интернета, они используют его для совершения различных покупок, используют различные банковские приложения, обмениваются различной информацией между друг другом. Данную информацию необходимо защищать, ведь утрата данных может привести к последствиям как для одного человека, так и для целого государства. Для совершения своих действий пользователи используют веб-приложения, они подвержены многочисленным атакам как из вне, так и изнутри системы. Вредоносные атаки стали более изощренными. Злоумышленники изучают и используют уязвимости систем обнаружения атак, увеличилось количество угроз безопасности, связанных с угрозами нулевого дня направленных на пользователя.

Веб-приложения являются основным сетевым решением для предоставления стандартных веб-сервисов. Функционирование этих приложений основана на разработке серверной части, а также клиентской. Серверная часть включает в себя веб-сервер, веб-приложение и сервер базы данных; он использует такие языки программирования как Python, PHP, а клиентская часть работает в веб-браузере пользователя с использованием CSS / HTML, Javascript. Связь между клиентом и сервером происходит при помощи протоколов HTTP. На рисунке 1 представлена архитектура веб-приложений на стороне сервера и клиента.

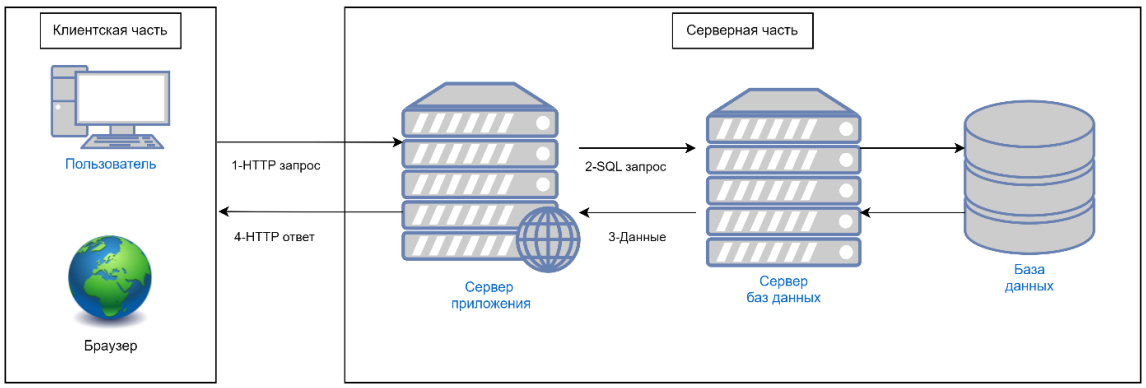


Рис. 1. Архитектура веб-приложений на стороне сервера и клиента

Веб-приложения стали неотъемлемой частью повседневной жизни людей из-за их доступности и удобства. Однако эта возросшая популярность – палка о двух концах. Действительно, веб-приложения являются основным способом для злоумышленников поставить под угрозу критически важные услуги в жизненно важных секторах, таких как

здравоохранение, образование, банковское дело [1].

Эти приложения обычно пишутся и разрабатываются программистами, которые могут быть не полностью осведомлены об аспектах кибербезопасности, что делает приложения уязвимыми для атак веб-приложений. Лучший способ снизить такие риски — написать безопасный код или добавить несколько уровней безопасности в архитектуру приложения, чтобы обеспечить всестороннюю безопасность, однако это сопряжено с большими затратами, а написать полностью защищенный исходный код практически невозможно. Вот почему исследователи и инженеры изобретают большое количество внешних систем для защиты веб-приложений и служб, например системы обнаружения вторжений, брандмауэры, а также различные системы защиты от мошенничества [2].

Согласно [1] веб-приложения имеют уязвимости, которые можно разделить на три основные категории. Первая это неправильная проверка ввода (Improper input validation) относится к неправильной проверке пользовательского ввода. SQL инъекции и межсайтовый скриптинг (XSS) являются примерами веб-атак, вызванных уязвимостью неправильной проверки ввода. Во-вторых, неправильное управление сеансом (Improper session management) относится к случаям, когда веб-сеанс не защищен должным образом, и, следовательно, приложение не может определить, являются ли веб-запросы вредоносными, пока они не будут связаны с надлежащим действительным идентификатором сеанса. Подделка межсайтовых запросов (CSRF) и перехват сеансов являются примерами веб-атак, вызванных неправильным управлением сеансом. Наконец, уязвимость с неправильной авторизацией и аутентификацией (Improper authorization and authentication) подразумевает логический недостаток в осуществлении политики контроля доступа, а также функций аутентификации. Нарушение контроля доступа является одной из веб-атак, которые могут произойти, если веб-приложение неправильно управляет процедурами аутентификации и авторизации, на рисунке 2 ниже представлены диаграммы, на которых можно увидеть распределение веб-атак по отраслям в России, а также сами атаки и уязвимости, которые использовали злоумышленники.

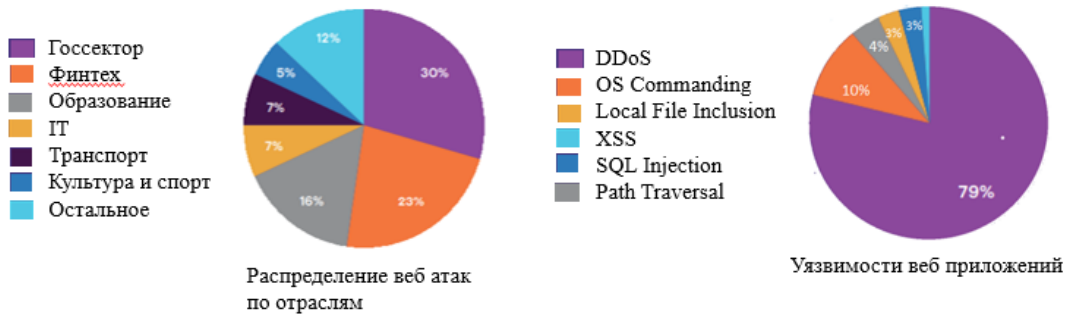


Рис. 2. Типы атак и распределение их по отраслям [3]

Существует много методов защиты веб приложений к самым элементарным можно отнести безопасное программирование ведь безграмотное написание программы может привести к утечке данных при использовании SQL инъекций или межсайтового скриптинга.

Механизмы безопасности обнаружения веб-вторжений были разработаны для обеспечения дополнительного уровня защиты веб-ресурсов. Этот дополнительный уровень веб-безопасности, широко известный как брандмауэры веб-приложений (Web-Application Firewalls ”WAF”) устанавливается на прикладном уровне, и веб-система обнаружения вторжений (Web-Intrusion Detection System ”WIDS”) устанавливается на сетевом уровне.

Одной из основных систем защиты веб-приложений является брандмауэр. Данная система обычно устанавливается на сервере, где работает веб-приложение. Он работает как фильтр между приложением и внешним миром, проверяя входящие запросы и исходящие ответы на наличие уязвимостей и злонамеренных действий.

Можно выделить два основных типа: брандмауэр на основе сигнатур и на основе машинного обучения.

Брандмауэр на основе сигнатур, наиболее распространенные в настоящее время, требуют, чтобы атаки происходили ранее, поскольку они проверяют трафик в поисках сигнатур, связанных с известными шаблонами атак. Поэтому они требуют постоянного обновления сигнатур. И они не способны идентифицировать атаки нулевого дня. Брандмауэры, основанные на обнаружениях аномалий, используют машинное обучение далее МО. Подход машинного обучения позволяет брандмауэру классифицировать запрос как вредоносный без применения методов сопоставления сигнатур. Цель состоит в том, чтобы уменьшить количество ложноположительных и ложноотрицательных вариантов, затрудняя злоумышленникам обман брандмауэр и снижая нагрузку, связанную с обновлениями сигнатур. Но модель машинного обучения должна быть обучена и переобучена. Машинное обучение дает хорошие результаты при применении к традиционным брандмауэрам, обнаруживая атаки нулевого дня и избегая необходимости постоянного обновления сигнатур и шаблонов трафика. Однако эти решения также создают новые проблемы и трудности, которые во многих случаях трудно решить или требуют больших затрат людских знаний и финансовых ресурсов. Среди этих трудностей можно выделить проблему, которая в значительной степени наблюдается при обучении некоторых математических моделей, используемых в машинном обучении. В этом смысле искажение обучающих данных является одной из самых больших проблем в области обнаружения брандмауэра. Трудно определить, были ли данные или информация сфальсифицированы, которые мы используем для обучения алгоритмов МО. Искаженные или ненадежные данные могут привести как к ложноположительным, так и к ложноотрицательным результатам. Еще одной существенной трудностью при применении МО к брандмауэрам является необходимость время от времени переучивать модели и вычислительные затраты, требуемые многими методами, используемыми для этой задачи. Кроме того, необходимо учитывать, что такая переподготовка требует затрат времени и человеческих ресурсов, а также необходимости иметь наборы данных, соответствующие необходимым ожиданиям [4].

Как было сказано раньше, существует множество различных методов и стратегий для построения моделей обнаружения веб-вторжений. Вторым из наиболее известных является система обнаружения вторжений, обычно данную систему размещают в центральном узле сети, таком как коммутатор, чтобы мониторить весь трафик, проходящий через сеть. Также системы обнаружения можно на два типа:

1. На основе сигнатур.
2. На основе обнаружения аномалий.

Системы, основанные на сигнатурах, используют методы обнаружения неправомерного доступа, сохраняя базу данных об известных атаках, проверяя входящие запросы на соответствие набору сигнатур (шаблонов), если какой-либо запрос подозревается в соответствии с одной из атак, система инициирует предупреждение или даже предотвращает запрос. Напротив, системы, основанные на обнаружении аномалий, вместо этого понимают шаблон нормального потока HTTP, извлекают признаки нормального трафика, а запрос, который отличается от стандартного, блокируется [5].

Однако в обеих моделях существует множество проблем. С одной стороны, современные системы, основанные на сигнатурах, поддерживают высокие показатели обнаружения, но это связано с затратами на поддержание обновленной базы данных о различных типах атак, в противном случае эти системы уязвимы для новых видов атак [6]. С другой стороны, детекторы аномалий превосходят системы, основанные на

сигнатурах, в обнаружении новых атак, но понимание шаблона легитимного потока HTTP-запросов труднодостижимо, поскольку данные HTTP очень изменчивы, зависят от приложения и изменяются со временем, то есть данный способ выдаёт много ложноположительных показаний.

Для реализации данных моделей часто используется искусственный интеллект, так в работе [7] обсуждается, как системы обнаружения веб-вторжений, основанные на сигнатурах, не справляются с раскрытием новых веб-уязвимостей, и предлагается система, которая использует несколько методов обнаружения аномалий. В предлагаемых подходах используются два параметра шаблонов доступа и параметров запроса. Обучение модели выполняется на обычных журналах HTTP, наивысшая оценка аномалии рассчитывается на основе нескольких моделей для обнаружения аномальных HTTP-запросов во время выполнения. Путем подсчета нескольких баллов аномалий, если какой-либо общий балл запроса превышает порог обнаружения (который рассчитывается на этапе обучения на основе наивысшего показателя аномалии для каждой программы и ее атрибутов), запрос помечается как потенциальная атака. Модель была проверена на 3 наборах данных данная модель показала очень низкий уровень ложных срабатываний (0,000419, 0,000650 и 0,000192). Однако предлагаемая модель работает только с HTTP-запросами GET и не обучается на тексте полезной нагрузки HTTP-сообщения.

Использование моделей глубокого обучения обсуждалось в работе [8]. Специализированная сеть CNN для обнаружения вредоносных веб-запросов путем анализа пакетов HTTP-запросов. Вклад заключается в том, что сама CNN извлекает признаки после нескольких шагов предварительной обработки. Модель была обучена на наборе данных (CSIC 2010). Процесс начинается с разделения строки URL на слова, применения встраивания слов одним из слоев специально разработанной сети CNN, а не третьей стороной. Авторы утверждают, что, когда встраивание слов изучается сетью и оптимизируется путем обратного распространения, результаты будут лучше. Сеть состоит из трех отдельных уровней:

1. Уровень максимального объединения: уменьшение размерности.
2. Уровень полного подключения: высокоуровневые рассуждения и извлечение признаков высокого порядка.
3. Слой Soft-Max: прогнозирование (классификационное отнесение запросов к нормальным, аномальным).

Модель показала относительно высокий уровень обнаружения (93,35 %), низкий уровень ложных срабатываний (1,37 %) и точность теста (96,49 %).

К проблемам данной системы обнаружения вторжений можно отнести отсутствие обучающих наборов данных, которые одновременно сбалансированы и охватывают широкий спектр веб-атак, существует проблема высокой стоимости создания помеченных наборов данных, а также сложность повторного использования набора данных в связи с тем, что у каждого веб-ресурса свои HTTP запросы.

Наборы данных для систем обнаружения веб-вторжений

UNSW-NB15: набор данных сочетает в себе фактические современные обычные действия и синтетические современные модели поведения при атаках. Он имеет девять типов атак, которые в основном связаны с сетевым вторжением. Обучающий набор данных включает 175 341 экземпляр, тогда как 82 332 экземпляра находятся в тестовом наборе. Он также в основном используется для обнаружения сетевых вторжений [1].

CICIDS-2017: Канадский институт кибербезопасности создал этот набор данных. Он содержит 2 830 540 различных экземпляров и 83 функции, содержащие 15 меток классов (1 обычная + 14 меток атак). Набор данных содержит только 2180 экземпляров веб-атак, что означает, что его недостаточно для оценки модели обнаружения веб-атак [6].

CSIC-2010: Набор данных содержит сгенерированный трафик, предназначенный для веб-приложения электронной коммерции. Это автоматически сгенерированный набор данных, который содержит 36 000 обычных запросов и более 25 000 аномальных запросов (т. е. Веб-атак) [8].

ECML / PKDD 2007: Набор данных является частью конференций ECML и PKDD по машинному обучению. Набор данных содержит 35 006 обычных трафика и 15 110 вредоносных веб-запросов. Набор данных был разработан путем сбора реального трафика, а затем обработан для маскировки имен и значений параметров – они были заменены на случайные значения [1].

Набор данных CSIC2012. После CSIC2010 аналогичный набор данных, который представляет собой основу для построения помеченных наборов данных для веб-трафика с целью оценки и тестирования брандмауэров веб-приложений (WAF). Набор данных включает 8363 обычных запроса, 16456 запросов, классифицированных как аномальные, и 40 984 записи, подпадающие под 8 типов атак. Классы атак включают: SQLi, Format String, XPATH, XSS, CRLFi, SSI, LDAPi, переполнение буфера [1].

Вывод

В данной статье рассмотрены основные методы и системы защиты веб-приложений, были выделены их достоинства и недостатки. Так, системы, использующие сигнатуры, хорошо справляются с уже известными атаками, но уязвимы для атак новых атак. Системы на основе сигнатур нуждаются в постоянном обновлении шаблонов, но в случае корректной отладки данная система способна эффективно предотвращать атаки. В настоящий момент идет упор на внедрение машинного обучения в системы обнаружения вторжений, но данные системы обнаружений веб-вторжений дороги в обслуживании и имеют проблему с наборами данных, ведь искаженные или ненадежные данные могут привести как к ложноположительным, так и к ложноотрицательным результатам. Еще одной существенной трудностью при применении МО к данным системам является необходимость время от времени переучивать модели и вычислительные затраты, требуемые многими методами, используемыми для этой задачи. Кроме того, необходимо учитывать, что такая переподготовка требует затрат времени и человеческих ресурсов, а также необходимости иметь наборы данных, соответствующие необходимым ожиданиям.

Литература

1. Alaoui R.L., Nfaoui E.H. Deep learning for vulnerability and attack detection on web applications: A systematic literature review //Future Internet. 2022. Т. 14. №. 4. С. 118.
2. Goutam A., Tiwari V. Vulnerability assessment and penetration testing to enhance the security of web application //2019 4th International Conference on Information Systems and Computer Networks (ISCON). IEEE. 2019. С. 601-605.
3. [Электронный ресурс]. Режим доступа: https://rt-solar.ru/analytics/reports/3289/ (дата обращения: 08.02.2023).
4. Calvo M., Beltrán M. An Adaptive Web Application Firewall. 2022.
5. Halim H.I., Kholief M., Maghraby F. Deep Learning Methods in Web Intrusion Detection: A Systematic Review. 2022.
6. Khraisat A. et al. Survey of intrusion detection systems: techniques, datasets and challenges //Cybersecurity. 2019. Т. 2. №. 1. С. 1-22.
7. Kruegel C., Vigna G., Robertson W. A multi-model approach to the detection of web-based attacks //Computer Networks. 2005. Т. 48. №. 5. С. 717-738.
8. Zhang M. et al. A deep learning method to detect web attacks using a specially designed CNN //International Conference on Neural Information Processing. Springer. Cham, 2017. С. 828-836.

УДК 004.056.55

БЕЗОПАСНАЯ СИСТЕМА ОБМЕНА МГНОВЕННЫМИ СООБЩЕНИЯМИ С ПОВЫШЕННОЙ КОНФИДЕНЦИАЛЬНОСТЬЮ ДАННЫХ

Прыгунов М.И.[1] (студент)
Научный руководитель – к.т.н., доцент Камальдинова З.Ф.[1]
1 – Самарский государственный технический университет
e-mail: maksim.prygunov@gmail.com

Аннотация
Цель данного проекта заключается в создании мессенджера с максимально возможным уровнем конфиденциальности. Это достигается отсутствием персональных данных пользователей на серверах, а также алгоритмом обмена сообщениями, при котором ключ (дешифратор) не проходит по сети.
Высокий уровень криптостойкости достигается использованием гибридной криптографической схемы, то есть использованием как ассиметричного, так и симметричного шифрования. Такое решение позволяет компенсировать недостатки симметричного шифрования при работе в сети, но сохранит главное преимущество проекта конфиденциальность личных сообщений, даже если у злоумышленника есть доступ к серверу и базе данных. Защита приложения ПИН-кодом отразит атаки на хранимые на устройстве ключи шифрования. Таким образом, технически переписка будет максимально защищена.
Ключевые слова
Обмен сообщениями, симметричное шифрование, хранение данных, безопасность данных, конфиденциальность, сеть.

Сейчас в мире представлено много решений так называемых мессенджеров, удобных для всех. Причем одна из самых значимых вещей, на которую обращает внимание пользователь при выборе подобного приложения, конфиденциальность. Но так ли хороши системы, которые нам предлагают?

Краткий ответ нет, если изучать вопрос подробнее, то можно узнать, что в большинстве случаев в базе данных информация хранится в открытом для администратора виде. Это дает возможность третьему лицу – администратору, читать личные сообщения пользователей. Получается, что конфиденциальность переписки держится на доверии этому самому администратору. Если человек дорожит своей перепиской, его не устроит подобное обстоятельство. Хранимые данные могут быть использованы администраторами для показа персонализированной рекламы, однако, есть риск того, что эти данные могут быть проданы третьим лицам без нашего согласия. Это может привести к нежелательной почте, спаму и другим формам нежелательной коммуникации. Кроме того, утечка личных данных может привести к серьезным последствиям, таким как кража личности и финансовых мошенничеств.

Еще одна проблема безопасность подобных данных. Администратор может совершать действия для достижения максимальной безопасности данных, но от утечек никто не застрахован. Как пример: год назад была слита база данных пользователей Яндекс Еды [1]. Стоит упомянуть, базы данных в большинстве своем не зашифровываются, это не выгодно для бизнеса, но, тем временем, та база содержала в себе фамилии, реальные адреса и телефоны, суммы заказов за последние полгода, причем речь не только о постоянных адресах жительства. Фактически по заказам можно отследить все перемещения пользователя по стране, его благосостояние. Для мошенников это очередная возможность. Где бы вы не находились – все будет на карте. Тогда в базе было семь миллионов человек, можно было найти себя и своих близких, в том числе знаменитостей и сотрудников государственных ведомств.

Некоторые мессенджеры громко заявляют о себе сквозным шифрованием. Что же, сквозное шифрование не спасает вас от всех проблем, остается уязвимость: человек по середине, когда кто-то может перехватывать все данные сети в одном узле и подделывать их для сервера. Тем более никто не гарантирует, что сервер сам не будет подделывать ключи и получать доступ к переписке, код у серверов закрытый. Если человек все же доверяет серверу, то существует необходимость проверки совпадения контрольных значений.

Еще одна проблема – анонимность [2], которая уже была частично затронута ранее. Человек, отдавая сервису свой номер телефона, может быть уверен, что при желании по этому номеру можно узнать, где он проживает. Более того, если у злоумышленника есть доступ к данным оператора, с помощью триангуляции можно узнать точное местонахождение человека в текущий момент.

Упомянутые проблемы натолкнули на идею создать максимально конфиденциальный мессенджер, где шифрование и дешифрование будет происходить только на устройствах, а на сервере будет хранится исключительно зашифрованная информация.

У устройств есть общий ключ, с помощью которого происходит расшифровка и шифрование сообщений. Этот ключ никуда не передается и всегда остается на устройствах. Более того, так как весь мир сейчас уже отказался от соединений без шифрования, зашифрованное приложением сообщение повторно зашифруется асимметричным шифрованием при передаче через HTTPS. Так что, на уровне сети сообщение будет зашифровано дважды, так достигается гибридная криптографическая схема.

Ключ один, но как начать переписку? Ключ необходимо создать на двух устройствах, для этого генерируется QR-код на одном из них, второй сканирует его и получает ключ к переписке, которая создается на сервере. Если соединяемые устройства обладают технологией NFC, их достаточно расположить рядом друг с другом для синхронизации ключа. Необходима личная встреча, мы идем на такую жертву ради уверенности в надежности переписки. К одной такой переписке может быть подключено несколько пользователей, а у каждого пользователя есть возможность пользоваться не только одним устройством. Чтобы подключить дополнительного пользователя – один из участников переписки должен передать ему ключ. На рисунке представлена наглядная схема работы сервиса в таком случае. Синхронизировать ключи между устройствами одного пользователя предлагается через его локальную сеть (рисунок).

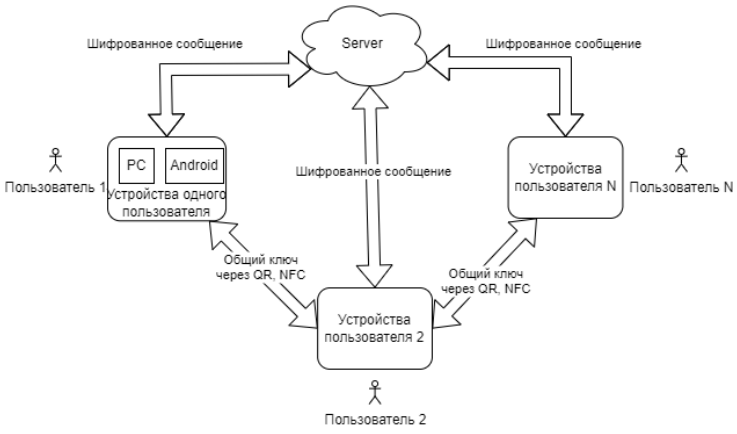


Рисунок. Схематичное изображение принципа работы [1]

В пользу анонимной регистрации приходится отказаться от способов восстановления пароля. Ровно также, как и в хороших блокчейн системах, у пользователя есть логин и пароль, никакой персональной информации. Ответственность

за анонимность логина перекладывается полностью на пользователя. Конечно, он может придумать логин вида «ivanov.ivan.2001», тогда уже об анонимности не может быть и речи, а пользоваться подобным приложением ему ни к чему. Хорошим решением также будет создание специальной системы ограничения регистрации подобных логинов.

Если описывать приложение архитектурно, то можно сказать следующее - алгоритм шифрования может быть выбран любой, в разработанном прототипе «Columba», размещенном в GitHub, используется стандарт AES с длиной ключа 128 бит, а чтобы одинаковое сообщение после шифрования принимало разный вид используются разные векторы инициализации [3]. Ключи шифрования, которые хранятся локально на устройствах, будут шифроваться повторно с помощью введенного пользователем пароля доступа, ровно как в хорошем банковском приложении.

Само сообщение содержит:

- вектор инициализации;
- зашифрованный текст;
- псевдоним отправителя;
- время отправления.

Код сервера-посредника оставить открытым не составит труда – он будет заниматься только сохранением сообщения в базе данных и его отправкой другим пользователям. Более того, можно провести эксперимент и сделать базу данных открытой - узнать, как долго будет расшифровываться одна случайная переписка. Расшифровав ее, злоумышленнику будет затруднительно понять кому она принадлежит - логин не содержит персональной информации. Более того, как хакеру выбрать ту самую переписку, которая ему интересна для расшифровки? Также стоит упомянуть о возможности не сохранять переписку на сервере, если того захочет пользователь. В таком случае теряется удобство для пользователя – пропадет синхронизация сообщений между устройствами одного пользователя, но вместе с автоматическим удалением сообщений после доставки и прочтения будет достигнут принципиально новый уровень безопасности и конфиденциальности данных.

Из преимуществ описанного подхода: безопасная переписка даже по незащищенному соединению, сама переписка конфиденциальна и анонимна. В недостатках то, что нужна очная встреча и появляются сложности синхронизации ключей на устройствах одного пользователя, также не стоит все еще исключать утечку со стороны пользователя.

Разработанный продукт позволяет вести безопасную переписку по незащищенному соединению. Безопасность и конфиденциальность переписки не будут зависеть от администратора системы, будут зависеть исключительно от длины ключа и самого пользователя. Чем больше бит занимает ключ, тем дольше будет проходить взлом шифра методом перебора. Со стороны пользователя есть угроза утечки пароля доступа к клиенту, а также не следует генерировать пароль при посторонних лицах, необходимо это учитывать при использовании.

Литература

1. Прыгунов М.И. Конфиденциальная система обмена мгновенными сообщениями. В сборнике: Перспективные информационные технологии (ПИТ 2022) [Электронный ресурс]: труды Международной научно-технической конференции / под ред. Прохорова С.А. Электрон. текстовые и граф. дан. (12,5 Мбайт). Самара: Издательство Самарского научного центра РАН. 2022. С. 120-122.
2. Гатчин Ю.А., Коробейников А.Г. Основы криптографических алгоритмов. Учебное пособие. СПб.: СПбГИТМО (ТУ). 2002.
3. Баричев С.Г., Гончаров В.В., Серов Р.Е. 2.4.2. Стандарт AES. Алгоритм Rijdael // Основы современной криптографии. 3-е изд. М.: Диалог-МИФИ. 2011. 176 с.

УДК 004.056

PORT SCANNING DETECTION USING MACHINE LEARNING MODELS

Stepanian A.A.[1] (student)
Scientific supervisor – tutor Khanov A.R.[1]
1 – Университет ИТМО
e-mail: mylenewit@gmail.com

The article was carried out within the graduation thesis topic "Development of a Software for Detecting Network Port Scanning Based on Machine Learning".

Abstract
This article describes the development of a software for detecting network port scanning using machine learning techniques. The software uses a dataset of network traffic to train a model that can accurately detect port scanning activity. The training process involves feature extraction and selection, as well as the selection of appropriate machine learning algorithms. Overall, the software represents a promising approach to detecting port scanning activity and improving network security.
Keywords
Port scanning, cybersecurity, network security, datasets, machine learning, supervised learning.

Nomenclature

AUC	Area Under the Curve.
CSV	Comma-separated values.
DT	Decision Tree.
FN	False Negative.
FP	False Positive.
FPR	False Positive Rate.
LR	Logistic Regression.
ML	Machine Learning.
MLP	Multi-layer Perceptron.
NB	Gaussian Naïve Bayes.
RF	Random Forest.
ROC	Receiver Operating Characteristic.
TP	True Positive.
TPR	True Positive Rate.

Introduction

Cybersecurity is a critical process that involves defending electronic systems, servers, mobile devices, and networks against malicious attacks. One of the key aspects of cybersecurity is network security, which focuses on protecting computer networks from intruders, including malware and targeted attacks.

As the size and complexity of networks continue to expand, IT experts face the challenge of building reliable and secure networks. Achieving maximum network security is a daunting task, as attackers are constantly seeking to exploit any vulnerabilities in the network to compromise its availability, integrity, or confidentiality.

One of the most common techniques used by attackers to gain information about a network is port scanning. This technique involves scanning a network to identify which computer ports are open. By identifying the ports that are open, attackers can determine which applications are currently running on the system [1], as certain programs listen on specific ports and respond to traffic in distinct ways.

Distinguishing between benign network traffic and malicious port scanning activity is an ongoing challenge in the field of cybersecurity. The increasing sophistication of attacks and the

growing number of devices connected to networks make it difficult to identify potentially harmful network activity. Port scanning is a particularly difficult activity to detect, as it can appear similar to legitimate traffic that is part of normal network operations.

To address this challenge, machine learning algorithms are offered in this paper. By training algorithms on large datasets of network traffic, it is possible to develop models that can accurately distinguish between normal network activity and port scanning. These algorithms can analyse patterns in network traffic and identify anomalies.

Using machine learning algorithms to detect port scanning activity can help network administrators to quickly identify potential threats and take appropriate action to prevent them from causing damage.

The difficulty of distinguishing between benign network traffic and malicious port scanning activity underscores the importance of using advanced tools and techniques to protect computer networks. By harnessing the power of machine learning algorithms, it is possible to improve network security and stay one step ahead of attackers.

Software development framework

Steps to develop and improve the accuracy of port scanning detection software:

1. Data Collection: a rich and diverse dataset of network traffic consisting of both port scanning attack traffic [2] and benign traffic [3] was collected. To ensure the accuracy and reliability of the data, an open-source data was carefully selected in CSV format, allowing for easy and efficient data processing and analysis.

2. Data Preparation: throughout the preprocessing stage, the data was carefully cleaned by removing any irrelevant or extraneous information that could hinder the accuracy of the final models. The data was then normalized to ensure that it conforms to a standardized range of values, allowing for accurate and consistent comparisons between different data points.

3. Feature Selection: a range of features was considered, including network protocol, packet size, and other relevant parameters. The features were then examined for correlations and associations with port scanning activity, and those found to have the strongest relationship were selected as the most relevant for use in machine learning algorithms.

4. Model Selection: the selection of the multi-layer perceptron (MLP), random forest (RF), decision tree (DT), gaussian naïve bayes (NB), and logistic regression (LR) models represents a significant contribution to the development of advanced machine learning algorithms for detecting port scanning activity. Through the use of these models, it is now possible to effectively detect and prevent cyber-attacks, enhancing network security and reducing the risk of costly data breaches.

5. Hyperparameter Tuning: to optimize the performance of the machine learning algorithms developed to detect port scanning activity, hyperparameter tuning was performed using grid search. This involved systematically searching through a range of hyperparameters for each of the selected models.

6. Model Evaluation: to evaluate the performance of the machine learning algorithms developed to detect port scanning activity, a range of metrics were used, including accuracy, precision, recall, ROC, AUC, and F1-score. These metrics were selected based on their ability to provide a comprehensive view of the performance of the models, taking into account both true positives and false positives.

Correlation analysis

In the field of machine learning, selecting the right features is crucial for building accurate and efficient models. Correlation analysis is an important technique for identifying the most relevant features in a dataset, as it allows researchers to measure the strength and direction of the relationships between different variables.

By using correlation analysis, it is possible to quickly identify features that are highly correlated with each other, and to remove any redundant or irrelevant features from the dataset.

This can help to improve the accuracy and efficiency of machine learning models, while reducing the risk of overfitting and improving the overall performance of the system.

To select the most relevant features for detecting port scanning activity using machine learning, correlation analysis was performed on the collected data using the Sklearn [4] heatmap function. A threshold of 0.85 was used to identify highly correlated features, which were then selected for use in the machine learning models. Any irrelevant features were dropped from the dataset, resulting in a new heatmap shown in Figure 1 that consisted only of the most important features.

As a result, the heatmap no longer consists of data points that are highly correlated with each other beyond the 0.85 threshold, except for the features themselves (fig.1).

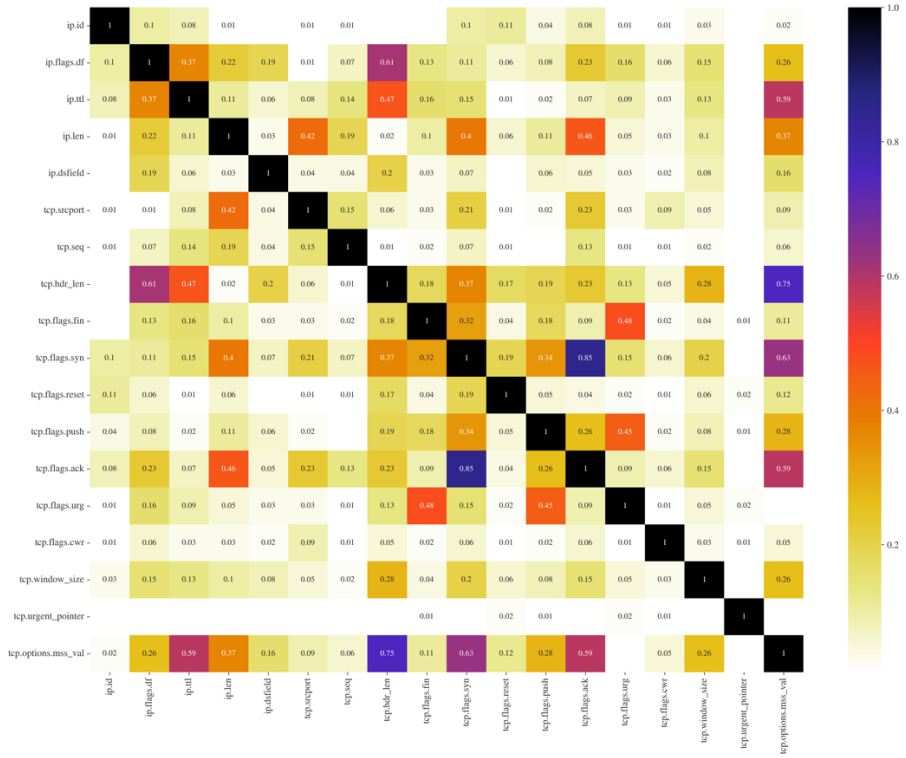


Fig. 1. Heatmap after dropping irrelevant features

Software development and evaluation

In this paper, it was found that several supervised machine learning models are particularly effective at distinguishing between malicious traffic and benign one. Specifically, MLP, DT, RF, LR, and NB algorithms were identified as being among the most successful at this task.

These models were selected based on their ability to handle complex, high-dimensional data and to learn from patterns in the data to accurately classify traffic as malicious or benign. Additionally, they were well-suited for the size and complexity of the dataset, which included large amounts of network traffic data.

Each of these algorithms was used to build and train a machine learning model that could accurately classify network traffic as either benign or malicious. The models were optimized using grid search as a hyperparameter tuning and cross-validation to achieve the best possible performance.

Once the models were trained and optimized, they were integrated into the software solution for detecting network port scanning.

For RF and DT, a grid search was conducted to identify the optimal combination resulting in the highest F1-score [5] performance. In the case of RF, the grid search was carried out with different values for the number of estimators, including 10, 15, and 20, using criteria of Gini

and entropy, maximum depth of 5 and 10. As for the DT, maximum depth was set up to 5, 10, and 15.

Model evaluation is a crucial step in machine learning, and there are various metrics to measure the performance of a model. One commonly used metric is the F1-score.

The F1-score is a measure of a model's accuracy [6] that takes into account both precision and recall. It is the harmonic mean of precision and recall, with a value that ranges from 0 to 1. A score of 1 represents perfect precision and recall, while a score of 0 represents poor performance.

To calculate the F1-score, first precision and recall must be assessed. Precision is the ratio of true positives (TP) to the total number of predicted positives, while recall is the ratio of true positives to the total number of actual positives.

$$\text{Precision} = \frac{\text{TP}}{\text{TP+FP}},$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP+FN}}.$$

Once precision and recall are calculated, F1-score can be assessed using the following formula:

$$\text{F1-Score} = 2 \times \frac{\text{Recall} \times \text{Precision}}{\text{Recall} + \text{Precision}}.$$

The validation step was performed using stratified K-fold with a value of k = 10. Table 1 provides a summary of the F1-score performance for all algorithms, including the average score from each K-fold.

Table 1

F1-score performance from each k-fold

K-FOLD	MLP	RF	DT	NB	LR
0	0.998633	0.997773	0.999931	0.776275	0.918413
1	0.999006	0.999655	0.999945	0.775338	0.917180
2	0.998592	0.999613	0.999972	0.776466	0.917518
3	0.998799	0.999779	0.999945	0.776678	0.917715
4	0.999006	0.999227	0.999890	0.775366	0.918185
5	0.999255	0.999061	0.999959	0.776366	0.917826
6	0.998688	0.998355	0.999903	0.774810	0.919038
7	0.999116	0.999641	0.999972	0.774606	0.917210
8	0.998868	0.999765	0.999931	0.777129	0.917376
9	0.998716	0.995592	0.999945	0.775962	0.916391

The strong indication of over-fitting from the dataset's machine learning models is reflected by high F1-scores, such as 0.999 for RF, DT, and MLP. The correlation between F1-score and ROC/AUC should also be emphasized, although the positive class probability estimation for each algorithm is used to calculate the ROC/AUC based on all samples in the dataset. This decision is aimed at providing further discussion on ROC analysis regarding the trade-off between threshold and TPR/FPR. Fig. 2 illustrates the ROC curve and the corresponding AUC for each algorithm, with an average AUC of 0.979.

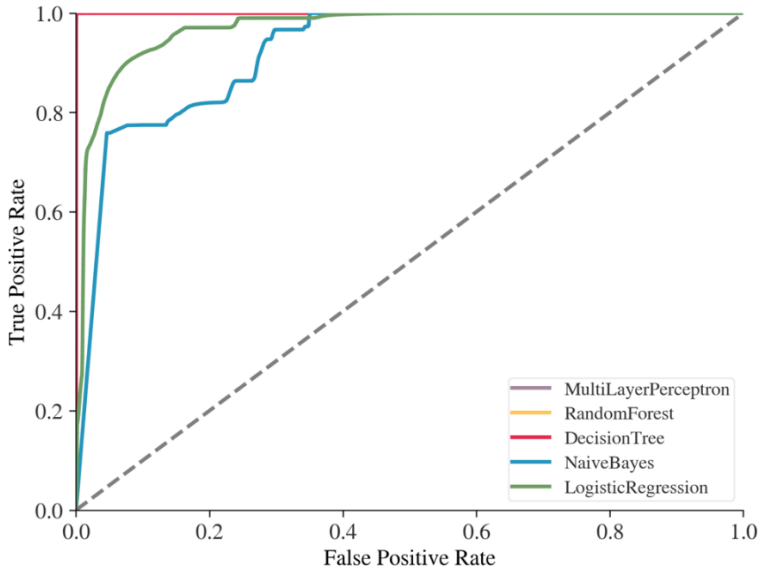


Fig. 2. ROC curve

The AUC value was calculated for each algorithm to evaluate its ability to distinguish between positive and negative classes, as shown in table 2. This metric provides a comprehensive performance measure that considers the entire range of TPR and FPR thresholds, thus making it a useful tool for comparing different classifiers.

Table 2

AUC value for each ML algorithm

	MLP	RF	DT	NB	LR
AUC	0.999929	0.999991	0.999944	0.924172	0.969724

Overall, a thorough performance evaluation of various machine learning algorithms has been conducted. The evaluation involved the use of different evaluation metrics such as accuracy, precision, recall, F1-score, ROC, and AUC. The results were summarized in tables and graphs, providing an overview of each algorithm's performance. This evaluation process allowed to select the most suitable algorithms for the development of a software that detects port scanning attacks.

Conclusion

The development of a software for detecting network port scanning based on machine learning has been successfully carried out. A dataset consisting of benign and malicious traffic was used to evaluate the performance of various machine learning algorithms, including MLP, DT, RF, LR, and NB. The algorithms were trained using grid search, and the F1-score was used as the performance metric. The best performing algorithms were RF, DT, and MLP with an F1-score of 0.999, although LR also showed a good performance with an F1-score of 0.919, indicating that a simple parametric model can be used to detect port scanning.

The ROC curve and AUC were also evaluated, with an average AUC of 0.979. Correlation analysis was used to select relevant features for the machine learning models. The heatmap generated showed that features with a correlation above 0.85 were dropped from the dataset. The resulting software provides an effective solution for detecting network port scanning, with a high level of accuracy and precision. Further research can be carried out to improve the performance speed of the software and to explore its potential applications in the area of network security.

References

1. Kumar M. Satheesh et al. "Artificial intelligence managed network defense system against port scanning outbreaks." 2019 International Conference on Vision Towards Emerging Trends in Communication and Networking (ViTECoN). IEEE. 2019.
2. Dataset of Probing Attacks performed with neap, unicornscan, hping3, zmap and masscan. [Электронный ресурс]. URL: zenodo. org/ record/ 3558350. (Publication date: 29.11.2019).
3. MAWILab dataset - traffic connecting the USA and Japan. [Электронный ресурс]. URL: fukuda-lab.org/mawilab/v1.1/2020.html. (Publication date: November 2020).
4. The Python Standard Library. [Электронный ресурс]. URL: docs.python.org/3/library. (Publication date: November 2020).
5. Bertoli G. De C., et al. "An end-to-end framework for machine learning-based network intrusion detection system." IEEE Access 9 (2021): 106790-106805.
6. Zhou J. et al. "Evaluating the quality of machine learning explanations: A survey on methods and metrics." Electronics 10.5 (2021). 593.

УДК 004.056

ФАКТОРЫ ВЛИЯЮЩИЕ НА СИСТЕМЫ БЕЗОПАСНОСТИ В МУЛЬТИАГЕНТНОЙ СИСТЕМЕ АВТОНОМНЫХ НЕОБИТАЕМЫХ ПОДВОДНЫХ АППАРАТОВ

Тимкин А.К.[1], Кича И.В.[1]
Научный руководитель – к.ф.-м.н., доцент ФБИТ Комаров И.И.[1]
1 – Университет ИТМО
e-mail: timkinalex@mail.ru

Аннотация
В настоящее время активно идет развитие автономных необитаемых подводных аппаратов (АНПА). Они применяются для решения широкого спектра задачи в различных областях. Развитие программной и аппаратной части позволяет управлять АНПА в группах и реализовывать мультиагентные системы. Однако возникает вопрос, связанный с обеспечением безопасности информационного взаимодействия в мультиагентной системе (МАС). В работе проведён анализ факторов, влияющих на системы информационной безопасности (ИБ) в мультиагентных системах автономных подводных аппаратов. Проведенный анализ позволит концентрировать внимание на факторах, оказывающих наибольшее влияние на безопасность информационного взаимодействия элементов МАС АНПА.

Ключевые слова
Мультиагентные системы, автономные необитаемые подводные аппараты, безопасность информационного взаимодействия, гидроакустический канал связи, робототехнические системы.

Введение

С каждым годом все больше растет интерес к подводной робототехнике, в том числе – автономным необитаемым подводным аппаратам, объединяемым в МАС.

С развитием вычислительных мощностей центральных вычислительных комплексов (ЦВК), а также развития аппаратной части АНПА открываются новые позволяют реализовывать новые решения в области автономной подводной робототехнике. Например, развитие АНПА тяжёлого класса идёт по пути увеличения их информационной и энергетической автономности [1].

Поскольку традиционные проблемы повышения функциональных характеристик АНПА являются трудоемкой задачей, но вместе с этим позволяет решать новые задачи, растет уровень интереса к вопросам связанных с обеспечением их ИБ. Особенно актуально вопрос стоит в развитии систем ИБ в МАС АНПА. Для дальнейшего качественного исследования факторов будут учитываться технические характеристики легкого и среднего класса [2], поскольку именно данные классы вызывают наибольший интерес с точи зрения реализации мультиагентных концепций.

Можно выделить несколько задач, которые требуют решения в данных системах:

1) обеспечение конфиденциальности, целостности и доступности информационных ресурсов отдельного аппарата и МАС в целом;
2) безопасность информации, передаваемой по каналам связи;
3) мониторинг статуса аппарата;
4) противодействие угрозам неправомерного использования АНПА.

Традиционно для решения задач обеспечения ИБ АНПА используются механизмы «жёсткой» безопасности: шифрование канала связи и применение специальных программно-аппаратных средств [3].

В настоящее время серийные гидроакустические системы связи являются наиболее часто используемой технологией передачи данных под водой на расстояниях в десятки

километров [4]. Из-за особенности морской среды существует множество факторов, которые непосредственно или косвенно влияют на качество систем ИБ.

На уровень ИБ МАС АНПА влияют следующие группы факторов:

1) внешняя обстановка;
2) скоротечность процессов коммуникации;
3) потенциальная возможность компрометации агента путём его захвата;
4) ограниченность бортовых вычислительных ресурсов;
5) наличием внешних и внутренних нарушителей ИБ.

Таким образом, исследование факторов, влияющих на систему ИБ в МАС АНПА является актуальной научно-технической задачей, решение которой позволит повысить безопасность информационного взаимодействия в группировке АНПА. Основной целью данной работы является определение факторов, влияющих на качество системы ИБ в МАС АНПА.

Факторы, влияющие на качество системы ИБ в МАС АНПА

Пусть $D = \{d_1, \dots, d_n\}$ – множество агентов АНПА (аутентичные и скомпрометированные), действующих в МАС. Для передачи данных и доступа к ресурсам системы агенты должны осуществлять проверку подлинности (аутентификацию).

Тогда задача исследования состоит в исследовании специфических факторов (N_R), влияющих на качество механизма аутентификации, в представленном в виде (1)[5]:

$$A = F(D, E, S, K, \dots, N_R) \quad (1)$$

где, $E = \{e_0, \dots, e_n\}$ – влияние окружающей обстановки, $S = \{s_1, \dots, s_l\}$ – состояние агентов, $K = \{k_1, \dots, k_m\}$ – информация необходимая для аутентификации.

Существует множество факторов, влияющих на систему аутентификации в МАС АНПА, но их все можно разделить на три главные группы:

- аппаратные характеристики;
- алгоритмы взаимодействия;
- внешние факторы (среды).

Нужно отметить, что данные факторы можно также разделить на несколько типов, управляемые и не управляемые. Например, разработчик может улучшать и технические характеристики АНПА и модифицировать аппаратно-программную часть, но не может повлиять на внешние факторы, такие как морская обстановка или внешняя атака на агента.

Аппаратные характеристики

Аппаратные характеристики – это управляемые факторы аппаратного комплекса АППА, которые влияют на качество аутентификации в МАС АНПА, в данные факторы входят:

1) энерговооружённость;
2) характеристики антенны;
3) производительность ЦВК.

Важной частью работы системы ИБ являются характеристики самих агентов, поскольку они определяют основные параметры взаимодействия в группе.

От энерговооружённости и, соответственно, допустимой потребляемой мощности зависит возможность работы агента в активном режиме. Характеристика антенны определяет дальность, глубину связи между агентами, а также накладывает ограничения на программную реализацию. ЦВК предопределяет возможности использования различных протоколов передачи информации.

Алгоритмы взаимодействия

Алгоритмы взаимодействия – являются управляемыми факторами программного обеспечения АППА. Их них можно выделить:

1) алгоритм связи;
2) протокол маршрутизации.

Алгоритмы связи являются ключевой частью работы системы информационного взаимодействия, в том числе - для решения задач обеспечения ИБ, поскольку определяют объём и скорость передаваемой информации, что очень важно в таких сложных условиях, как морская среда. Также алгоритмы коммуникации определяют потенциальное количество абонентов в мультиагентной сети, что может значительно влиять на качество системы, если канал связи не справляется с нагрузкой.

Внешние факторы

К числу наиболее существенных внешних (не контролируемых) факторов относятся;

1) подверженность агентов MAC атакам (внешним и внутренним);
2) акустическая обстановка.

Важным и высоко динамичным фактором является акустическая обстановка, поскольку она может значительно затруднять взаимодействиями между агентами.

По результатам исследования предлагается обобщённая классификация факторов, которые оказывают существенное влияние на качество информационного взаимодействия в мультиагентной группировке АНПА (рисунок).

С точки зрения обеспечения ИБ целевого процесса, для реализации которого создаётся группировка АНПА, эти факторы являются источниками всех видов угроз: нарушения конфиденциальности, целостности и доступности информации. В итоге это может привести к невыполнению постеленной задачи или даже разрушению группировки или её элементов.

Рисунок. Классификация факторов, оказывающих влияние на качество информационного взаимодействия в мультиагентной группировке АНПА

Использование предлагаемой классификации обеспечивает рациональное и целенаправленное выделение ресурсов для противодействия угрозам ИБ коммуникационного взаимодействия для достижения заданного уровня ИБ.

Заключение

Развитие робототехники является одной из приоритетных стратегических задач нашей страны. Особенно стоит уделить внимание морской и подводной робототехнической технике, обладающей большим потенциалом для решения широкого спектра задач. В свою очередь, обеспечение ИБ МАС АНПА является необходимым условием их эффективного применения.

В контексте актуальных задач обеспечения ИБ АНПА перспективными направлениями дальнейших исследований являются:

– получение количественных зависимостей характеристик безопасности этапов информационного взаимодействия от значений отдельных факторов и групп устойчивых сочетаний факторов;

– разработка методик формирования требований к отдельным подсистема информационного взаимодействия в группировке АНПА в зависимости от предполагаемых условий их функционирования и требований к уровню ИБ.

Литература

1. Борейко А.А. и др. Система управления АНПА большой автономности на базе мультиагентного подхода //Подводные исследования и робототехника. 2019. №. 2. С. 23-31.
2. ГОСТР 56960-2016 Аппараты необитаемые подводные.
3. Решотка А.В. Анализ способов распределения ключевой информации в робототехнических комплексах военного назначения //Состояние и перспективы развития современной науки по направлению «Информационная безопасность». 2021. С. 616-624.
4. Урик Р.Д. Основы гидроакустики. Л.: Судостроение. 1978. 446 с.
5. Чан З.Х. Разработка модели и методов аутентификации в группировке беспилотных воздушных судов. // Диссертация на соискание учёной степени кандидата технических наук, Национальный исследовательский университет ИТМО (Университет ИТМО). Санкт-Петербург 2022.

УДК 004.414.2

КОНТРОЛЬ КАЧЕСТВА ДАННЫХ С ПОМОЩЬЮ БЛОКЧЕЙН ТЕХНОЛОГИЙ

Токаев А.Д.[1] (студент)
Научный руководитель – к.ф.-м.н. Корхов В.В.[1]
1 – Санкт-Петербургский государственный университет
e-mail: tokaev_artem@mail.ru

Аннотация
В данной работе рассматривается актуальная проблема повышения качества данных в информационных системах. Автор подчеркивает, что качество данных является важным критерием для эффективной работы информационных систем, и необходимо обеспечить контроль качества данных в процессе их обработки. Рассматриваются существующие методы контроля качества данных и возможности применения технологий блокчейн для решения проблемы. Особое внимание уделяется реализации механизмов контроля качества данных на основе приватных блокчейнов. Автор приходит к выводу о расширении возможностей внедрения технологии блокчейна в информационные системы при использовании механизмов оценки качества данных.

Ключевые слова
Качество данных, IoT, блокчейн, смарт-контракт, Ethereum.

Выбранная тема в настоящее время особо актуальна, так как с ростом технологического прогресса появляется все больше различных данных. В настоящее время блокчейн-технология активно развивается и находит широкое применение во многих сферах, которые используют ее уникальные возможности. Однако блокчейн-приложения могут сохранять некорректные или ошибочные данные, которые нельзя будет изменить. Это приводит к необходимости разработки механизмов, которые контролируют качество данных. Большее количество данных поступает из области «Интернета вещей» (далее – IoT). Датчики, используемые в различных информационных системах, могут выходить из строя и передавать неверные или ошибочные данные, а также не передавать их вовсе. Одним из способов решения проблемы точности и достоверности данных в блокчейн-приложениях являются смарт-контракты, которые фиксируют неверные данные, помечая их или сигнализируя об ошибках. Использование смарт-контрактов позволяет улучшить качество данных и повысить их достоверность, что в свою очередь повышает эффективность работы систем и приложений.

Цель данной работы заключается в создании механизма контроля качества данных с использованием блокчейн-технологий. Для достижения этой цели решаются следующие задачи:

1) обозначение проблем качества данных и причин их возникновения;
2) изучение существующих методов контроля качества данных с использованием блокчейн-технологий;
3) создание смарт-контрактов для оценки качества данных.

Качество данных (далее – КД) – это характеристика, которая показывает степень полезности данных [1]. КД – это неотъемлемая характеристика данных, которая представлена через ее атрибуты, которые охватывают различные аспекты использования данных. Согласно стандартам ISO 9000: 2015, ключевыми критериями качества данных являются:

- релевантность;
- точность;
- целостность;
- полнота;

– последовательность;
– объем данных;
– своевременность.

В практических случаях эти параметры несут разные веса для разных организаций, поэтому конечная модель качества данных сильно отличается в зависимости от специфики. В настоящее время отсутствуют универсальные метрики оценки качества данных. Основными способами сбора данных на сегодняшний день являются:

– датчики (регистрация и измерение событий);
– транзакции (регистрация событий);
– публикации (факты и утверждения);
– интервью (мнения, воспоминания);
– открытые источники (веб, социальные сети);
– тексты (выявление упоминаний об объектах и фактах);
– исторические (хранимые данные);
– метаданные:
– авторство;
– время;
– место.

Возможные причины низкого качества данных:

– человеческий фактор – ошибки ввода;
– текстовые данные – сложности установления семантики текста, разный лексикон, разные методы анализа;
– контекст – время, условия;
– отсутствие данных – конфиденциальность;
– неоднородное представление – из разных источников, разные стандарты;
– различная точность – от разного оборудования, получение разными методами, различные требования.

В зависимости от исследуемой области проблемы качества данных могут меняться из-за смены способов сбора данных и типов данных. Согласно [2], данные любого характера и размера подвержены искажениям, многие из которых связаны с недостающими значениями, дубликаты, противоречия, аномальные значения и отсутствие инклюзивности. Эти проблемы можно рассматривать с разных точек зрения: нарушения структуры, формата, случайная или преднамеренная подстановка значений, потеря ассоциаций и т. д.

Смарт-контракты (далее – С-М) – компьютерный код, развернутый в блокчейне для выполнения конкретных задач после выполнения заранее определенных условий. С-М может реализовать и выполнить условие, выраженное в бизнес-процессе. Они могут передавать актив, выполнять другие С-М, взаимодействовать с другими внешними службами (вне цепочки).

Блокчейн (Block – блок, chain – цепь) – это децентрализованная база данных, которая предназначена для хранения последовательных блоков с набором характеристик. Блокчейн позволяет избавиться от посредников, то есть напрямую обращаться к базе данных. Так как блоки, которые уже были созданы, изменить нельзя, эта технология создает некий контроль качества данных. Подделать существующие данные в такой системе невозможно, поэтому качество записанных в них данных становится критически важным. Любое значение данных, записанное в них, де-факто становится достоверной истиной, которой доверяют все стороны в системе. Без какого-либо контроля качества данных, которые записываются в блокчейн, существует риск принять за достоверные данные неверные или ложные данные, что может привести к сбою процессов в информационных системах.

Транзакция – группа логически объединённых последовательных операций по работе с данными, обрабатываемая или отменяемая целиком.

Использование блокчейн-технологии рассмотрим на примере задачи по перевозке опасных грузов, которые в наше время решают различные мониторинг-системы, использующие датчики такие как GPS трекер, RFID-метки и сенсоры.

Для контроля перевозки жидкости с определенной температурой возгорания в специальных резервуарах используются различные методы и технологии. Одним из наиболее распространенных методов является мониторинг температуры и давления в резервуарах в режиме реального времени. Для этого используются датчики, которые могут быть установлены на самом резервуаре или подключены к нему через специальное оборудование. Влияние таких технологий как IoT и блокчейн предоставляют новый уровень прозрачности и проверки процесса транспортировки в режиме реального времени. Также стоит отметить, что в IoT проблема неоднородности данных частично решена использованием всевозможных датчиков. Поскольку КД влияет на результаты передачи данных, то безопасность, надежность и доступность информации в хранилище данных зависит от КД, используемых в системе [3].

Существующие решения, позволяющие обеспечивать контроль качества данных в рассматриваемой задаче, реализуются на основе публичного блокчейна [4] или при помощи искусственного интеллекта [5]. В первом случае есть ряд минусов, такие как: низкая скорость транзакции, недостаточная конфиденциальность (данные, хранящиеся в блокчейне, находятся в общем доступе), высокие затраты на транзакции (публичные блокчейны взимают комиссию). Во втором случае обучение модели нейросети на распознавание некачественных данных достаточно трудозатратно, а непосредственное использование данных моделей требует больших вычислительных мощностей в решении данной задачи. Поэтому было принято решение реализации механизмов контроля качества данных на основе приватных блокчейнов.

Можно определить два архитектурных варианта реализации контроля качества данных:

1. Оценка КД с использованием приложений, потребляющих данные или производящих данные.

Недостатки первого варианта:

– отсутствие механизма, препятствующего переполнению блокчейна низкокачественными данными, когда производители не контролируют качество своих данных;

– отсутствие гарантий КД для потребителя, который будет использовать данные в блокчейн-приложении. Получается, что следует слепо доверять производителю/потребителю данных;

– если есть посредник, то встает вопрос доверия к его данным.

2. Оценка КД в блокчейне. Этот вариант позволит преодолеть вышесказанные ограничения, позволяя узлам выполнять проверку КД до того, как транзакции с данными будут зарегистрированы в блокчейне.

Выделяют следующие методы контроля качества данных:

– анализ границ значений: проверка того, что данные находятся в ожидаемом диапазоне значений;

– использование резервных датчиков: обеспечивает возможность проверки точности измерений с помощью других независимых датчиков;

– методы фильтрации данных: используются для устранения шума и других аномалий, возникающих в данных;

– проверка целостности данных: проверка, что данные, полученные от датчиков, не были повреждены или изменены во время передачи по сети.

Без какого-либо контроля качества данных, которые записываются в блокчейн, существует риск принять за достоверные данные неверные или ложные данные, что

может привести к сбою процессов в информационных системах. Поэтому задача свелась к оценке данных, поступающих с датчиков температур, при помощи смарт-контракта. Таким образом, был разработан смарт-контракт, в котором реализована логика оценки качества данных и оценивались следующие критерии:

– точность относительно эталонного значения (определяется как мера близости значения данных v к эталонному значению v′);

– оценка среднеквадратичного отклонения (определяется с точки зрения стандартного отклонения измеренных ранее значений);

– полнота – измерение степени присутствия требуемых данных в наборе полученных данных;

– своевременность – измерение временной достоверности данных.

Каждая транзакция содержит показание датчика и его временную метку, параметры КД устанавливаются при создании смарт-контракта. Все элементы, нужные для оценки данных, находятся внутри транзакции. Все измерения КД немедленно оцениваются, когда транзакция добавляется в новый блок.

Смарт-контракт был написан на языке Solidity и реализован в приватной блокчейн-сети Ethereum, развернутой на локальной машине. Было смоделировано поведение датчика, который передает свои показания каждые 5 секунд в течение 1 минуты для смарт-контракта использующего и не использующего элементы КД. Оценивалось быстродействие и затраты на выполнение транзакций. Результаты, приведенные в таблице, показали, что влияние элементов контроля качества данных на время обработки незначительны. Следовательно, можно предположить, что использование элементов контроля КД в существующих системах мониторинга, использующих блокчейн, не будет вызывать задержку. Однако влияние элементов КД на стоимость исполнения транзакции (gas`a) достаточно велико. Но при использовании приватной блокчейн-сети стоимостью транзакции можно управлять и обеспечивать более эффективное использование ресурсов сети.

Таблица

Результаты эксперимента

Смарт-контракты	Время обработки транзакции	Затраты на выполнение транзакции	Прирост
Используя КД	Менее 1 секунды	246070 gas	115,6%
Не используя КД	Менее 1 секунды	114110 gas	–

Таким образом, внедрение данного смарт-контракта или его логики в существующие смарт-контракты мониторинг-системы позволит получить более подробные данные, содержащие в себе метрики, описанные выше. Данные метрики позволят успешнее анализировать данные.

Система мониторинга температур, использующая блокчейн со смарт-контрактами, имеет следующие преимущества:

– надежность: блокчейн гарантирует целостность данных и защиту от подмены информации;

– безопасность: смарт-контракты позволяют автоматически контролировать выполнение условий, например, сигнализировать о превышении определенных температурных порогов, что может предотвратить дополнительные риски и повысить безопасность;

– прозрачность: блокчейн обеспечивает прозрачность и доступность данных для всех участников системы, что может повысить доверие к продукту и улучшить сотрудничество между заинтересованными сторонами.

В свою очередь, система мониторинга температур без использования блокчейн имеет следующие недостатки:

- не гарантирует целостность данных и защиту от подделок;
- не позволяет автоматически контролировать выполнение условий;
- может быть менее прозрачной и недоступной для всех участников системы.

В данной работе представлены существующие проблемы и причины низкого качества данных в системах, использующих блокчейн, и механизм контроля КД в блокчейн системе мониторинга. Механизм основан на смарт-контракте, написанном на языке Solidity, и реализован в приватной блокчейн-сети Ethereum на локальной машине. Получилось, что элементы контроля КД не оказывают сильного влияния на задержку транзакций.

В будущем планируется разработать несколько экземпляров смарт-контрактов, которые будут также содержать логику оценки КД, и определить оптимальный вариант в зависимости от количества показаний датчиков в транзакции.

Литература

1. Jayawardene V., Sadiq Sh. and Indulska M. An analysis of data quality dimensions // ITEE Technical Report. 2015. №2013-01 and 2015-027.
2. Degtyarev A., Bogdanov A., Shchegoleva N., Khvatov V. Data Quality in a Decentralized Environment // Lecture Notes in Computer Science. 2020. №12251. С. 58-71.
3. Gellings C.W., Samotyj M., Howe B. The future’s smart delivery system // IEEE Power and Energy Magazine. 2004. V. 2. № 5. P. 40-48.
4. Cappiello C., Comuzzi M., Daniel F., Meroni G. On the Need for Data Quality Assessment in Blockchains // IEEE Internet computing. 2021. V. 25. № 3. P. 71-78.
5. Casado-Vara R., de la Prieta F., Prieto J., Corchado J. M. Blockchain framework for IoT data quality via edge computing // blocksys 2018 - proceedings of the 1st blockchain-enabled networked sensor systems. part of sensys 2018.

УДК 004.855.5

ИСПОЛЬЗОВАНИЕ ТРАНСФОРМЕРОВ ДЛЯ ЛИНГВИСТИЧЕСКОЙ ИДЕНТИФИКАЦИИ АВТОРА

Хазагаров А.А.[1] (аспирант), **Воробьева А.А.**[1] (сотрудник)
Научный руководитель – к.т.н., Воробьева А.А.
1 – Университет ИТМО
e-mail: khazagarov@itmo.ru

Аннотация
В работе рассмотрено исследование использования нейронных сетей типа Трансформер и RNN для лингвистической идентификации автора. В качестве данных для обучения использовались тексты различных авторов, написанные на русском языке. В работе проведены эксперименты с обоими типами нейронных сетей, обучая их на разных наборах данных и сравнивая их результаты. Результаты показали, что нейронная сеть типа Трансформер показывает лучшие результаты по сравнению с рекуррентными нейронными сетями, особенно при обучении на большом объеме данных.

Ключевые слова
Нейронные сети, Трансформер, RNN, лингвистическая идентификация автора, анализ текстов, кибербезопасность.

Лингвистическая идентификация автора текстов является актуальной и важной проблемой, которая может быть применена в различных отраслях, таких как криминалистика, кибербезопасность и многие другие. Данная задача относится к области обработки естественного языка (NLP - natural language processing). Существует множество различных подходов к определению авторства, в том числе использование RNN (рекуррентные нейронные сети) [2]. Недавно представленная сеть типа Трансформер показывает хорошие результаты в области NLP. В этой статье мы сравним эти два подхода и оценим их сильные и слабые стороны в контексте идентификации авторства.

При анализе известных решений было выявлено, что использование традиционных статистических методов для лингвистической идентификации автора имеет ряд ограничений и недостатков [1]. В частности, эти методы часто требуют больших объемов данных для обучения, и могут иметь проблемы с классификацией авторов, которые пишут в разных жанрах и стилях. Рекуррентные нейронные сети (RNN) для лингвистической идентификации автора, также имеют некоторые ограничения, например производительность [1]. Предлагается рассмотреть новый тип нейронных сетей Трансформер, которые показывают лучшие результаты в задачах обработки естественного языка. Преимущества и особенности предлагаемого подхода включают в себя более высокую точность и эффективность в задачах лингвистической идентификации автора. Кроме того, нейронные сети типа Трансформер также обладают возможностью более эффективно работать с большими объемами текстовых данных и могут быть использованы для более широкого спектра задач анализа текстов. Цель данной статьи – рассмотреть применение нейронных сетей типа Трансформер для задачи лингвистической идентификации автора и сравнить полученные результаты с традиционным подходом на основе рекуррентных нейронных сетей.

Задача лингвистической идентификации автора заключается в определении авторства текста на основе некоторых языковых признаков [4]. Как правило, для этой задачи используются статистические методы, которые анализируют наборы языковых признаков, таких как частоты слов, пунктуации, длины предложений и других. В данной статье рассматривается подход на основе нейронных сетей, который позволяет автоматически извлекать признаки из текстов. Однако, в нейронную сеть нельзя подать

необработанный текст, для этого нужно привести его к векторной форме. Векторизация текста – это процесс преобразования текстовых данных в числовой вид, который может быть использован для обучения машинных моделей [1]. Существует несколько вариантов векторизации текста для данной задачи, включая:

– мешок слов (Bag of words) — это метод, в котором каждое уникальное слово из текста рассматривается как отдельный признак, и для каждого документа создается вектор, который содержит частоту встречаемости каждого слова. Этот подход не учитывает порядок слов в тексте;

– TF-IDF – это метод, использующий весовые коэффициенты для каждого слова, основанные на частоте его встречаемости в документе и количестве документов, в которых это слово встречается. Этот подход также не учитывает порядок слов в тексте;

– n-граммы – это метод, который учитывает порядок слов в тексте, разбивая его на последовательности из N слов. Вектор каждого документа строится на основе частоты встречаемости каждой N-граммы;

В данной работе было выбрано использование N-граммов для векторизации текста, так как данный метод показывает наилучшую точность. Это связанно с тем, что данный подход позволяет учитывать порядок слов в тексте, что может повысить точность классификации в задаче лингвистической идентификации автора. Оптимальное значение N было определено на основе экспериментов и составило 4. На рисунке представлена зависимость потребления памяти и точности идентификации от увеличения размерности n-граммы для LSTM. Все тексты, используемые в эксперименте, не подвергались нормализации, так как согласно исследованиям, это негативно влияет на точность идентификации [1].

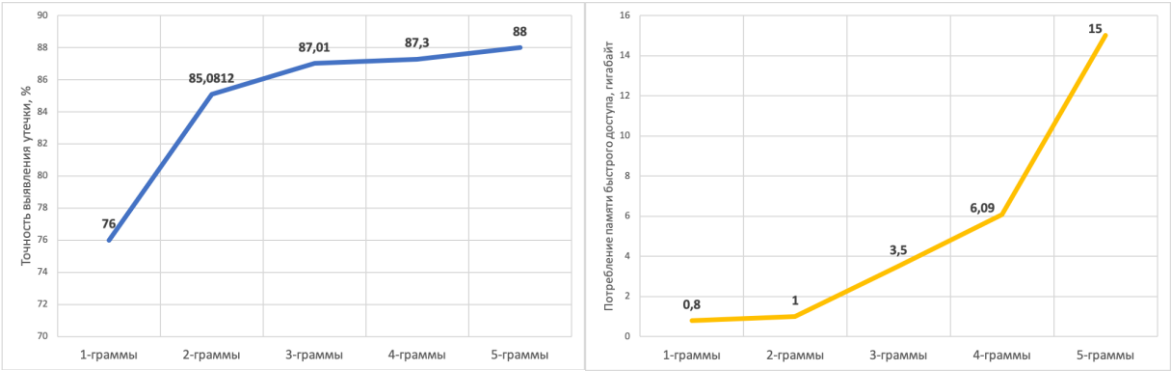


Рисунок. Зависимость основных характеристик от увеличения размерности n-грамм

Трансформер – это относительно новая архитектура нейронных сетей, которая была представлена в 2017 году [3]. Она была разработана для эффективной обработки последовательностей, таких как тексты, без использования рекуррентных связей. Трансформер состоит из нескольких блоков, каждый из которых включает в себя механизмы внимания и нормализации. Важным преимуществом Трансформера является возможность параллельной обработки данных, что делает его более эффективным, чем RNN, особенно для обработки больших объемов данных. Для сравнения эффективности двух подходов, был использован набор данных с текстами, написанными 1200 авторами. Каждый автор написал от 300 до 50 текстов. Для каждого автора была создана обучающая выборка, состоящая из 80% текстов, и тестовая выборка – из 20% текстов. Точность идентификации автора текста измеряется с помощью метрики accuracy, которая показывает, какая доля текстов была классифицирована правильно. В данной работе была использована 10-фолд кросс-валидация, которая позволяет оценить точность на различных наборах данных.

Результаты показали, что нейронная сеть на основе Трансформера превзошла по точности традиционный подход на основе LSTM на тестовой выборке. Точность классификации при использовании Трансформера составила 88,5%, в то время как при использовании LSTM она составила 87%, согласно таблице. Это говорит о том, что Трансформер более эффективен в задаче лингвистической идентификации автора текстов.

Таблица

Сравнение нейронных сетей

Тип сети	RNN	LSTM	GRU	CNN	Transformer
Точность	73	87	85	80	88,5

Следует отметить, что использование нейронных сетей типа Трансформер требует большего количества вычислительных ресурсов, чем традиционные методы. Также, наличие большого количества данных для обучения является важным фактором для получения хороших результатов. В целом, результаты данной работы показывают преимущества использования нейронных сетей типа Трансформер в задаче лингвистической идентификации автора текстов. Этот подход может быть использован в различных областях, где необходимо быстро и точно определить авторство текста.

Основные результаты данной работы заключаются в том, что нейронная сеть на основе Трансформера показала более высокую точность в задаче лингвистической идентификации автора текстов, чем традиционный подход на основе рекуррентных нейронных сетей. Точность классификации при использовании Трансформера составила 88,5%, в то время как при использовании LSTM она составила 87%. Эти результаты указывают на перспективность использования нейронных сетей типа Трансформер в задаче лингвистической идентификации автора текстов. Однако, необходимо учитывать, что для эффективного использования данного подхода требуется большое количество данных для обучения и вычислительных ресурсов. Одним из возможных направлений дальнейшего развития проблемы может быть использование лингвистической идентификации автора для выявления скрытых утечек информации в текстовых стегасообщениях. Также можно исследовать возможность комбинации нейронных сетей различных типов для достижения еще более высокой точности классификации. В целом, данная работа открывает новые возможности для использования нейронных сетей в задаче лингвистической идентификации автора текстов и может стать отправной точкой для дальнейших исследований в этой области.

Литература

1. Khazagarov A., Vorobeva A., Korzhuk V. Preventing Hidden Information Leaks Using Author Attribution Methods and Neural Networks//Proceedings of the 29th Conference of Open Innovations Association FRUCT. 2021. pp. 177-184.
2. Dmitrin Y.V. Comparison of deep neural network architectures for authorship attribution of Russian social media texts [Электронный ресурс]. http://www.dialog-21.ru/media/4560/_-dialog2018scopus.pdf (дата обращения: 26.04.2023).
3. Transformer – новая архитектура нейросетей для работы с последовательностями [Электронный ресурс]. https://habr.com/ru/articles/341240/ (дата обращения: 26.04.2023).
4. Воробьева А.А. Отбор информативных признаков для идентификации Интернет-пользователей по коротким электронным сообщениям // Научно-технический вестник информационных технологий, механики и оптики. 2017. Т. 17. № 1. С. 117–128. doi: 10.17586/2226-1494-2017-17-1-117-128.

УДК 004.056.5; 621.3.095; 621.373.8

РАЗРАБОТКА ЛАБОРАТОРНОГО СТЕНДА ДЛЯ ИССЛЕДОВАНИЯ ОПТИЧЕСКОЙ СИСТЕМЫ СВЯЗИ

Чернов Р.И.[1], Горошков В.А.[1], Островский Д.П.[1], Еремук В.В.[1]
Научный руководитель – д.т.н., 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».

Аннотация
Разработан лабораторный оптический стенд для исследования методов передачи информации в оптическом канале связи. Осуществлена разработка, изготовление и подбор элементов и узлов лабораторного стенда. Разработанный лабораторный стенд предназначен для исследования информационной защищенности и помехоустойчивости оптического канала связи.
Ключевые слова
Оптический канал связи, информационная защищенность, помехоустойчивость, лабораторный стенд, связь.

Актуальным направлением исследований являются исследования в области беспроводных оптических систем связи. Данные системы связи находят применение [1-3] в случаях, когда организация радиорелейных или оптоволоконных систем связи является нецелесообразной или невозможной, например, из-за особенностей ландшафта местности или загруженности радиочастотного пространства. Беспроводные оптические системы лазерной связи отличаются [4] низким энергопотреблением, высокой скоростью передачи информации, сложностью перехвата информации злоумышленником, небольшой продолжительностью и стоимостью проектирования и производства. Для исследования беспроводной оптической лазерной системы связи необходима разработка лабораторного стенда. Разработка лабораторного стенда осуществляется на основе структурной схемы оптического канала связи [5] с поляризационным разделением и дифференциальным приемом сигнала (рис. 1).

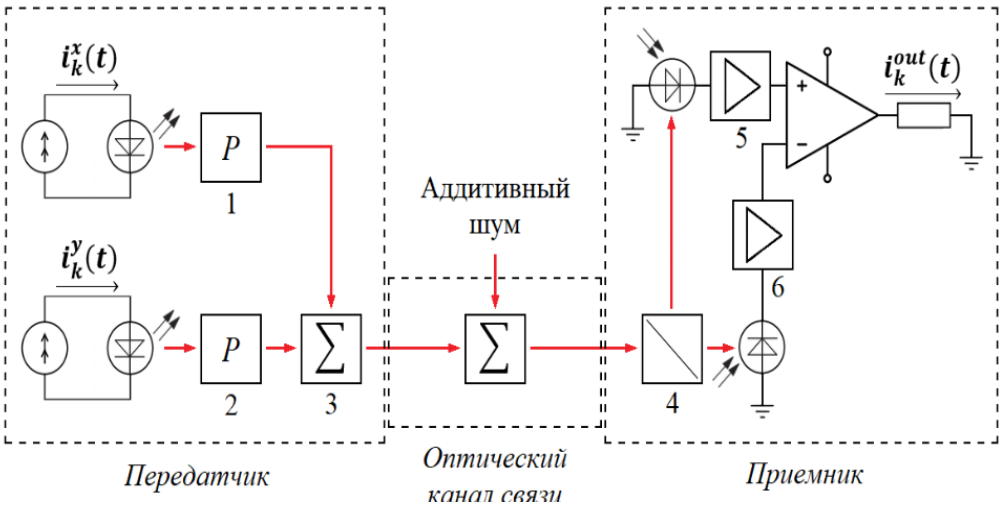


Рис. 1. Структурная схема оптического канала связи

На схеме под номером 1 и 2 изображены линейные поляризаторы, плоскости поляризации которых взаимно ортогональны. Под номером 3 изображен сумматор

оптических сигналов, 4 – двулучепреломляющая призма, 5 и 6 – усилители токовых сигналов, регистрируемых фотодиодами.

Таким образом, структура системы беспроводной оптической системы связи включает в себя передающее устройство, приемное устройство и канал связи, в котором передаваемый сигнал испытывает воздействие шума. Осуществлена разработка электрических схем передающего и приемного устройства. Для передающего устройства разработана схема питания лазерных диодов (рис. 2).

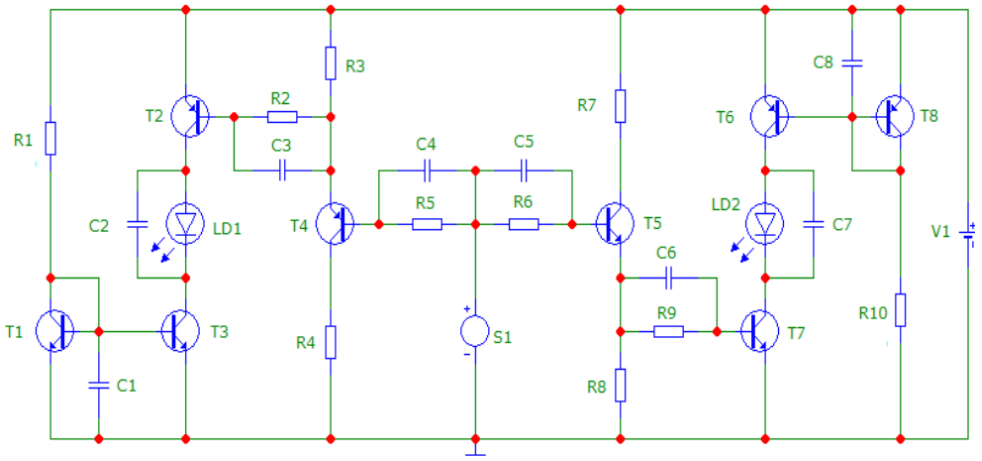


Рис. 2. Схема питания лазерных диодов

Схема питания лазерных диодов построена на базе комплементарной пары биполярных транзисторов. Стабилизация питания обеспечена за счет схемы токового зеркала. Ниже представлен график выходного токового сигнала схемы питания лазерных диодов (рис. 3). Небольшие выбросы тока не представляют опасности для работы лазерных диодов и не влияет на их интенсивность излучения.

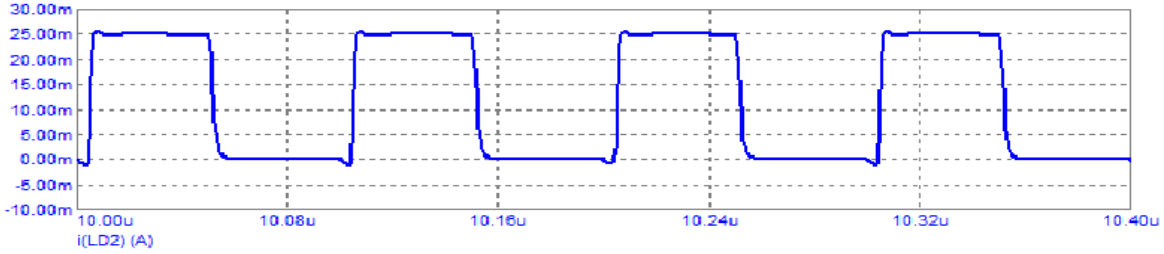


Рис. 3. Выходной токовый сигнал

Для приемного устройства разработана схема преобразования токовых сигналов фотодиодов в напряжение (рис. 4).

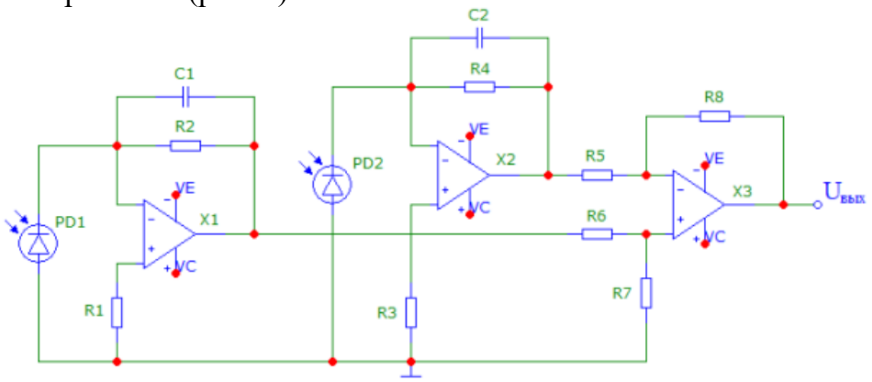


Рис. 4. Схема приема и усиления оптического сигнала

Схемы разработаны с учетом ограничений на величину тока, проходящего через лазерные диоды, а также с учетом быстродействия фотодиодов и необходимостью дифференцирования сигналов. Спектральные характеристики приемного и передающего устройства согласованы. Осуществлена разработка и изготовление оптической плиты на которой размещаются и позиционируются оптические компоненты.

Таким образом, разработан лабораторный стенд лазерной системы связи (рис. 5).

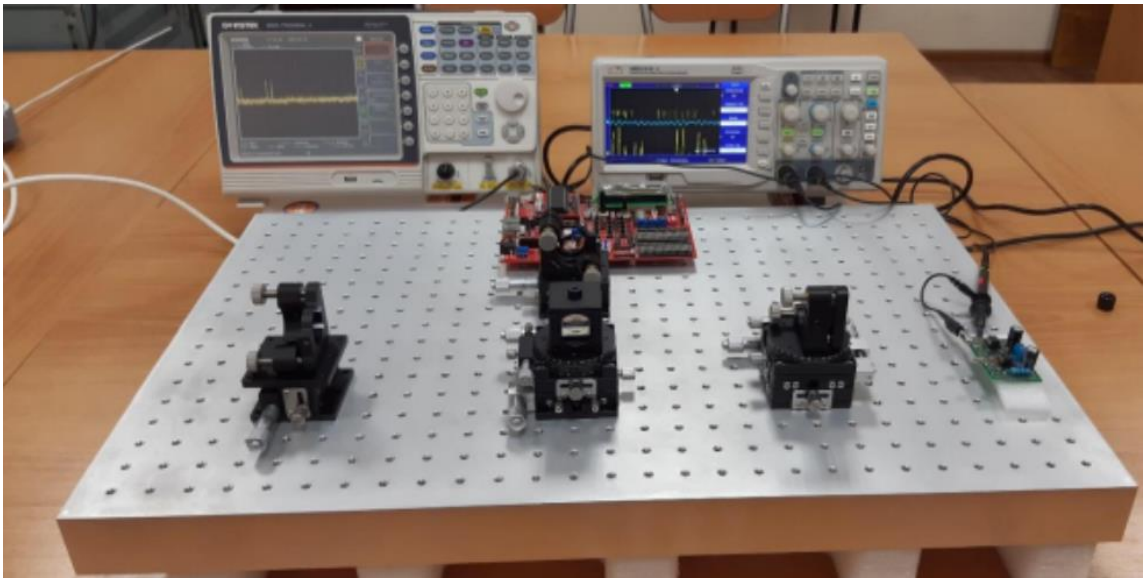


Рис. 5. Лабораторный стенд для исследований беспроводной лазерной системы связи

Лабораторный оптический стенд может быть использован для проведения экспериментов по исследованию метода поляризационного разделения и дифференциального приема сигнала, исследованию приема лазерного излучения, отраженного от поверхностей различной структуры, исследованию пропускной способности оптического канала связи с точки зрения использования ресурса частота-время и других.

Литература

1. Роенков Д.Н., Яронова Н.В. Атмосферные оптические линии связи //Автоматика, связь, информатика. 2016. №. 11. С. 7-10.
2. Белов В.В. и др. Атмосферные бистатические каналы связи с рассеянием. Часть 1. Методы исследования //Оптика атмосферы и океана. 2013. Т. 26. №. 4. С. 261-267.
3. Воронцов М.А. и др. Частота появления ошибочных битов в системах беспроводной оптической связи с частично когерентным передающим пучком //Оптика атмосферы и океана. 2012. Т. 25. №. 11. С. 936-940.
4. Смайлов Н.К., Абдыкадыров А.А., Керей А.Ж., Жадигер Т.А. Моделирование оптической связи в свободном пространстве //Universum: технические науки. 2022. №. 4-4 (97). С. 14-21.
5. Разработка методов повышения уровня информационной защищенности и помехоустойчивсти оптического канала связи. XI Конгресс молодых учёных. Тихомиров А.В., Чернов Р.И., Еремук В.В., Козин О.В. 2022. Сборник научных трудов. Том 1. С. 477-479.

УДК 004.052.32

РАЗРАБОТКА АРХИТЕКТУРЫ СИСТЕМЫ МОНИТОРИНГА БЕСПИЛОТНОГО МАНЕВРОВОГО ЛОКОМОТИВА

Якимов Я.Д.[1]
Научный руководитель – к.т.н., доцент ФБИТ Попов И.Ю.[1]
1 – Университет ИТМО
e-mail: yaroslav@itmo.ru

Аннотация
В данной работе рассматривается система мониторинга беспилотного маневрового локомотива. Предложенная система направлена на улучшение показателей функциональной безопасности блока обнаружения препятствий, показателя среднего времени восстановления, благодаря внедрению инструментов аналитики и выявления аномалий.
Ключевые слова
Функциональная безопасность, упреждающее техническое обслуживание, обнаружение аномалий, автономная железнодорожная система, мониторинг, надежность.

На данный момент актуальными темами технологического развития являются популяризация и автоматизация перевозок с помощью беспилотных транспортных средств (далее БТС). Современные БТС имеют внутри себя достаточно сложный и многокомпонентный аппаратно-технический комплекс, который использует различные технологии [1]. Обеспечение эффективного и бесперебойного функционирования системного, сетевого и вспомогательного оборудования является ключом к работе БТС без сбоев и с наименьшим риском. В связи с этим возникает необходимость оперативного сбора и агрегации разнородных метрик, поступающих от различных узлов, например, сетевого оборудования или сенсоров, для их дальнейшей обработки. Этот процесс существенно усложняется необходимостью наличия человекочитаемого представления данных большого объема о техническом состоянии транспортных систем в удобном для анализа и оценки виде.

На пилотируемых поездах вывод данных протоколируется только в виде текстовых файлов на самих поездах и доступны машинисту на приборной панели. Кроме того, аппаратная составляющая также не позволяет осуществлять сбор и отправку информации о данных параметрах во внешние системы, например, установленные на железнодорожных станциях. Описанная ситуация повлекла за собой слабую развитость области мониторинга и аналитики состояния железнодорожных локомотивов. Однако, при разработке беспилотных локомотивов, благодаря наличию вычислительных мощностей и технической оснащенности, в том числе и модулями связи, возможности агрегации данных стали более доступными. Это позволяет уже сейчас выявлять и учитывать появляющиеся ошибки в реализации программного обеспечения (далее ПО), например, при управлении аппаратными системами. Впоследствии можно будет использовать данную информацию для исправления выявленных ошибок на конкретных граничных параметрах.

Необходимость в проведении анализа различных метрик очевидна, так как это напрямую оказывает влияние на функциональную безопасность как беспилотных маневровых локомотивов, так и в целом БТС. Сбои в работе систем детектирования несут в себе угрозу жизни пассажиров, пешеходов и экономические потери от возможных поломок [2].

Соответственно, целью работы является повышение отказоустойчивости системы блока обнаружения препятствий беспилотного маневрового за счет:

1) включения в архитектуру блока системы мониторинга;
2) выявления критических показателей для безопасной эксплуатации БОП;

3) уменьшение времени на обнаружение возможного отказа.

В контексте функциональной безопасности для оценки надежности и эффективности систем и компонентов можно использовать несколько показателей эффективности. Эти показатели включают, среди прочего, среднее время наработки на отказ (MTBF), среднее время до отказа (MTTF) и среднее время ремонта (MTTR) [3]. Далее рассматривается обоснование выбора MTTR в качестве ключевого показателя эффективности в рамках функциональной безопасности по сравнению с другими параметрами.

MTTR конкретно измеряет среднее время, необходимое для восстановления вышедшей из строя системы или компонента до их рабочего состояния. Этот показатель напрямую связан с эффективностью процесса технического обслуживания и ремонта, которые являются важнейшими аспектами функциональной безопасности приложений. Уменьшенный MTTR подразумевает сокращение времени нахождения в нерабочем состоянии, тем самым повышая доступность системы и обеспечивая быстрое восстановление функций, связанных с функциональной безопасностью.

В приложениях важно обеспечить доступность функций, когда это необходимо для снижения рисков, связанных с опасными явлениями. Сосредоточив внимание на MTTR, организации могут уделять особое внимание своевременному восстановлению этих функций и поддерживать высокий уровень доступности системы. С другой стороны, хотя наработка на отказ и MTTF могут предоставлять информацию об ожидаемой частоте или времени до отказа, они напрямую не влияют на доступность функций, связанных с безопасностью, во время и после сбоев.

Сосредоточив внимание на MTTR, организациям рекомендуется применять стратегии упреждающего технического обслуживания, такие как техническое обслуживание на основе состояния (CBM) и прогнозирующее техническое обслуживание (PdM). Эти стратегии необходимы для минимизации риска непредвиденных сбоев и обеспечения максимальной доступности системы в приложениях функциональной безопасности.

Сокращение MTTR с помощью систем мониторинга позволяет организациям демонстрировать соответствие стандартам функциональной безопасности и нормативным актам, таким как IEC 61508, ISO 61511 и другим [4]. Эти стандарты требуют оценки и обеспечения надежности и доступности системы, а также применения соответствующих мер по снижению рисков. Системы мониторинга предоставляют необходимые данные для верификации и валидации функций, связанных с безопасностью, и оценки производительности системы в соответствии с установленными требованиями. Сокращая MTTR и повышая надежность системы, системы мониторинга способствуют выполнению нормативных требований, тем самым обеспечивая безопасность людей, имущества и окружающей среды. Как показано на рисунке 1, MTTR складывается из четырех составляющих: обнаружение отказа, подготовка, сам ремонт и возвращение в работу.

Как следствие, системы мониторинга приобретают все большее значение в различных инфраструктурных приложениях, где они способствуют эффективному управлению и техническому обслуживанию критически важных активов. Далее представлены основные принципы построения систем мониторинга для различных инфраструктурных приложений, подчеркивается важность процессов сбора данных, анализа и принятия решений в этих системах. Основные принципы построения мониторинга любой инфраструктуры:

- устойчивость и высокая доступность системы мониторинга;
- проактивный мониторинг;
- формирование оповещений и проверок;
- масштабируемость с учетом расширения инфраструктуры;
- аналитика.

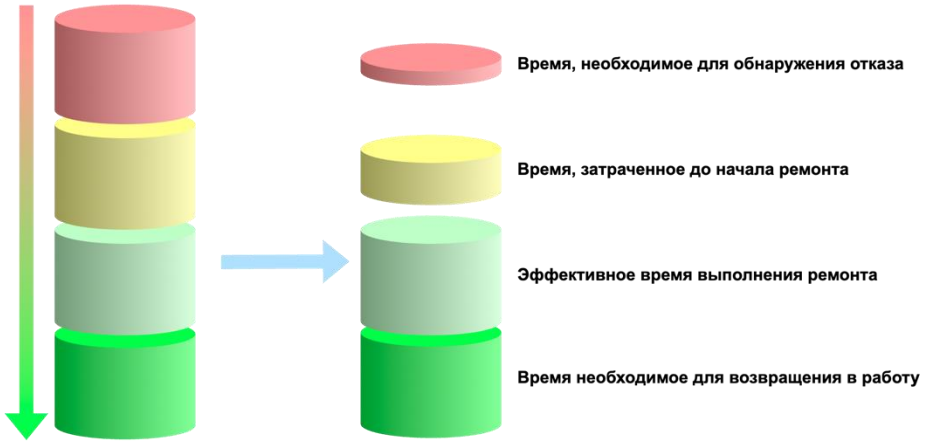


Рис. 1. Изменение показателя MTTR

Предлагаемая архитектура состоит из клиентов, которые включают в себя микросервисы, аппаратное обеспечение и другие системы. За этими клиентами следят агенты мониторинга, которые собирают и передают данные в базу данных временных рядов. Доступ к этой базе данных осуществляется с помощью программного обеспечения для оповещения, визуализации, аналитики и проактивного мониторинга технического обслуживания, что позволяет операторам взаимодействовать с этими инструментами и обеспечивать надлежащее функционирование автономной железнодорожной системы. Разработанная архитектура представлена на рисунке 2.

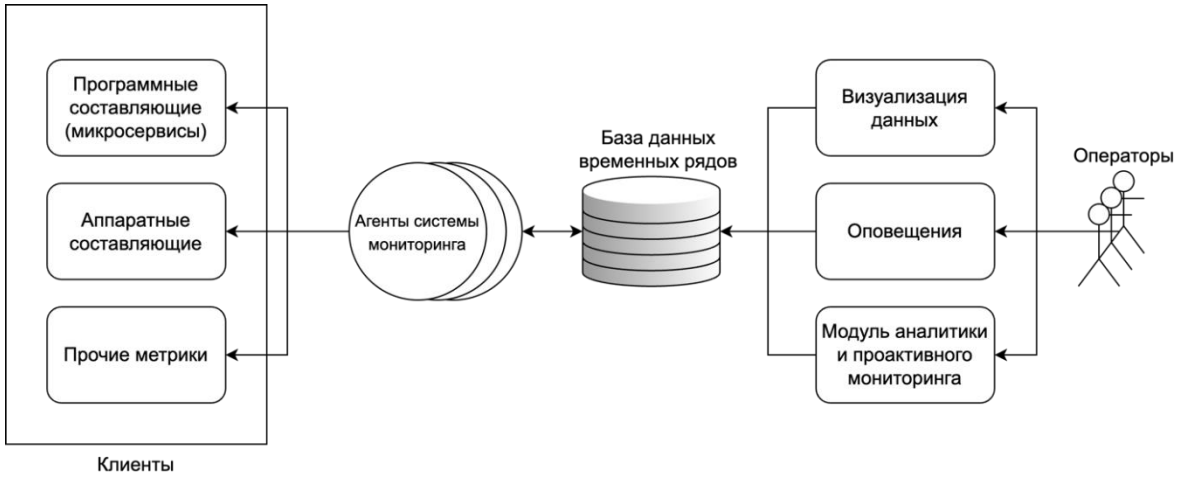


Рис. 2. Визуальное представление архитектуры мониторинга

Рассмотрим более подробно функционал каждого из элементов данной архитектуры. Клиенты формируют основу архитектуры, состоящей из микросервисов, аппаратного обеспечения и других систем, которые совместно работают для поддержания работы автономного поезда. Благодаря модульному подходу архитектура позволяет легко добавлять, модифицировать или удалять отдельные компоненты, что способствует масштабируемости и адаптируемости системы. Агенты по мониторингу играют жизненно важную роль в поддержании функциональной безопасности, постоянно отслеживая клиентов и собирая все необходимые метрики. Эти агенты должны быть высоконадежными и эффективными, гарантируя быструю обработку данных и передачу в базу данных. База данных временных рядов служит центральным хранилищем для хранения данных, собранных агентами мониторинга, и управления ими.

Используя формат временных рядов, база данных обеспечивает эффективное хранение, извлечение и запрос данных с отметками времени, позволяя операторам отслеживать производительность и обнаруживать тенденции или закономерности с течением времени. Программное обеспечение оповещения действует как первая линия защиты в архитектуре, гарантируя оперативное уведомление операторов о любых потенциальных проблемах или аномалиях в системе. Анализируя данные из базы данных временных рядов, программное обеспечение генерирует своевременные оповещения, позволяя операторам принимать быстрые и надлежащие меры для поддержания функциональной безопасности. Программное обеспечение для визуализации предоставляет операторам четкие и всесторонние визуальные представления о производительности и статусе системы. Благодаря интуитивно понятным и интерактивным информационным панелям операторы могут легко отслеживать и оценивать работоспособность системы, что позволяет им принимать обоснованные решения и эффективно поддерживать функциональную безопасность. Программное обеспечение для аналитики и проактивного мониторинга технического обслуживания использует передовые алгоритмы и дополнительно методы машинного обучения для анализа данных из базы данных временных рядов. Рассмотрим такие ключевые аспекты предложенной системы, как модульность, управление данными, упреждающее обслуживание и общая надежность системы.

Модульность: Модульный характер предлагаемой архитектуры с ее различными клиентами обеспечивает большую масштабируемость и адаптивность по сравнению с традиционными монолитными системами. Возможность добавлять или изменять отдельные компоненты без ущерба для всей системы обеспечивает более гибкое и экономичное техническое обслуживание и модернизацию.

Управление данными: База данных временных рядов в предлагаемой архитектуре обеспечивает высокоэффективный метод хранения, извлечения и анализа данных. Формат временных рядов обеспечивает быстрый доступ к историческим данным и данным в режиме реального времени, позволяя операторам выявлять тенденции и принимать обоснованные решения относительно функциональной безопасности.

Упреждающее техническое обслуживание: Интеграция аналитики и программного обеспечения для проактивного мониторинга в архитектуру отличает ее от других методологий, которые могут полагаться исключительно на стратегии реактивного технического обслуживания. Используя передовые алгоритмы и методы машинного обучения, программное обеспечение позволяет операторам прогнозировать и устранять потенциальные проблемы до того, как они приведут к сбоям системы или угрозе безопасности.

Общая надежность системы: Направленность предлагаемой архитектуры на непрерывный мониторинг, эффективное управление данными и упреждающее техническое обслуживание способствует общему повышению уровня надежности системы.

Предлагаемая архитектура предлагает комплексное и передовое решение для поддержания функциональной безопасности автономных поездов. Благодаря интеграции всех элементов архитектура отвечает всем представленным ранее требованиям и способна конкурировать с другими архитектурами. Этот подход предполагает поддержание высокого уровня функциональной безопасности и надежности, одновременно способствуя развитию технологии беспилотных поездов.

В рамках данной работы были проведены эксперименты, используя различные методы обнаружения аномалий, которые призваны снизить показатель MTTR. Для проведения экспериментов вводилось допущение, что тренировочные данные не имеют аномалий. Количество экспериментов выбиралось эмпирически на основании опыта аналогичных исследований по расчету данных алгоритмов. Приведем усредненные значения времени вычисления в таблице. В то время как тест Граббса и метод 3-Сигм

заметно быстрее из-за их простоты и зависимости от статистических показателей, PCA и K-means относительно быстрее среди методов, основанных на машинном обучении. Напротив, методы глубокого обучения, такие как автокодеры, LSTM и CNN, имеют более длительное время вычислений, поскольку они предполагают обучение сложных моделей. SVM представляет собой усредненный метод между традиционным машинным обучением и методами глубокого обучения с точки зрения времени вычислений.

Таблица

Время вычисления для различных методов обнаружения аномалий

Метод	Время вычисления (секунды)
K-средних	2,50
PCA	1,80
Тест Граббса	0,01
3-сигма	0,01
Автоэнкодеры	120
LSTM	150
CNN	180
SVM	4

В заключение, сравнение методов обнаружения аномалий, включая K-средних, PCA, тест Граббса, 3-Сигма, автокодеры, LSTM, CNN и SVM, показывает широкий спектр методов с различными характеристиками производительности. Традиционные статистические методы, такие как тест Граббса и 3-Сигма, обеспечивают быстрое время вычислений, но могут быть не столь эффективны при обработке сложных шаблонов данных или данных большой размерности. Методы, основанные на машинном обучении, такие как K-средних и PCA, обеспечивают баланс между вычислительной эффективностью и точностью обнаружения, что делает их подходящими для наборов данных среднего размера или ситуаций, когда требуется баланс между скоростью и производительностью. При выборе подходящего метода обнаружения аномалий следует учитывать несколько факторов, таких как размер набора данных, сложность шаблонов данных, размерность данных и желаемый баланс между вычислительной эффективностью и точностью обнаружения. Статистические методы, такие как тест Граббса и 3-Сигма, рекомендуются для простых наборов данных или ситуаций, когда решающее значение имеет быстрое вычисление. Методы, основанные на машинном обучении, такие как K-средних и PCA, подходят для наборов данных среднего размера, в то время как методы глубокого обучения, такие как автокодеры, LSTM и CNN, лучше всего подходят для больших многомерных данных со сложными шаблонами. SVM предлагает универсальное решение для широкого спектра сценариев обнаружения аномалий, сочетающее в себе преимущества как традиционного машинного обучения, так и методов глубокого обучения.

В рамках дальнейшей работы для оперативного построения оповещений были выбраны методы 3-сигм и К-средних, как наиболее подходящие для данных, полученных в режиме реального времени. Данные методы имеют малое время вычисления и достаточную точность для выявления потенциальных аномалий при работе систем блока обнаружения препятствий беспилотного поезда.

В традиционном сценарии время, затрачиваемое на расследование проблем и MTTR, может быть значительно больше из-за отсутствия стратегий упреждающего

обслуживания и зависимости от реактивного обслуживания. В рамках данной работы приводится первичная оценка эффективности предложенной архитектуры и были рассмотрены несколько примеров для оценки времени, затрачиваемом на исследование проблем и MTTR в обычных сценариях по сравнению с сценариями с предлагаемой архитектурой. Данные о времени являются усредненными в рамках испытаний беспилотного маневрового локомотива и предоставлены инженерами АО «НИИАС». Приведем один из экспериментов, описываемый следующим образом: микросервис, отвечающий за обработку облака точек лидара начал выдавать некорректные значения по обнаруженным объектам, причиной этого являлась утечка памяти в другом микросервисе, вызвавшая дефицит оперативной памяти на вычислителе.

В традиционной системе команда технического обслуживания получает предупреждение о некорректном обнаружении объекта в поле зрения блока обнаружения препятствий после того, как сбой уже произошел. Команда тратит 2 часа на диагностику проблемы и выявление первопричины. После определения проблемы команда тратит еще 2 часа на ее устранение и восстановление работоспособности системы. В этом сценарии время, затраченное на исследование проблемы, составляет 2 часа, а MTTR - 4 часа.

В предлагаемой архитектуре со стратегиями упреждающего технического обслуживания для мониторинга железнодорожной системы в режиме реального времени используются методы 3 сигмы и K-средних. При обнаружении аномалии в проблемном микросервисе команда технического обслуживания немедленно получает предупреждение о превышении использования оперативной памяти, до того, как система выйдет из строя. Они тратят всего 1 час на диагностику проблемы, поскольку система предоставляет им подробную информацию об аномалии и возможных корреляциях. Затем команда тратит еще 2 часа на устранение проблемы и предотвращение сбоя. В этом сценарии время, затраченное на исследование проблемы, составляет 1 час, а MTTR - 3 часа.

Предварительная оценка эффективности предложенной архитектуры мониторинга и аналитики составляет не менее 25% по сравнению с традиционными сценариями. Сокращение времени расследования и MTTR способствуют повышению общей функциональной безопасности, поскольку потенциальные проблемы обнаруживаются и устраняются до того, как они перерастут в системные сбои. В свою очередь, это приводит к повышению надежности и доступности системы, сводя к минимуму риск аварий или сбоев в работе автономной железнодорожной системы.

Список использованных источников

1. Van Brummelen J. et al. Autonomous vehicle perception: The technology of today and tomorrow //Transportation research part C: emerging technologies. 2018. Т. 89. С. 384-406.
2. Leilabadi S. H., Schmidt S. In-depth analysis of autonomous vehicle collisions in california //2019 IEEE Intelligent Transportation Systems Conference (ITSC). IEEE. 2019. С. 889-893.
3. Alavian P. et al. The (α, β)-precise estimates of MTBF and MTTR: Definitions, calculations, and induced effect on machine efficiency evaluation //IFAC-PapersOnLine. 2019. Т. 52. №. 13. С. 1004-1009.
4. Smith D.J., Simpson K.G.L. The Safety Critical Systems Handbook: A Straightforward Guide to Functional Safety: IEC 61508 (2010 Edition), IEC 61511 (2015 Edition) and Related Guidance. Butterworth-Heinemann. 2020.

ИНФОКОММУНИКАЦИОННЫЕ ТЕХНОЛОГИИ

УДК 004.78

THE DEVELOPMENT OF IOT APPLICATION ON EDGE COMPUTING PLATFORM

Dayoub A.[1], Filyanin I.V.[1]
Research supervisor - associate prof., candidate of technical sciences Grudinin V.A.[1]
1 – ITMO university

Abstract
As the number of the Internet of things (IoT) devices connected to the internet increases, the amount of data explodes. leading to an increased demand for bandwidth during data transmission and heightened computing and storage requirements for central servers. To address these issues and enable real-time data processing while reducing network bandwidth utilization, edge computing has emerged as a potential solution. This article is devoted to the development of the IoT application on the edge computing platform to solve problems like latency and network bandwidth occupation for IoT applications.

Keywords
Multi-access Edge Computing (MEC), Internet of things IoT, IoT application.

Introduction

Multi-access Edge Computing (MEC) is considered as an effective method to solve the limited resources problems in IoT. In MEC environments, servers with certain computing capabilities are deployed on the edge of the network, which are closer to the IoT devices and can provide computing services for the IoT devices.

According to the According to the Cisco Annual Internet Report (2018-2023) [1], the number of IoT devices connected to the internet is expected to reach 30.9 billion by 2023, and IoT devices are expected to generate 73.1 zettabytes (ZB) of data by 2025, up from 18.3 ZB in 2019. This includes data generated by all types of IoT devices, including sensors, cameras, and other connected devices. With all this generated data and the increasing number of IoT devices the demand for low latency applications and high bandwidth networks will increase, and the MEC technology is the best solution meet the needs of the IoT application and that's what we are discussing in this paper.

While traditional cloud computing offers benefits such as abundant computing resources, scalability, and cost-effectiveness, it presents several challenges, including single-point-of-failure risks and significant latency (potentially hundreds of milliseconds) between the user and the centralized cloud. In contrast, many IoT applications require decentralized systems that incorporate location awareness, reliability, and low latency. For time sensitive IoT services, latency must not exceed 1 millisecond, and reliability must be no lower than 99.99% [2].

Main part

IoT edge computing refers to the processing of data generated by Internet of Things (IoT) devices at or near the data source, rather than sending all the data to a centralized server. This approach enables faster analysis and response to the collected data, as well as reducing the amount of data that needs to be transmitted.

A conventional IoT architecture consists of three layers: the perception layer, network layer, and cloud platform layer [3]. The perception layer primarily collects field data and controls the system, while the network layer is responsible for transmitting perception data and control instructions. The platform layer completes tasks such as equipment and connection management, application enablement, and business analysis.

The application of edge computing in IoT scenarios provides an effective supplement to traditional IoT systems. Compared with the conventional IoT architecture, the IoT edge

architecture includes an additional edge layer located between the perception and platform layers as shown in figure 1.

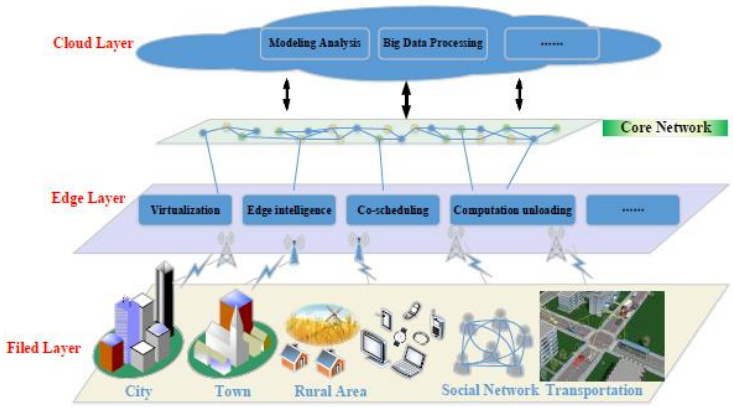


Fig. 1. IoT Edge Architecture

The edge layer serves as a functional sink for the cloud platform, capable of conducting real-time data analysis and processing. It also facilitates real-time control and decision-making within the system, which reduces the use of unnecessary data and conserves network bandwidth, minimizes system delays, and mitigates energy consumption caused by data transmission. Additionally, the edge-side IoT system operates as a closed-loop system, thereby significantly enhancing the system's robustness and reliability [4].

The development of IoT application on Edge computing platform

Our system has three main parts, which are the IoT sensors, the Edge computing platform and the cloud as shown in figure 2.

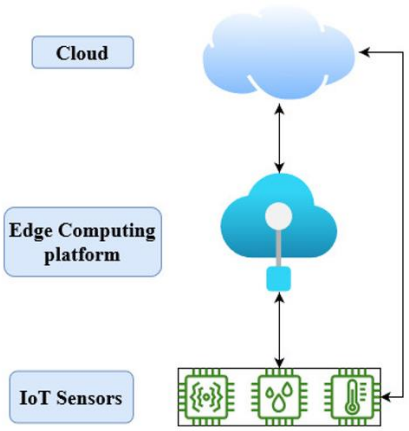


Fig. 2. The schema of the IoT edge application

The process is to deploy the IoT application on the edge computing platform, the IoT application simply will receive the data from IoT devices and then apply processing on this data.

To apply this scenario, we will use AdvantEDGE [5], which is a Mobile Edge Emulation Platform (MEEP) based on Docker and Kubernetes. The platform offers an emulation environment that allows for the testing and experimentation of edge computing technologies, applications, and services.

Here to determine the performance we should measure many parameters like latency bandwidth and error rate. And thanks to AdvantEDGE we can measure all these parameters.

After that we will compare this to the case when the IoT applications run on cloud and see the difference between the two cases.

The first step of the development is to design the network model of the scenario as shown in figure 3. In our network we have the three main components which are UE, edge server and cloud. Also, we have other network components like zone, PoA, operator and Internet.

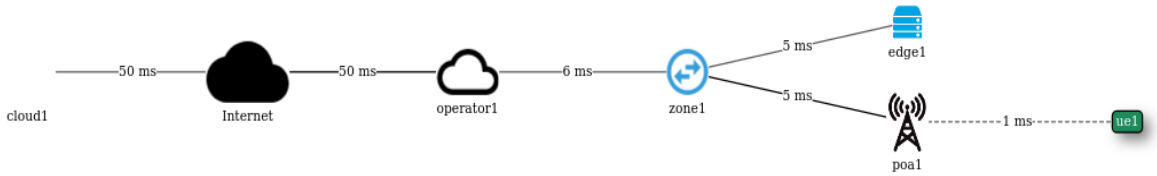


Fig. 3. Network design

In our network components we set the Uplink/Downlink throughput to 1000/1000 Mbps, and the latency between each two components as recommended from the platform as shown in figure 3.

After designing our network, we added the application that will simulate the IoT application and for that we will use iPerf which is a tool for network performance measurement and tuning. iPerf is a versatile tool that is compatible across multiple platforms, capable of generating standardized performance metrics for any network. This tool offers both client and server functionality and can establish data streams to measure the throughput between the two endpoints in one or both directions. The typical output generated by iPerf includes a time-stamped report that provides information on the volume of data transferred and the measured throughput [6].

Technically we will have three applications with iPerf, one works as a client and the other two work as servers as shown in figure 4.

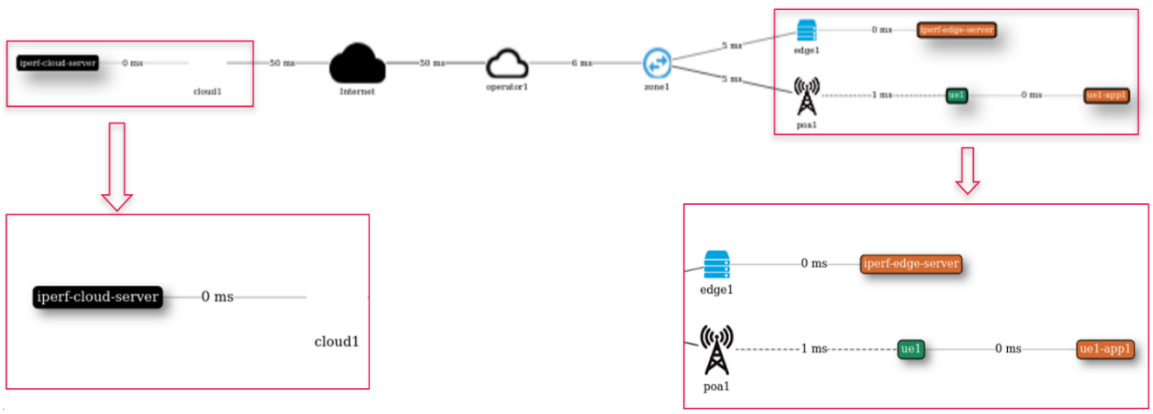


Fig. 4. Network design with the iPerf applications

The application that runs on the UE works as a client and sends the data to the server and the edge server, it uses UDP protocol, and it sends 50 Mbps. The applications that run on edge and cloud work as a server and receive the data from the client, it used UDP protocol and listen on port 80.

After we configure the three main applications we come across the execution of the model and see the results, and we measure the latency as shown in figure 5.

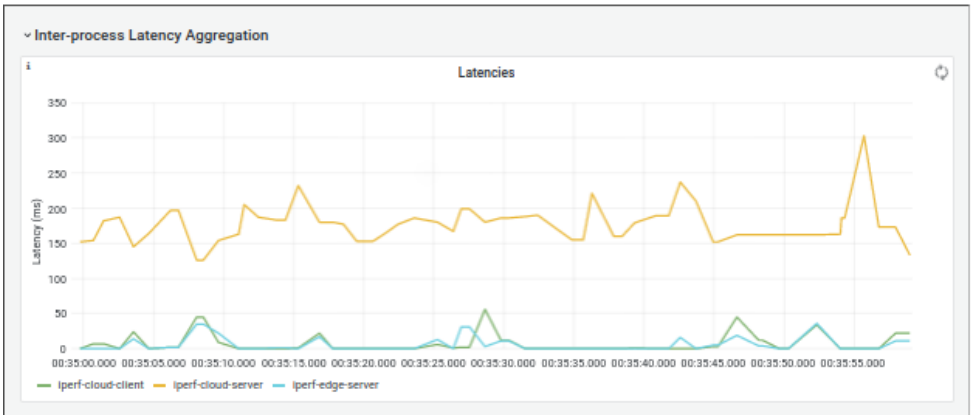


Fig. 5. The latency from the UE to edge and cloud

We can clearly see that the latency between the UE and the cloud is very big for IoT applications (150 -200) ms, but the latency from the UE to the edge server is (0-30) ms which is better for IoT applications.

Conclusion

In this paper we have presented a simulated model of an IoT application developed on an edge computing platform, and we come across the fact that using MEC with IoT we will be able to achieve latency around (0-30 ms) rather than (150-200 ms) using only cloud and reduce the data load over the network. This model will open the doors to other studies on the idea of developing the IoT applications with the help of MEC, because the MEC is one of the main enabler technologies for IoT. What we can recommend after this study is to use edge server for critical application and use cloud for non-critical applications.

References

1. Cisco, white paper, Cisco Annual Internet Report (2018-2023) https://www. cisco. com/c/ en/us/ solutions/collateral/executive-perspectives/annual-internet-report/white-paper-c11-741490.html. (дата обращения: 21.02.2023).
2. Mahbub M., Gazi Md. S. Apu, Provat S. al Arabi, and Islam Md. S. “Multi-Access Edge Computing-Aware Internet of Things: MEC-IoT,” in 2020 Emerging Technology in Computing, Communication and Electronics (ETCCE). Dec. 2020. pp. 1–6.
3. Jabraeil Jamali M.A. et al. “IoT architecture,” Towards the Internet of Things: Architectures, Security, and Applications, pp. 9–31. 2020.
4. Pan J. and Elhannon J. Mc. “Future Edge Cloud and Edge Computing for Internet of Things Applications,” IEEE Internet Things J. vol. 5. no. 1. pp. 439–449. Feb. 2018.
5. Advantedge platform, on github [electronic recourses] https://interdigitalinc.github.io/AdvantEDGE/. (дата обращения: 21.02.2023).
6. Iperf tools. [electronic recourses] https://iperf.fr/. (дата обращения: 21.02.2023).

УДК 004.934.2

АНАЛИЗ ВОВЛЕЧЕННОСТИ И ЭМОЦИЙ СОБЕСЕДНИКОВ ВИРТУАЛЬНОЙ КОММУНИКАЦИИ

Двойникова А.А.[1] (младший научный сотрудник)
Научный руководитель – д.т.н., профессор Карпов А.А.[1]
1 – Санкт-Петербургский Федеральный исследовательский центр Российской академии наук
e-mail: dvoynikova.a@iias.spb.su

Работа выполнена при поддержке Совета по грантам Президента РФ (грант № НШ-17.2022.1.6).

Аннотация
В работе проводится исследование многомодальных и многозадачных подходов к анализу вовлеченности, полярности и интенсивности эмоций собеседников виртуальной коммуникации. Для экспериментальных исследований использовался аудиовизуальный корпус данных RECOLA. Предложенный многомодальный и многозадачный подход основан на применении спектрограмм и нейросетевых признаков ResNet, для акустической и визуальной информации соответственно, которые подаются на вход сверточным нейронным сетям. С помощью предложенного подхода удалось достичь точности распознавания 3 классов вовлеченности, полярности и интенсивности эмоций по показателю взвешенной полноты 69,29%, 59,02%, 42,62%, соответственно.

Ключевые слова
Вовлеченность, интенсивность эмоций, полярность эмоций, многомодальность, многозадачность.

В последнее время люди стали чаще коммуницировать между собой посредством систем телеконференций. Во время онлайн встреч понимание невербальных знаков может быть затруднено. Причиной этому могут быть различные факторы такие, как неустойчивое интернет-соединение, видеокамера с низким разрешением, наложение акустических шумов во время разговора и другие. Разработка системы анализа вовлеченности и эмоций собеседников во время разговора в системах телеконференций поможет решить описанные выше проблемы.

Исследуемая задача находится на стыке двух областей: психологии и искусственного интеллекта. Для того, чтобы построить эффективную систему распознавания вовлеченности и эмоций, необходимо проанализировать эти состояния человека в области психологии. Вовлеченность – это проявление эмоциональной, когнитивной и поведенческой составляющих человека с различной интенсивностью [1]. Как видно из определения, вовлеченность и эмоции коррелирующие понятия. Многие психологи в своих исследованиях выявляют следующую закономерность: если студент испытывает позитивные эмоции, то его вовлеченность в образовательный процесс повышается и наоборот. Также чрезмерная интенсивность любой эмоции приводит к пониженной вовлеченности студента [1]. Исходя из особенностей вовлеченности и эмоций с точки зрения психологии необходимо разработать многозадачный подход к распознаванию исследуемых состояний диктора. Также стоит учесть, что на вовлеченность влияют не категории эмоций, а интенсивность и полярность этих эмоций. Поэтому разрабатываемый подход для анализа невербальных проявлений коммуниканта должен содержать в себе анализ вовлеченности, полярности и интенсивности эмоций. На сегодняшний день тема автоматического многозадачного распознавания эмоций и вовлеченности является нерешенной и актуальной задачей.

Многомодальные подходы к распознаванию вовлеченности и эмоций позволяют построить более робастную систему [2]. Наиболее репрезентативными модальностями для проявления вовлеченности и эмоций дикторов являются видео и аудио [1]. В данной

статье анализируется влияние многозадачного и многомодального подхода к распознаванию вовлеченности, полярности и интенсивности эмоций собеседников в виртуальную коммуникацию.

Анализ предыдущих исследований. В последнее время тема автоматического распознавания вовлеченности студентов в обучающий процесс, собеседника в коммуникацию или пользователя в онлайн-игры и пр. является актуальной и востребованной [1]. Большинство существующих исследований направлены на анализ визуальной информации человека [3], реже на анализ акустических характеристик речи во время разговора [4]. На сегодняшний день не было найдено исследований, направленных на анализ вовлеченности по лингвистической составляющей речи собеседника [1]. Это может быть связано с тем, что видео модальность является более репрезентативной для анализа вовлеченности, даже в таких случаях, когда собеседник не ведет диалог. Для разработки наиболее робастной системы анализа вовлеченности участников исследователи используют многомодальные подходы [1].

Тема автоматического анализа эмоционального состояния диктора на протяжении многих лет активно исследуется учеными. Существует два направления исследований: анализ категориальных эмоций и анализ непрерывных эмоций [5]. Как правило, непрерывные эмоции описываются с помощью двух факторов: интенсивности и полярности эмоций. Исследований в области автоматического распознавания непрерывных эмоций значительно меньше, чем категориальных, ввиду малого количества информационных ресурсов для этой задачи. Поэтому задача анализа непрерывных эмоций является актуальной. В задачах автоматического распознавания эмоций применяется анализ визуальной [6], акустической [7] и лингвистической информации [8]. При этом также существует большое количество исследований, направленных на многомодальный анализ эмоций [2]. Данные исследования показывают, что многомодальный анализ позволяет достичь наибольшей точности распознавания как категориальных, так и непрерывных эмоций.

Некоторые исследователи применяют многозадачные подходы для решения нескольких задач одновременно [9]. Как правило, такие задачи являются коррелирующими между собой. В таком случае распознавание одной категории помогает повысить точность распознавания другой категории.

Корпус данных. Для экспериментальных исследований был использован многомодальный корпус данных RECOLA [10]. Данный корпус включал в себя видеозаписи диалогов между 2 участниками. Диалоги проводились с помощью системы телеконференций – Skype. Всех участников перед диалогом настраивали на какую-либо эмоцию (нейтральную, положительную, отрицательную) посредством показа эмоциональных видеороликов. Участники не были ограничены в длительности диалога. Однако, авторы корпуса каждый диалог обрезали до первых 5 минут, придерживаясь логики, что начало диалога более репрезентативно в проявлении невербальных сигналов собеседников. Всего в корпусе 23 видеозаписи (общим объемом 1,5 ч), в каждом видео присутствует только 1 диктор, всего 10 мужчин и 13 женщин разной национальности. Общение между дикторами происходило на французском языке.

Аннотация корпуса происходила по непрерывным меткам интенсивности и полярности эмоций (от -1 до 1) с интервалом в 4 секунды. Также видеозаписи были размечены по таким состояниям, как вовлеченность, согласие, доминирование, эффективность, взаимопонимание на всем интервале времени.

Для построения системы классификации полярности и интенсивности эмоций метки, предложенные авторами RECOLA, были сгруппированы в 3 класса: низкий (значения меньше -0,1, средний (значения от -0,1 до 0,1) и высокий (значения больше 0,1). Метки вовлеченности были также сгруппированы в 3 класса, участки пауз диктора помечались как низкая вовлеченность, значения 3-5 – средняя, и 6 – высокая.

Предложенный подход. В данной статье анализируется эффективность многомодального и многозадачного подходов к распознаванию вовлеченности, полярности и интенсивности эмоций участников виртуальной коммуникации. На рисунке 1 представлены baseline подходы к распознаванию вовлеченности, полярности и интенсивности эмоций отдельно для каждой используемой модальности: видео и аудио.

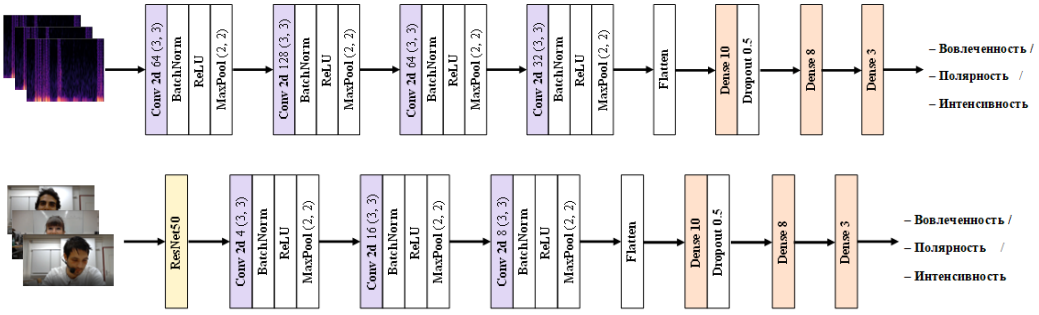


Рис. 1. Baseline подход анализа аудио (сверху) и видео (снизу) для распознавания вовлеченности, полярности и интенсивности эмоций по отдельности

Как видно из рисунка 1 в качестве акустических признаков использовались спектрограммы, в качестве визуальных – нейросетевые признаки ResNet50. Оба подхода содержали в себе сверточные блоки, которые в последствии подавались на полносвязные слои. В baseline подходе задачи распознавания вовлеченности, полярности и интенсивности эмоций решались отдельно, и для каждой задачи использовалась одинаковая архитектура, представленная на рисунке 1. Распознавание вовлеченности, полярности и интенсивности эмоций происходит на 3 класса для каждого: низкий, средний и высокий.

Как было сказано во введении такие человеческие состояния, как вовлеченность и эмоции являются коррелирующими понятиями. Исходя из этого предлагается гипотеза, заключающаяся в том, что многозадачный (англ. multitask) подход может более полно анализировать данные и лучше распознавать коррелирующие задачи. На рисунке 2 показаны многозадачные подходы анализа аудио и видео для задачи распознавания одновременно вовлеченности, интенсивности и полярности эмоций собеседников.

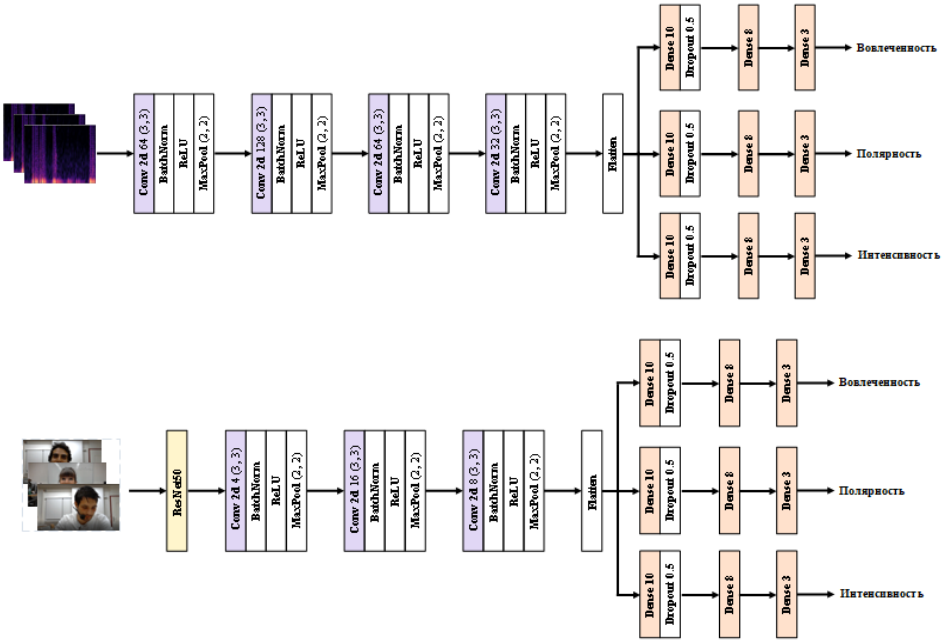


Рис. 2. Многозадачный подход для распознавания одновременно вовлеченности, полярности и интенсивности эмоций

Как видно из рисунка 2 архитектура моделей является схожей с baseline архитектурами, за исключением параллельных последних полносвязных слоев.

Исходя из анализа существующих исследований, многомодальный (англ. multimodal) подход позволяет построит более робастную систему классификации. Поэтому в этой статье также анализируется подход к отдельным задачам распознавания вовлеченности, полярности и интенсивности эмоций с использованием визуальной и акустической информации одновременно. Такой подход изображен на рисунке 3.

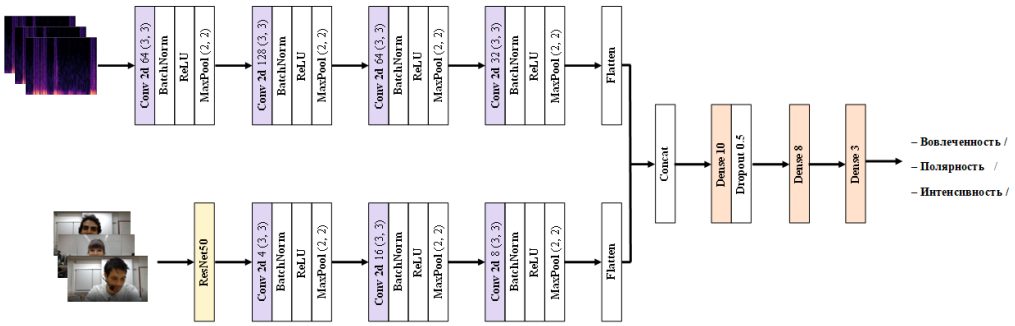


Рис. 3. Многомодальный подход для анализа вовлеченности, полярности и интенсивности эмоций по отдельности

Из рисунка 3 видно, что объединение модальностей происходит на уровне принятия решений с помощью полносвязных слоев.

Наконец, в статье предлагается многомодальный и многозначный подход к распознаванию вовлеченности и эмоций, представленный на рисунке 4.

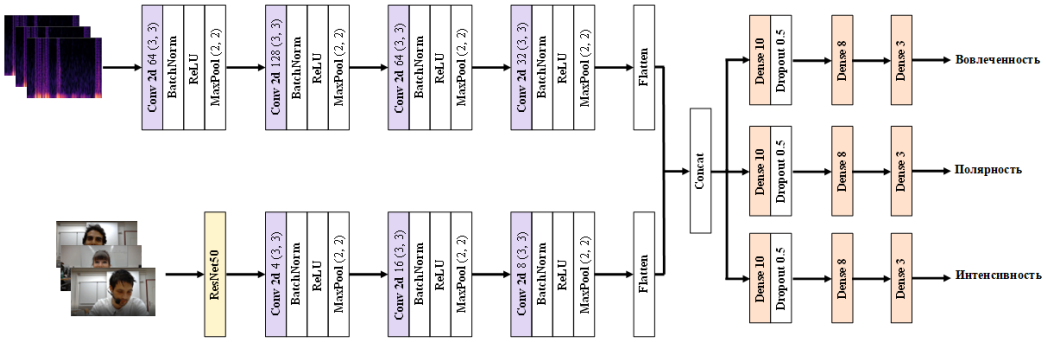


Рис. 4. Многомодальный и многозадачный подход для распознавания вовлеченности, полярности и интенсивности эмоций

Предложенный подход на рисунке 4 позволяет построить робастную и объективную систему анализа визуальной и акустической информации для распознавания вовлеченности, полярности и интенсивности эмоций собеседников виртуальной коммуникации.

Экспериментальные исследования. Все описанные в разделе выше подходы были апробированы на многомодальном корпусе RECOLA [10]. Для этого аудиоданные из корпуса были нарезаны по 1 секунде, из этого промежутка был извлечен 1 кадр из видеопоследовательности. Для обучения предложенных архитектур нейронных сетей использовались следующие параметры: скорость обучения – 0,001, размер пакета обучения – 6, количество эпох – 100 (используется прерывание обучение, если функция потерь не уменьшалась на протяжении 6 эпох). Результаты экспериментальных исследований показаны в таблице.

Таблица

Результаты экспериментальных исследований предложенных подходов к распознаванию вовлеченности, полярности и интенсивности эмоций, взвешенной полноты, %

Подход	Вовлеченность	Полярность	Интенсивность
Baseline видео (рис. 2 снизу)	29,14	8,14	40,33
Baseline аудио (рис. 2 сверху)	64,05	57,21	42,65
Многозадачный видео (рис. 3 снизу)	64,10	58,94	32,12
Многозадачный аудио (рис. 3 сверху)	69,29	59,02	24,38
Многомодальный (рис. 4)	64,05	57,21	42,65
Многомодальный и многозадачный (рис. 5)	64,05	59,02	24,31

Исходя из результатов, представленных в таблице, можно сделать вывод, что аудио модальность является более репрезентативной для анализа всех трех исследуемых задач: вовлеченности, полярности и интенсивности эмоций. Это связано с тем, что система анализа аудиоинформации лучше распознает средние и высокие классы, а они содержат большее количество экземпляров относительно низких классов. При этом видео модальность хорошо предсказывает низкие значения вовлеченности, полярности и интенсивности, поэтому показатель взвешенной полноты является достаточно низким при baseline подходе.

Также стоит отметить, что многозадачный подход позволяет улучшить точность распознавания вовлеченности и полярности эмоций на 34,96%, 50,8% и 5,24%, 1,81% для видео и аудио модальности, соответственно. Однако точность распознавания интенсивности эмоции при многозадачном подходе значительно снижается. Данные результаты экспериментальных исследований подтверждают, что полярность эмоций и вовлеченность – это зависимые друг от друга состояния человека во время коммуникации. При этом интенсивность эмоций не коррелирует с вовлеченностью также, как и с полярностью эмоций. Исходя из данных экспериментов можно сделать вывод, что такие состояния как вовлеченность и полярность эмоций необходимо анализировать одновременно.

Многомодальный подход не позволил повысить точность распознавания вовлеченности и эмоций. С помощью данного подхода получилась максимальная точность как при baseline подходе с использованием только акустической информации. Это может быть связано с особенностью данных в корпусе RECOLA. Также предложенный многомодальный и многозадачный подход не позволил превысить точность, полученную при многозадачном подходе. Это может быть также связано с использованием нескольких каналов информации одновременно.

Заключение. В данной статье были проведены экспериментальные исследования с многозадачными и многомодальными подходами к распознаванию вовлеченности, полярности и интенсивности эмоций собеседников виртуальной коммуникации. На данных корпуса RECOLA удалось достичь точности распознавания 3 классов вовлеченности, полярности и интенсивности эмоций по показателю взвешенной полноты = 69,29%, 59,02%, 42,62%, соответственно.

Литература

1. Двойникова А.А., Кагиров И.А., Карпов А.А. Аналитический обзор методов автоматического распознавания вовлеченности пользователя в виртуальную коммуникацию // Информационно-управляющие системы. 2022. №. 5 (120). С. 12-22.
2. Двойникова А.А., Маркитантов М.В., Рюмина Е.В., Уздяев М.Ю., Величко А.Н., Рюмин Д., Ляксо Е.Е., Карпов А.А. Анализ информационного и математического

обеспечения для распознавания аффективных состояний человека // Информатика и автоматизация. 2022. Т. 21. №6. С. 1097-1144.
3. Dresvyanskiy D., Minker W., Karpov A. Deep learning based engagement recognition in highly imbalanced data // Speech and Computer. 2021. С. 166-178.
4. Двойникова А.А. Распознавание вовлеченности собеседников с помощью анализа мел-спектрограмм // Сборник трудов XI Конгресса молодых ученых. 2022. Т. 2. С. 38-42.
5. Russell J. A. Culture and the categorization of emotions // Psychological bulletin. 1991. Vol. 110. No. 3. P. 426 – 450.
6. Ryumina E., Dresvyanskiy D., Karpov A. In search of a robust facial expressions recognition model: A large-scale visual cross-corpus study // Neurocomputing. 2022. Т. 514. С. 435-450.
7. Verkholyak O.V., Kaya H., Karpov A.A. Modeling short-term and long-term dependencies of the speech signal for paralinguistic emotion classification// SPIIRAS Proceedings. 2019. Vol. 18. No. 1(62). Pp. 30-56.
8. Dvoynikova A., Verkholyak O., Karpov A. Analytical review of methods for identifying emotions in text data // CEUR Workshop Proceedings. 2020. Vol. 2552. Pp. 8-21.
9. Velichko A. et al. Complex Paralinguistic Analysis of Speech: Predicting Gender, Emotions and Deception in a Hierarchical Framework // INTERSPEECH 2022. 2022. С. 4735-4739.
10. Ringeval F. et al. Introducing the RECOLA multimodal corpus of remote collaborative and affective interactions // 2013 10th IEEE international conference and workshops on automatic face and gesture recognition (FG). IEEE. 2013. С. 1-8.

УДК 004.934.8'1

ИССЛЕДОВАНИЕ ВЛИЯНИЯ УСЛОВИЙ ПРОВЕДЕНИЯ ЗВУКОЗАПИСЕЙ НА РЕЗУЛЬТАТ РАСПОЗНАВАНИЯ ДИКТОРОВ ПРИ ИСПОЛЬЗОВАНИИ АВТОМАТИЧЕСКОГО МЕТОДА СРАВНЕНИЯ

Инкин М.Е.[1] (магистрант)**, Столбов М.Б.**[1] (к.т.н., доцент)
Научный руководитель – к.т.н., доцент Столбов М.Б.[1]
1 – Университет ИТМО
e-mail: proppkrot@gmail.com; mbstolbov@itmo.ru

Аннотация
В работе рассматривается влияние условий проведения тестовых звукозаписей на результат идентификации дикторов при использовании автоматических систем распознавания дикторов (АРД). Приведены теоретические и экспериментальные исследования, подтверждена теоретическая модель оценки качества сигналов для АРД. Обосновано применение метода оценки отношения сигнал-реверберированный сигнал (ОСР) для оценки качества сигнала.
Ключевые слова
автоматическое распознавание диктора, биометрия, качество звукозаписи, отношение сигнал-шум, отношение сигнал-реверберация.

Системы автоматического распознавания дикторов (АРД) находят применение в разных областях жизни – коммерческой деятельности (банки, контакт-центры), органах безопасности, компьютерных системах безопасности и других.

В основу АРД положено сравнение двух фонограмм – образцовой, записанной для известного диктора, и тестовой, записанной для диктора, близость речи которого с образцовым диктором необходимо оценить. Образцовая фонограмма записывается обыкновенно в контролируемых, «идеальных» условиях в то время, как запись тестового образца может проводиться различными способами в различных условиях. При этом на надежность АРД влияет ряд факторов записи тестового образца: кодеки, тип устройства звукозаписи, вариативность речи дикторов, отношение сигнал-шум и время реверберации [1, 2].

Основными параметрами, связанными с условиями звукозаписи, являются отношение сигнал-шум (ОСШ) и время реверберации (ВР). Исследованию влияния этих факторов на работу систем АРД посвящен ряд работ [2, 3]. При этом ОСШ и ВР рассматриваются как подлежащие оценке характеристики фонограмм. Значения ОСШ и реверберационных искажений зависят от условий проведения звукозаписи – акустических характеристик помещения, времени реверберации, уровня шума и дистанции, с которой осуществляется звукозапись речи целевого диктора. Поэтому для операторов, осуществляющих запись тестовых фонограмм для системы АРД, целесообразно дать рекомендации по выбору условий звукозаписи.

Целью данной работы является исследование влияния условий проведения звукозаписей на результат идентификации дикторов при использовании автоматических методов идентификации дикторов. Экспериментальная часть исследование проводилась с использованием системы АРД производства ЦРТ [4].

Опишем теоретическую модель. Сигнал на микрофоне можно представить в виде суммы речевого сигнала (РС) s(i), реверберации r(i) и шума n(i):

$$x(i) = s(i) + r(i) + n(i). \quad (1)$$

Поскольку речевой сигнал, реверберацию и шум можно считать независимыми, интенсивность сигнала складывается из интенсивностей речевого сигнала Is, реверберации Ir и шума In:

$$Ix = Is + Ir + In. \quad (2)$$

Интенсивность речи в свободном пространстве зависит от расстояния между диктором и микрофоном:

$$Is = Is(d_0) \times (\frac{d_0}{dist})^2, \quad (3)$$

где $Is(d_0)$ – интенсивность речи на дистанции d_0 от диктора, dist – дистанция между диктором и микрофоном.

Рассмотрим модель пространственно-равномерного поля шума. Относительный уровень шумовых искажений РС характеризуется величиной ОСШ:

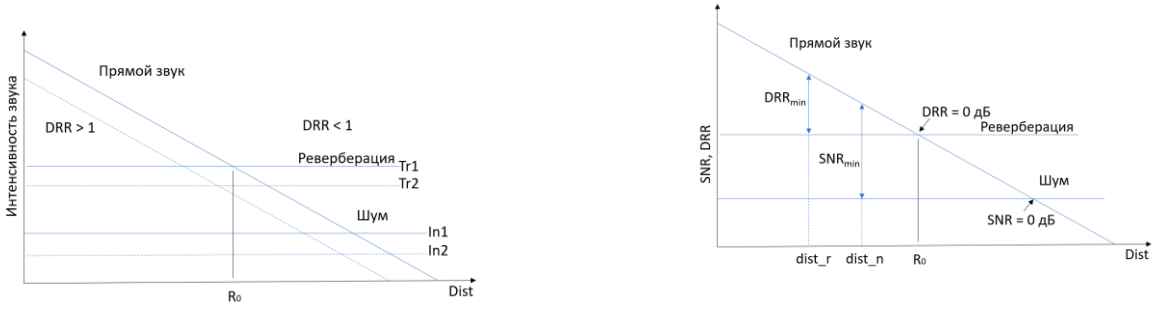
$$SNR(dist) = \frac{Is}{In} = SNRo \times \left(\frac{d_0}{dist}\right)^2, \quad (4)$$

Рассмотрим модель пространственно-равномерного поля реверберации.. Реверберацию можно рассматривать как нестационарный шум, порождаемый речью диктора. Интенсивность этого шума пропорциональна интенсивности речи диктора. Относительный уровень реверберационных искажений РС характеризуется отношением сигнал-реверберация (ОСР) [5]:

$$DRR(dist) = \frac{Is}{Ir} = \left(\frac{R_0}{dist}\right)^2, \quad (5)$$

где DRR (direct-to reverberation ratio) – ОСР, R_o – радиус реверберации, т.е. дистанция, на которой интенсивности прямого звука и реверберации равны. [6].

На рисунке 1(а) показаны графики зависимости от дистанции интенсивностей РС, реверберации Ir1 и шума In1. На рисунке 1(б) показаны графики зависимости от дистанции ОСШ и ОСР, а максимальные дистанции эффективного распознавания дикторов по ОСР и ОСШ обозначены как dist_r dist_n.



(а) (б)

Рис. 1. (а) Зависимость от дистанции интенсивностей РС, реверберации и шума, (б) зависимость от дистанции ОСШ и ОСР (б). Логарифмический масштаб

Экспериментальная оценка зависимости надежности решения конкретной системы АРД от SNR, DRR позволяет для заданных акустических условий определить допустимую дистанцию звукозаписи, при которой требуемая надежность АРД будет обеспечена.

Как следует из уравнений, для оценки ОСШ и ОСР необходимо знание значений величин $Is(d_o)$, In, R_o. Оценки этих значений могут быть получены из литературы либо экспериментально.

Рассмотрим теоретические оценки. Примем за величину интенсивности нормальной речи на расстоянии 1 м значение 68 дБ [7]. Величины интенсивности шума в разных условиях приведены в работе [8]. Значения ОСШ для разных дистанций диктор-микрофон приведены в таблице 1.

Таблица 1

Значение ОСШ для разных дистанций диктор-микрофон, дБ

\

Интенсивность шума	Дистанция диктор-микрофон					
	25 см	50 см	1 м	2 м	4 м	8 м
Тихая квартира, 35 дБ	46	39	33	27	21	15
Тихое учреждение, 40 дБ	40	34	28	22	16	10
Шумное учреждение, 65 дБ	15	9	3	-3	-9	-15

Величина ОСР зависит от радиуса реверберации. Известно, что в комнате средних размеров $R_o \approx 2$ м, в комнате больших размеров (класс) $R_o \approx 4$ м.

Величина радиуса реверберации зависит от времени реверберации и размеров помещения. Для оценки радиуса реверберации предложено несколько формул. Воспользуемся эмпирической формулой [6]:

$$R_0 = 0.06\left(\frac{V[m^3]}{Tr[c]}\right)^{\frac{1}{2}}, \quad (6)$$

где V – объем помещения [м³], Tr – время реверберации [с]. Например, для помещения V = 120 м³, Tr = 0.5 с получим оценку $R_o \approx 1$ м.

Рассчитанные значения ОСР для разных дистанций диктор-микрофон приведены в таблице 2.

Таблица 2

Значения ОСР для разных дистанций диктор-микрофон

Радиус реверберации	Дистанция диктор-микрофон					
	25 см	50 см	1 м	2 м	4 м	8 м
Свободное пространство	82	74	68	62	56	50
1 м	+12	+6	0	-6	-12	-18
2 м	+18	+12	+6	0	-6	-12
4 м	+24	+18	+12	+6	0	-6

Предельными значениями для рассмотренной системы являются: ОСШ>15 дБ, Tr <250 мс, оценка ОСР не заявлена. Нами была проведена экспериментальная оценка зависимости надежности от ОСШ, ОСР. Целью экспериментов было исследование зависимости меры сходства речи дикторов, записанных в образцовых условиях и в помещениях в условиях шума и реверберации. Исследования проводились в помещении 3x10x10 = 300 м^3, Tr = 890 мс. На рисунке 2 представлены графики зависимости меры сходства от расстояния диктор-микрофон для разных значений ОСШ.

Из графиков следует, что максимально допустимое расстояние для выполнения распознавания диктора с помощью выбранной АРД при малом уровне шума (ОСШ 16 дБ) равно 4 метрам. Допустимое расстояние сокращается при уменьшении ОСШ.

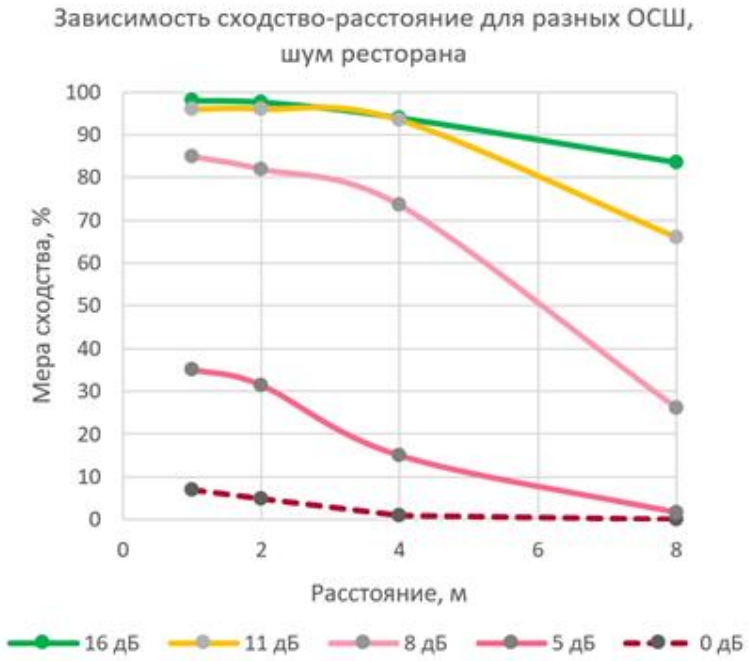


Рис. 2. Зависимость меры сходства от расстояния при разных отношениях сигнал-шум

Были также проведены эксперименты по оценке радиуса реверберации. В основу было положено разделение сигнала удаленного микрофона на прямую (речевую) и диффузную (реверберационную) компоненты с помощью двухканального адаптивного компенсатора помех. Исход экспериментов положительный: на стационарном шуме полученная оценка радиуса реверберации совпала с теоретической, однако на РС величина оценки радиуса реверберации оказалась больше теоретической. Возможно, это связано с нестационарностью речевых сигналов. Данный вопрос требует дальнейшего исследования.

В нашей работе мы обосновали выбор в качестве фактора надежности распознавания отношения сигнал-реверберация. Разработана экспериментальная методика оценки радиуса реверберации в помещении. Экспериментально исследована зависимости меры сходства в системе АРД от ОСР и ОСШ для различных типов шумов.

Предложенная модель позволяет сформулировать рекомендации о типе помещений, дистанции диктор-микрофон при записи тестовых образцов, что должно способствовать обеспечению надежной идентификации.

Литература

1. Аусев Е.В. Исследование точности верификации фразы от кодирования // IX Всероссийский конгресс молодых ученых Сборник тезисов докладов конгресса молодых ученых. Электронное издание. СПб: Университет ИТМО. 2020.
2. Аусев Е.В. Исследование систем верификации диктора в условиях использования удаленного микрофона. Диссерт. Магистрская. СПб. 2020. 75 с.
3. Волкова М.В. Автоматическая оценка качества речевых сигналов для систем голосовой биометрии и антиспуфинга // Диссертация на соискание учёной степени кандидата технических наук 2022. СПб. 2020. 224 с.
4. Novoselov S. et al. STC speaker recognition systems for the VOiCES from a distance challenge //arXiv preprint arXiv:1904.06093. 2019. (дата обращения: 28.04.2023).
5. Чебанов А.Д. Приближенная оценка времени реверберации для залов различного функционального назначения: учебно-методические указания // М.: МАРХИ. 2012. 36 с.

6. Ананьев А.Б. Математика для звукорежиссера. Выпуск 4. Учебное пособие. Киев: 2015. 55 с.
7. Алдошина И.А., Вологодин Э.И., Ефимов А.П. и др. Электроакустика и звуковое вещание: учебное пособие для вузов // Под редакцией Ю.А. Ковалгина. М.: Горячая линия-Телеком. Радио и связь. 2007. 872 с.
8. Тэйлор Р. Шум // М.: Мир. 1978.

УДК 004.05

РАЗРАБОТКА АВТОМАТИЗИРОВАННОЙ СИСТЕМЫ АУДИТА ПОКАЗАТЕЛЕЙ ЭФФЕКТИВНОСТИ БИЗНЕС-ПРОЦЕССОВ

Кузнецова Е.Г.[1] (магистрант)
Научный руководитель – к.э.н., доцент Карасева Е.И.[2]
1 – Университет ИТМО
e-mail: egkuznetcova@itmo.ru

Аннотация
В настоящее время всё больше предприятий задумываются над вопросами эффективности функционирования своих бизнес-процессов. Однако получить полную картину эффективности возможно лишь путём серьёзного и тщательного анализа большого количества показателей, которые характеризуют именно конкретное предприятие. Для предприятий становится актуальным внедрять современные информационные системы, которые способны сформировать подробную статистику, выделив нужные показатели для дальнейшего удобства анализа.
Ключевые слова
Веб-приложение, показатели эффективности, бизнес-процессы, моделирование, автоматизация.

В проектировании, разработке, науке очень часто широко применяются модели бизнес-процессов. Они являются мощным инструментом для представления сложных связей или структур для их лучшего понимания и визуализации. При отсутствии таких моделей необходимо было бы осмысливать совсем неструктурированные данные. Хорошие модели способны подчеркнуть характерные особенности структур [1].

Существует достаточно много методик моделирования бизнес-процессов. В рамках данной работы рассмотрена популярная и используемая нотация IDEF0 и визуальный алгоритмический язык ДРАКОН, как отечественный аналог. Методику IDEF0 можно считать следующим этапом развития хорошо известного графического языка описания функциональных систем SADT (Structured Analysis and Design Technique). Целью методики является построение функциональной схемы исследуемой системы, описывающей все необходимые процессы с точностью, достаточной для однозначного моделирования деятельности системы. В основе методики лежат четыре основных понятия: функциональный блок, интерфейсная дуга, декомпозиция, глоссарий.

Для данной работы были выбраны три основных бизнес-процесса (обработка заявки на обучение, заключение договора на обучение и проведение курса обучения) центра развития дополнительного образования, главной целью которого является предоставление образовательных услуг и подготовка будущих абитуриентов. На рисунке 1 представлена процедура обработки заявки на обучение в методике IDEF0.

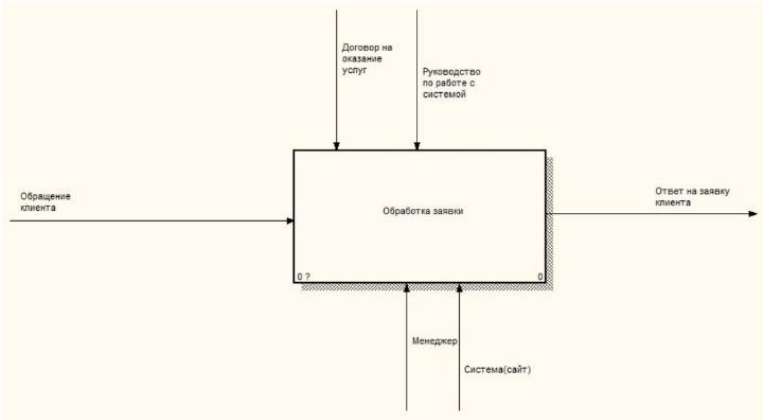


Рис. 1. Контекстная диаграмма бизнес-процесса «Процедура обработки заявки»

ДРАКОН: Дружелюбный Русский Алгоритмический язык, Который Обеспечивает Наглядность (DRAKON), был разработан в рамках космической программы «Буран». Разработка языка велась с 1986 г. при участии Федерального космического агентства (Научно-производственный центр автоматики и приборостроения им. акад. Н. А. Пилюгина, Москва) и Российской академии наук (Институт прикладной математики им. М. В. Келдыша). Главной задачей разработчиков стало создание единого универсального языка программирования и моделирования, способный своей доступностью и мощностью заменить специализированные языки, которые используются в организациях для моделирования и разработки. Язык построен путем формализации, эргономизации и неклассической структуризации блок-схем алгоритмов, описанных в ГОСТ 19.701—90 и ISO 5807—85 [2]. На рисунке 2 представлен бизнес-процесс обработки заявки на обучение в ДРАКОН.

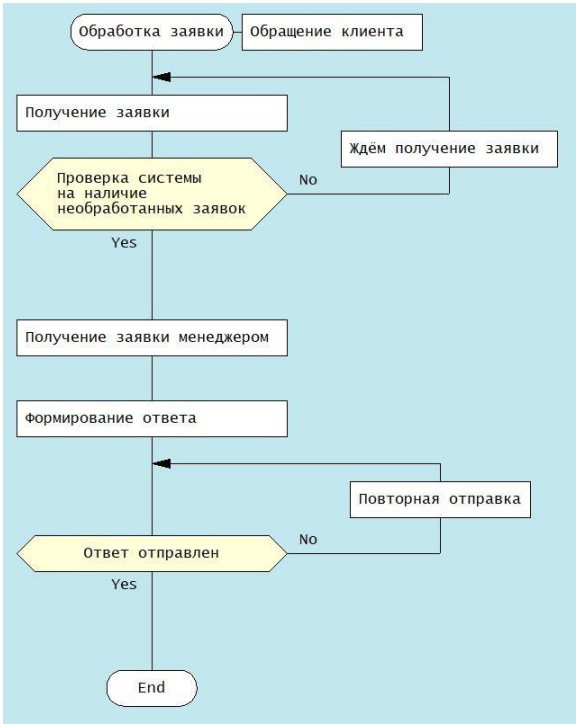


Рис. 2. ДРАКОН-схема бизнес-процесса «Процедура обработки заявки»

Сравнение методик показало, что наиболее неудобной является методология IDEF0, поскольку в ней отсутствует ветвление. Использование ДРАКОН-схем позволяет избавиться от этого недостатка. В ДРАКОН-схемах выделяются подпроцессы, что позволяет добиться максимальной наглядности графического описания бизнес-процесса.

Далее было необходимо выделить показатели эффективности для центра развития дополнительного образования для алгоритмического языка ДРАКОН. Всего было определено пять показателей:

– прибыльность – отражает эффективность деятельности организации, ее финансовые результаты и показывает, сколько денежных средств из вырученных от реализации является прибылью;

– длительность – важный временной показатель, определяется как отношение затраченного в день времени обработки 1 заявки к общему рабочему времени;

– качества – показатель, определяющий насколько успешно реализовываются процессы;

– регулируемости – определяет, как соотносятся количество регламентной документации в бизнес-процессах и экземпляры (действия) в бизнес-процессах;

– контролируемости – позволяет охарактеризовать эффективность управления бизнес-процессами, принадлежащими и управляемыми собственниками процессов.

Для решения вопроса автоматизации анализа показателей эффективности бизнес-процессов было принято создать программную систему. Выбор типа приложения был осуществлен в пользу веб-приложения, доступного в сети Интернет, которое предоставляет удобство в использовании без привязки к конкретной операционной системе. Основной возможностью пользователя приложения является внесение данных о бизнес-процессах, что доступно с компьютера с помощью возможностей веб-браузера.

Отличительной особенностью приложения служит функция анализа показателей эффективности бизнес-процессов. После загрузки данных о бизнес-процессах на серверную часть происходит сравнение и подсчёт нужных значений. В результате пользователь получает отчет в виде получившихся значений и вывод о их эффективности. Отчет сохраняется в хранилище, после чего, пользователь, проанализировав получившиеся значения, может их изменить и повторить проверку снова.

Таким образом, достигается простота анализа и автоматизация рутинной работы для пользователей.

Процесс проектирования приложения включил в себя:

- определение функциональных возможностей разрабатываемого приложения;
- формулирование требований к приложению;
- создание архитектуры приложения.

Разрабатываемое приложение должно предоставлять следующие функциональные возможности:

- регистрация пользователей в системе;
- авторизация пользователей с помощью пары логин-пароль;
- загрузка данных о бизнес-процессах, хранение и обработка;
- определение показателей эффективности бизнес-процессов по двум методам (IDEF0 и ДРАКОН);
- объединение показателей эффективности в единый отчёт по конкретному бизнес-процессу;
- хранение отчётов проанализированных бизнес-процессов.

Для пользователей необходимо предусмотреть разграничение прав доступа к информации, а также любым действиям с ней - её добавление, редактирование и удаление. При работе с приложением допустимы роли неавторизованного и авторизованного пользователя, а также технического администратора.

Далее необходимо было выбрать архитектуру приложения. Часто используемой программистами архитектурой является - трехуровневая модель, которая соответственно состоит из трёх составных частей и представлена на рисунке 3.

Проектирование архитектуры приложения является важным процессом, который включает в себя выбор решений, оказывающих в дальнейшем большое влияние на ряд параметров, например, производительность, качество, удобство обслуживания, масштабируемость и др. Именно при проектировании определяются ключевые сценарии использования приложения, требования по атрибутам качества, фундаментальные части архитектуры и формируются предложения по применению компонентов на всех слоях будущего приложения [3].

При таком чётком разделении трёх слоёв архитектуры, разрабатываемая система с точки зрения программного кода, станет читабельнее, к тому же такая архитектура обеспечивает, при дальнейшей необходимости лёгкую масштабируемость, по сравнению с двухуровневой архитектурой клиент-сервер.

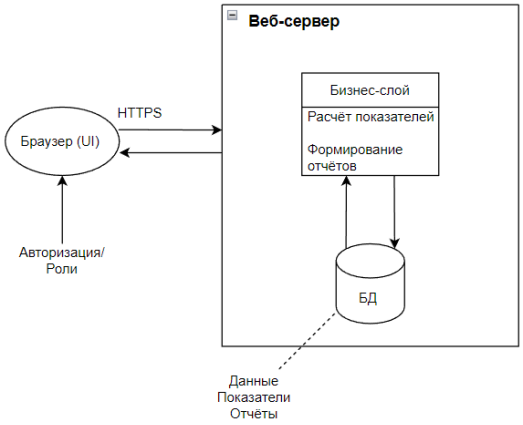


Рис. 3. Архитектура приложения

Для реализации нужной системы аудита показателей эффективности бизнес-процессов были выбраны следующие технологии:

1. Язык программирования - Python.
2. Фреймворк - Django.
3. База данных - SQLite.

Python – язык высокого уровня программирования общего назначения, который ориентирован на повышение читаемости кода и производительности разработчика. Также у этого языка программирования довольна большая область применения – он применяется как во внутренней разработке, как и в приложениях на основе искусственного интеллекта.

Популярность Python обеспечивает большое количество фреймворков (Pyramid, Pylons, TurboGears, Flask, CherryPy и Django), упрощающие процесс написания кода, сокращение времени на разработку, а также высокая производительность при обработке большого количества данных.

Для данной разрабатываемой системы было принято решение использовать фреймворк Django. Фреймворк использует шаблон проектирования model – view – controller. Также в составе Django присутствует собственный веб-сервер для разработки, который автоматически способен определять изменения в файлах исходного кода проекта и перезапускаться, это позволяет ускорять процесс разработки на Python. Такой фреймворк был создан для того, чтобы переходить от прототипов к готовым сервисам быстрее и позволяет сосредоточиться на бизнес-логике.

В данной работе было рассмотрено клиент-серверное веб-приложение, предоставляющее пользователю определённые в процессе проектирования функциональные возможности.

После размещения исходного кода приложения на хостинге и регистрации доменного имени веб-приложение будет доступно сотрудникам центра дополнительного образования в любом браузере в сети Интернет.

После проведённой работы можно сделать вывод, что при разработке приложения необходимо учитывать, все требования и функциональные возможности, предъявляемые к системе, т. к. это будет напрямую влиять на то, как будет спроектировано приложение, а в последствие и на то, как приложение будет работать.

Использование данного приложения будет способствовать автоматизации аудита бизнес-процессов центра развития дополнительного образования, а также определению эффективности показателей бизнес-процессов, что позволит сотрудникам и руководству центра оценить и проанализировать свои подходы к бизнес-процессам, понять, что необходимо изменить в работе, или какие моменты необходимо усилить, чтобы сделать работу сотрудников более результативной, а саму организацию конкурентноспособной на рынке.

Литература

1. Игнатьев А.В. Методы и средства проектирования информационных систем и технологий: учебное пособие / А.В. Игнатьев; М-во образования и науки Рос. Федерации, Волгогр. гос. архит.-строит. ун-т. Волгоград: ВолгГАСУ. 2014. 56 с.
2. Пароджанов В.Д. Алгоритмы и жизнеритмы на языке ДРАКОН. Разработка алгоритмов. Безошибочные алгоритмы. М. 2020. 376 с.
3. Руководство Microsoft по проектированию архитектуры приложений (2-е издание) //©2009 Корпорация Майкрософт. С. 28-152.

УДК 004.58

РАСПОЗНАВАНИЕ ЖЕСТОВ В УМНОМ ДОМЕ

Кумашев Э.[1], Аминов Н.С.[1]
Научный руководитель – к.т.н., доцент А.А. Капитонов[1]
Университет ИТМО
e-mail: kumashev22@gmail.com

Аннотация
В настоящее время все больше исследований направлено на решение задач компьютерного зрения и искусственного интеллекта. Наиболее популярными из них являются задачи, направленные на взаимодействие человека с машинным интерфейсом. Управление электроникой через распознавание жестов не является новой задачей, ранее уже были описаны несколько вариантов реализаций, представленные различными носимыми и не носимыми устройствами, с использованием компьютерного зрения и без, с моделями формы и ручного скелета. В целях достижения задачи по управлению устройствами в системе «умного дома» через распознавание жестов, в данной работе было решено сконцентрироваться на распознавании кистей рук с использованием модели ручного скелета.
Ключевые слова
Распознавание жестов, умный дом, компьютерное зрение, машинное обучение, построение взаимодействия.

Распознавание жестов кистей рук с использованием модели ручного скелета

С ростом цифровых технологий все больше исследований сосредоточено на решении задач, связанных с компьютерным зрением и искусственным интеллектом. Среди таких задач являются те, которые направлены на создание интерфейсов для взаимодействия между человеком и машиной, например: определение позы человека, распознавание лица, жестов кистей рук и, даже жестов лица [1-3].

Для широкого распространения данного функционала машины должны быстро и точно их определять. Это возможно сделать, так как каждая из вышеупомянутых задач имеет математическую интерпретацию вычислительного устройства [4].

Жесты рук человека являются наиболее важным средством для невербального общения между людьми. Они варьируются от простых манипулятивных жестов, используемых для указания и перемещения объектов, до более сложных коммуникативных, выражающих наши чувства и позволяющих нам общаться с другими.

Модель ручного скелета — это один из подходов к решению проблемы распознаванию жестов ладони. Он допускает то, что жесты на ладони могут быть определены как набор прикосновений, пальцев и запястья, которые могут быть представлены в виде скелета рук. Модель ручного скелета строит 3D-модель скелета рук и использует ее для сохранения жестов. Для построения модели ручного скелета используются технологии компьютерного зрения и машинного обучения. С помощью камеры происходит захват изображения рук. Затем на этом изображении производится обнаружение и отслеживание кисти и суставов пальцев, чтобы получить координаты точек скелета руки. После получения координат скелета руки происходит процесс изъятия жестов. Для этого используются алгоритмы машинного обучения, которые учитывают большое количество данных. Модель ручного скелета имеет ряд преимуществ перед другими подходами к распознаванию жестов ладони.

Во-первых, она позволяет распознавать жесты без использования маркеров или других специальных устройств на руке. Во-вторых, модель ручного скелета достаточно устойчива к изменениям освещения, фону и другим условиям.

Было решено использовать модель ручного скелета, как подхода для моделирования руки. Данный способ является точнее, чем другие подходы и является менее требовательным к условиям окружения.

Существует 4 обобщенных стадий использования модели ручного скелета в задаче распознавания жестов:

1. Сбор данных (Создание своего или использование существующего датасета жестов).
2. Разметка датасета.
3. Обучение датасета.
4. Считывание жестов режиме реального времени.

По причине стабильности в следствии большого набора предобученных данных жестов руки (модель была обучена на порядка 30 тысячах изображений из реального мира, а также на нескольких визуализированных синтетических моделях рук, наложенных на различные фоны), за основу подхода модели ручного скелета было решено использовать модуль Hands библиотеки MediaPipe, представленный в июне 2019 года. MediaPipe Hands [5] – это высокоточное решение для отслеживания рук и пальцев, использующее машинное обучение (ML), для вывода 21 трехмерного ориентира ладони (рис. 1).

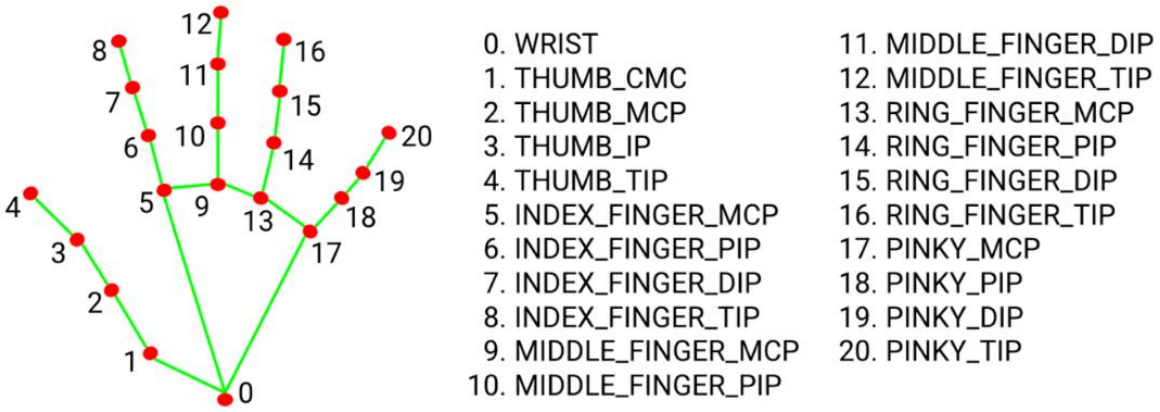


Рис. 1. 21 ориентир ладони

В ходе поиска релевантных решений было найдено три варианта:

1. Depth AI (под авторством: geaxgx, szabi-luxonis szabi-luxonis, michaelhwn)
2. American Sign Language (ASL) Recognition using Hand Landmarks (под авторством: michaelhwn, yelu728).
3. hand-gesture-recognition-using-mediapipe (под авторством: Kazuhito00, он же KazuhitoTakahashi).

Depth AI и American Sign Language (ASL) Recognition using Hand Landmarks.

Несколько фактов: оба этих решения имеют высокую точность распознавания (порядка 90%), оба были разработаны выходцами из одной компании разработчиков, однако для использования данных решений требуется использования камер OpenCV AI Kit: OAK—D [6-8].

hand-gesture-recognition-using-mediapipe (рис. 2).

Особенности данного решения:

1. Умная система конвертации изображений в ориентиры (landmarks).
2. Простота подхода.
3. Не требует специальных камер для реализации.

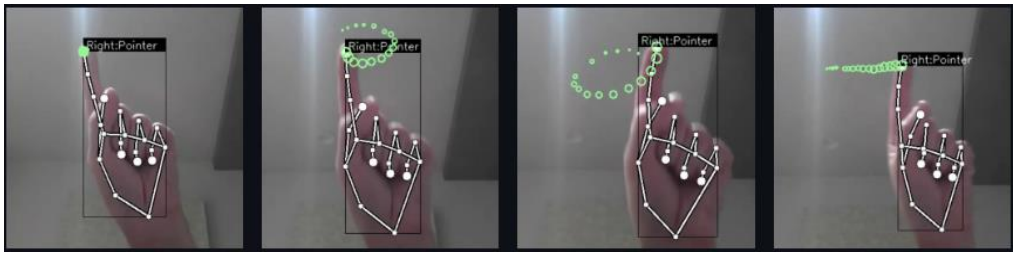


Рис. 2. Механизм работы универсального подхода [11]

Системы умного дома с открытым исходным кодом

Системы домашней автоматизации с открытым исходным кодом представляют собой доступную и открытую альтернативу коммерческим проектам, периодически обеспечивающим поддержку устройств и услуг, неподдерживаемых коммерческими альтернативами. Несмотря на то, что основной функционал умного дома поддерживается всеми существующими решениями, каждая отдельно взятая система обладает собственным набором функций и ограничений, что усложняет выбор конкретного решения.

OpenHab был выбран для проекта так как он имеет множество пользовательских интерфейсов и REST API [9]. Он состоит из 5 элементов: Вещи, Каналы, Привязки, Элементы и Ссылки. Вещи добавляются в систему и имеют функциональность, которая раскрывается через Каналы [10]. Привязки обеспечивают доступ к Вещам через систему. Элементы предоставляют функциональные возможности, а Ссылки соединяют Каналы и Элементы (рис. 3).

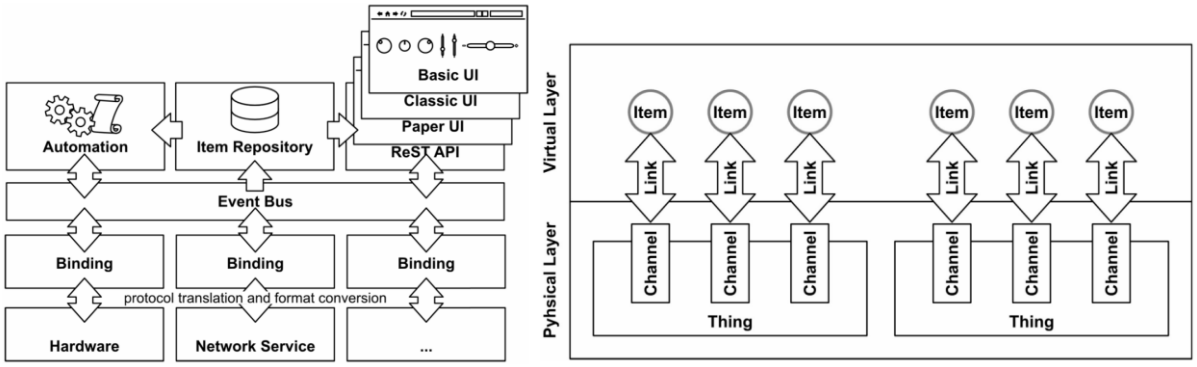


Рис. 3. Концептуальная архитектура системы openHAB

Разработка системы

Как было упомянуто выше, за основу системы распознавания жестов были взяты материалы KazuhitoTakahashi [11]. План разработки будет состоять из следующих этапов:

1. Сбор данных.
2. Обучение данных.
3. Распознавание жестов.
4. «Умный дом» и подключение устройств.
5. Разработка взаимодействия умного дома и распознавания жестов.

Сбор данных

Получение ориентиров.

Исходя из того, что модуль Hands MediaPipe умеет распознавать 21 ориентир ладони (рис. 1), необходимо более глубже описать процесс рисования ориентиров. Для

этого на основании исследований KazuhitoTakahashi [11] прямо в коде были определены ориентиры по точкам.

Далее необходимо написать функцию подсчета ориентиров исходя из размера окна программы.

Следующим шагом является преобразование полученных ориентиров в относительные координаты, с последующим преобразованием в одномерный список и нормализацией.

Добавление жестов.

Данный пункт является основным в пункте сбора данных для последующей классификации жестов. Для последующей обработки и хранения ориентиров ладони для жестов был создан файл keypoint.csv. Для добавления жестов необходимо переключиться в режим “Logging Key Point”, с возможностью добавить 10 жестов (Нажимая кнопку номера жестов 0–9 для записи ориентиров жестов в файл keypoint.csv [11].

Также было решено добавить количество кадров в секунду (FPS), модуль для подсчета (FPS).

На картинке ниже можно посмотреть на интерфейс полученной программы (рис. 4).

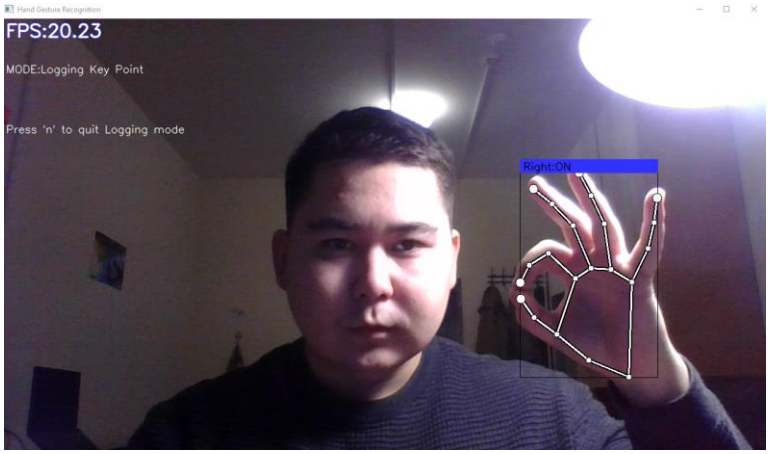


Рис. 4. Интерфейс добавления ключевых точек (жестов)

Следующий этап – это сбор данных, поворачивая жесты с разных сторон чтобы сделать собранные данные наиболее достоверными. Как уже было сказано ранее, записанные ориентиры жестов будут храниться в файле keypoint.csv. Для каждого жеста было собрано порядка 1500 ориентиров.

Обучение данных

Для обучения данных был создан файл в keypoint_classification.ipynb [11] в Jupyter Notebook. Обучение данных будет состоять из нескольких этапов:

1. Подготовка к обучению.
2. Загрузка обучающих данных.
3. Построение модели.
4. Обучение модели.
5. Построение матрицы путаницы.
6. Преобразование данных в модель Tensorflow-Lite.

1. Подготовка к обучению.

Вначале импортируем модули csv для работы с собранными ориентирами, а также модули numpy, tensorflow для работы с массивами и sklearn train_test_split для разделения набора данных на тестовую и обучающую выборки, а также определим значение RANDOM_SEED, цифра 42 была взята как сумма ориентиров двух ладоней.

Следом идет этап указания путей и параметра NUM_CLASSES, что говорит нам о том, сколько жестов у нас есть [11].

2. Загрузка обучающих данных.

Этап загрузки обучающих данных подразделяется на загрузку ориентиров, в данном случае для двух ладоней, затем загрузку информации к какому классу принадлежат ориентиры из файла keypoint.csv, и наконец разделение выборки на тестовую и обучающую, с разделением 25 и 75% соответственно [12].

3. Построение модели.

Для построения модели была использована модель нейронных сетей Sequential, которая позволяет строить модели таким образом, чтобы слои шли друг за другом [13].

Перед последующим обучением необходимо скомпилировать модель, при компиляции модели указываются параметры обучения [14].

Также необходимо отметить, что в качестве оптимизатора был использован оптимизатор “Adam”, так как Adam – один из самых эффективных алгоритмов оптимизации в обучении нейронных сетей. Он сочетает в себе идеи RMSProp и оптимизатора импульса. Вместо того чтобы адаптировать скорость обучения параметров на основе среднего первого момента (среднего значения), как в RMSProp, Adam также использует среднее значение вторых моментов градиентов. В частности, алгоритм вычисляет экспоненциальное скользящее среднее градиента и квадратичный градиент, а параметры beta1 и beta2 управляют скоростью затухания этих скользящих средних [15]. Он достаточно часто применяется при обучении нейронных сетей и хорошо себя зарекомендовал для поиска минимума функций потерь [16], что как раз и необходимо для достижения максимальной точности распознавания, код данной процедуры представлен листингом 1.

Листинг 1. Компиляция модели.

```
model.compile(
     optimizer='adam',
     loss='sparse_categorical_crossentropy',
     metrics=['accuracy']
 )
```

Функция ошибки в данной модели указана параметром loss в которой использована категориально-перекрестная энтропия, данная функция ошибки применима в задачах классификации, когда классов больше двух. И последний параметр, параметр качества metrics, то есть доля правильных ответов.

4. Обучение модели.

В основе обучения данных был взят метод «Обучения с учителем», поэтому значения X_train передаются как обучающая выборка, а y_train как ответы. Для тестирования было использовано обучение в течение 1000 эпох, код данной процедуры представлен листингом 2.

Листинг 2. Код для обучения модели.

```
model.fit(
    X_train,
    y_train,
    epochs=1000,
    batch_size=128,
    validation_data=(X_test, y_test),
    callbacks=[cp_callback, es_callback]
)
```

Как можем видеть из рисунка 5, вначале обучения точность составляла ~25%.

```
Epoch 1/1000
28/45 [=================>............] - ETA: 0s - loss: 1.6621 - accuracy: 0.2575
Epoch 1: saving model to model/keypoint_classifier\keypoint_classifier.hdf5
45/45 [==============================] - 1s 16ms/step - loss: 1.6184 - accuracy: 0.2828 - val_loss: 1.4699 - val_accuracy: 0.
5516
Epoch 2/1000
27/45 [=================>............] - ETA: 0s - loss: 1.4668 - accuracy: 0.3692
Epoch 2: saving model to model/keypoint_classifier\keypoint_classifier.hdf5
45/45 [==============================] - 0s 4ms/step - loss: 1.4383 - accuracy: 0.3875 - val_loss: 1.2779 - val_accuracy: 0.6
167
```

Рис. 5. Старт обучения модели

Но до 1000-й эпохи доходить не пришлось, уже на 152 эпохе была достигнута точность в свыше 95% (рис. 6), при неоднократных повторных запусках обучения, конечные значения, а также количество эпох варьировались, но все результаты оставались в округе ~95%.

```
Epoch 151/1000
28/45 [=================>............] - ETA: 0s - loss: 0.3912 - accuracy: 0.8585
Epoch 151: saving model to model/keypoint_classifier\keypoint_classifier.hdf5
45/45 [==============================] - 0s 4ms/step - loss: 0.3878 - accuracy: 0.8586 - val_loss: 0.1979 - val_accuracy: 0.9
474
Epoch 152/1000
26/45 [================>.............] - ETA: 0s - loss: 0.3843 - accuracy: 0.8630
Epoch 152: saving model to model/keypoint_classifier\keypoint_classifier.hdf5
45/45 [==============================] - 1s 12ms/step - loss: 0.3933 - accuracy: 0.8610 - val_loss: 0.1984 - val_accuracy: 0.
9510
Epoch 152: early stopping

<keras.callbacks.History at 0x19818a7fd00>
```

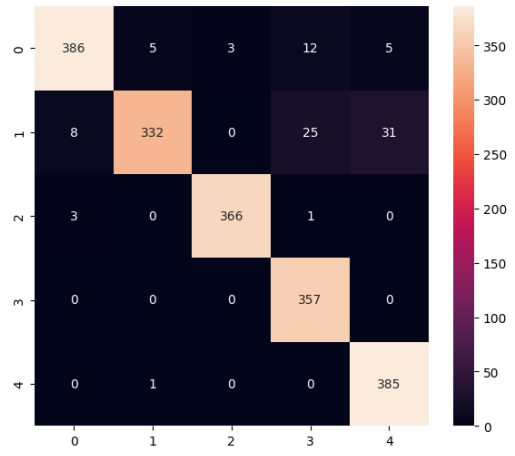
Рис. 6. Финиш обучения модели

```
1/1 [==============================] - 0s 91ms/step
[4.6154761e-01 5.0788212e-01 2.1538176e-03 1.2545590e-04 2.8290955e-02]
1
```

Рис. 7. Результаты проверки предсказаний модели Keras
по классам по очереди, для полученной модели

5. Построение матрицы путаницы.

Для проверки точности предсказаний, а также измерения производительности классификации была построена матрица путаницы с отчетом, результаты вывода на рисунке 8.




```
Classification Report
              precision    recall  f1-score   support

           0       0.97      0.94      0.96       411
           1       0.98      0.84      0.90       396
           2       0.99      0.99      0.99       370
           3       0.90      1.00      0.95       357
           4       0.91      1.00      0.95       386

    accuracy                           0.95      1920
   macro avg       0.95      0.95      0.95      1920
weighted avg       0.95      0.95      0.95      1920
```

Рис. 8. Матрица путаницы и отчет о классификации

Так как тестирование ведется на маломощном устройстве, одновременно в будущем планируется установка данного решения на еще менее производительное устройство типа микрокомпьютера, было решено использовать менее требовательный Tensorflow-Lite, метод, который был описан в проекте [11].

```
[4.6154782e-01 5.0788188e-01 2.1538185e-03 1.2545596e-04 2.8290953e-02]
1
```

Рис. 9. Результаты проверки предсказаний модели Tensorflow-Lite по классам по очереди

Исходя из результатов проверки на рис. 9 и в сравнении с рис. 7, можно сделать вывод, что преобразование модели из Keras в Tensorflow-Lite прошло успешно и без потерь.

Распознавание жестов

В ходе обучения из предыдущего раздела, кроме двух жестов пред обученных KazuhitoTakahashi [11] «Open», «Close», было обучено три жеста, это жесты «ОК», «ON», и «OFF». С результатами можно ознакомиться на рис. 10.

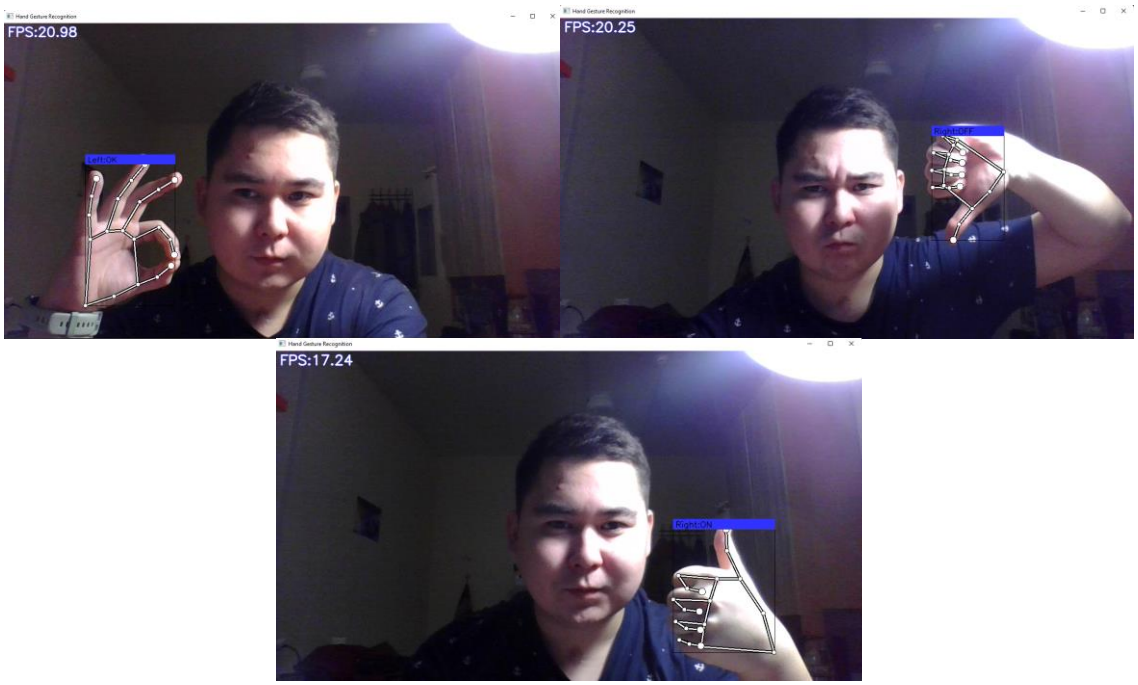


Рис. 10. Жесты

«Умный дом» и подключение устройств

Следующим шагом после установки системы openHAB на виртуальную машину, являлось подключение устройств. Взаимодействие осуществляется посредством протокола MQTT. Для последующего исследования было выбрано устройства типа WiFi Реле, производства компании Sonoff модели Basic R2. Такой выбор обусловлен дешевизной, а также простотой устройства данного типа. Однако в следствие отсутствия поддержи openHAB прошивки установленной для устройства компанией-производителем, был осуществлен поиск альтернативных прошивок, в ходе поиска была найдена прошивка Tasmota [17] из-за удобного интерфейса настройки и стабильности прошивки.

Дальнейшая перепрошивка, а также его настройка и подключение реле к системе, осуществлялась с использованием файла прошивки Tasmota, скачанной с сайта

производителя, приложение NodeMCU-PyFlasher версии 5.0, USB-TTL кабеля и на основании материалов Jordan Rounds [18].

Подключение производилось по схеме:

GNG → GND; VCC → 3V3; TXD → RX; RXD → TX.

После перепрошивки устройства была проведена соответствующая настройка openHAB, а именно: Создание вещи (Thing) MQTT Broker (Для чего отдельно потребовалось установить брокер Mosquitto [19] на виртуальную машину), создание вещи для Реле, создание каналов (Channels), элементов (Items) (рис. 11).

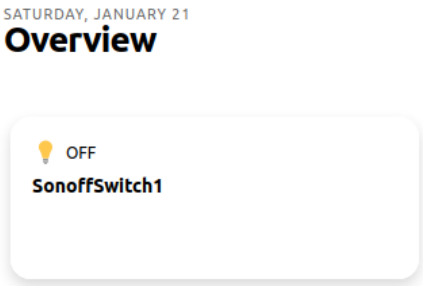


Рис. 11. Вид созданной иконки управления устройством в openHAB

В итоге устройство Sonoff Basic R2 было подключено к openHAB, и может управляться из интерфейса программы, при этом после физического нажатия на кнопку включения/выключения реле, его статус в системе также изменяется на соответствующий. С другой стороны реле можно подключить к любому устройству, например к настольной лампе.

Разработка взаимодействия «умного дома» и распознавания жестов

После достижения точности распознавания в ~95% Рис. 6, можно переходить к основной теме данного исследования. Взаимодействовать с openHAB извне можно посредством REST API, о чем сказано в самой документации [9]. То есть процесс взаимодействия строится на отправке curl запросов (см. листинг 3). Команда отправки POST запроса для отключения реле путем отправки HTTP-запроса, включающего MQTT Topic.

Листинг 3. Команда отправки POST запроса для отключения My_Item путем отправки HTTP-запроса, включающего MQTT Topic.

```
curl -X POST --header "Content-Type: text/plain" --header "Accept:
application/json" -d "OFF" "http://{openHAB_IP}:8080/rest/items/My_Item"
```

Исходя из вышесказанного, чтобы совместно использовать распознавание жестов и openHAB, было решено добавить логику взаимодействия в основной код распознавания жестов.

Первым шагом стало написание функции триггера, а также его операторов, код данной процедуры представлен в листинге 4.

Листинг 4. Код функции триггера.

```
def trigger(mode):
    """
    Функция триггера реле
    :param mode: "ON" or "OFF"
    :return: отправляет запрос триггера в openhab
    """
    os.system(f'curl -X POST
"http://192.168.31.87:8080/rest/items/SonoffBasic_SonoffSwitch1" -H \
```

```
                        "Accrpt: */*" -H "Content-Type: text/plain" -d
"{mode}"')
    print(f'triggered to: {mode}')

# Операторы триггера
if check_repeats(symbol, 'ON'):
trigger('ON')
if check_repeats(symbol, 'OFF'):
    trigger('OFF')
```

На данном этапе, возникла проблема с частотой срабатывания, то есть реле быстро начинало включаться/выключаться, когда были случайные распознавания. То есть было обнаружено, что распознавание жестов происходит в каждом кадре, а учитывая, что для тестируемого устройства, количество кадров в секунду в диапазоне 15–20, было решено добавить функцию проверки последовательности, код данной процедуры представлен листингом 5. Это означает, что функция проверяет последовательность (продолжительность) показанных жестов по-кадрово, и если список из одинаковых жестов больше 14 это примерно 0,75 секунд, то запускается определенное действие. Число 14 было выбрано как оптимальный баланс между скоростью показа жестов и действиями устройства идущего за ним, тем самым обеспечивая комфортное пользование системой при экспериментальных тестах.

Листинг 5. Код функции проверки последовательности.

```
def check_repeats(lst, word):
    """
    Функция проверки последовательности показанных жестов
    :param lst: список показанных жестов
    :param word: жест используемый в качестве триггера
    :return: повторяется ли жест на протяжении времени или нет
    """
    try:
        action_list = lst[-15:]
    except:
        action_list = lst
    for i in range(len(action_list) - 1):
        counter = 1
        while (i + counter < len(action_list) and action_list[i] ==
action_list[i + counter] and action_list[i] == word):
            counter += 1
        if counter > 14:
            return True
    return False
```

Жесты из рис. 10 были выбраны как жесты управления реле.

Заключение

В этой статье представлен способ построения системы распознавания жестов в умном доме, которая включает в себя разработку системы распознавания жестов, построение системы умного дома и построения взаимодействия между этими системами.

Нельзя не отметить, что в работе мы столкнулись с рядом препятствий, такими как, например, выбор подхода к моделированию жестов, анализ существующих подходов к распознаванию жестов и последующий ему выбор подходящего решения, анализ существующих решений умного дома с открытым исходным кодом, построение своего умного дома на основе openHAB с учетом необходимости использования стороннего ПО для реле, разработка взаимодействия двух систем с учетом низкой производительности устройств на которой производились тесты и реализация.

Результаты этой статьи можно резюмировать следующим образом: данная система может быть развернута в умном доме, на производстве, в любых местах, где можно взаимодействовать с камерой и IoT-устройствами. Данная система имеет большое преимущество в виде обученной модели TensorFlow-Lite, что значит, что она может быть развернута на менее производительных устройствах типа микрокомпьютер, например Raspberry 3, 4.

Построенные системы по-отдельности, так и вместе могут нести практическую пользу людям с ограниченными возможностями во взаимодействии с «роботами», обычным пользователям, везде, где может потребоваться взаимодействовать с интерфейсом ПО посредством жестов, начиная от врачей хирургов, делающих операции, продолжая пилотами, операторами систем и так далее.

Литература

1. Kadumudi F.B. et al. The manufacture of unbreakable bionics via multifunctional and self-healing silk–graphene hydrogels // Advanced Materials. 2021. V. 33. №. 35. P. 2100047.
2. Портная М., Ильин И., Ковалев О. «Tilt – браслет с распознаванием жестов». 2019. [Электронный ресурс]. Режим доступа: https://vc.ru/tribuna/92445-tilt-braslet-s-raspoznavaniem-zhestov/ (дата обращения: 17.12.2021).
3. Dinh D.L., Kim T.S. Smart home appliance control via hand gesture recognition using a depth camera //Smart Energy Control Systems for Sustainable Buildings. Springer, Cham, 2017. P. 159-172.
4. «Hand Tracking». 2021. [Электронный ресурс]. Режим доступа: https://www.computervision.zone/courses/hand-tracking/ (дата обращения: 25.03.2022)
5. «MediaPipe Hands». 2019. [Электронный ресурс]. Режим доступа: https://google.github.io/mediapipe/solutions/hands.html. (дата обращения: 25.12.2021).
6. «American Sign Language (ASL) Recognition using Hand Landmarks» – 2021 [Электронный ресурс]. Режим доступа: https://github.com/ cortictechnology /hand_asl_recognition (дата обращения: 14.12.2022).
7. «Hand tracking with DepthAI». 2023 [Электронный ресурс]. Режим доступа: https://github.com/geaxgx/depthai_hand_tracker (дата обращения: 12.01.2023).
8. «Classifying American Sign Language Alphabets on the OAK-D» – 2021 [Электронный ресурс]. Режим доступа: https://www.cortic.ca/post/classifying-american-sign-language-alphabets-on-the-oak-d (дата обращения: 15.12.2022).
9. openHAB REST API – 2023 [Электронный ресурс]. Режим доступа: https://www. openhab.org/docs/configuration/restdocs.html (дата обращения: 15.09.2022).
10. Fischer I., «Weitergepuzzelt-die wichtigsten neuen erweiterungen der smart-home-steuersoftware iobroker sein,'» in Proc. Heise Medien. 2019. P. 178-179.
11. «Hand-gesture-recognition-using-mediapipe». 2023 [Электронный ресурс]. Режим доступа: https://github.com/Kazuhito00/hand-gesture-recognition-using-mediapipe (дата обращения: 20.01.2023).
12. Примеры разделения датасета на train и test с Scikit-learn [Электронный ресурс]. Режим доступа: https:// pythonru.com/ baza-znanij/ sklearn-train- test-split (дата обращения: 23.01.2023).
13. The Sequential model [Электронный ресурс]. Режим доступа: https:// keras.io/ guides/ sequential_model/ (дата обращения: 15.01.2023).
14. Обзор Keras для TensorFlow [Электронный ресурс]. Режим доступа: https:// habr.com/ ru/post/482126/ (дата обращения: 30.12.2022).
15. Реализуем и сравниваем оптимизаторы моделей в глубоком обучении – 2020. [Электронный ресурс]. Режим доступа: https:// habr.com/ru/ company/ skillfactory/ blog/525214/ (дата обращения: 24.01.2023).
16. Строим градиентные алгоритмы оптимизации Adam, RMSProp, Adagrad, Adadelta [Электронный ресурс]. Режим доступа: https://proproprogs.ru/tensorflow/tf-stroim-

gradientnye-algoritmy-optimizacii-adam-rmsprop-adagrad-adadelta (дата обращения: 25.01.2023).
17. Open source firmware for ESP devices – Tasmota [Электронный ресурс]. Режим доступа: https://tasmota.github.io/docs/ (дата обращения: 10.10.2022).
18. Rounds J. Sonoff Basic R2 + Tasmota (no solder) – 12.2019 [Электронный ресурс]. Режим доступа: https://medium.com/@jordanrounds/sonoff-basic-r2-tasmota-aa6f9d4e033f (дата обращения: 05.10.2022).
19. Eclipse Mosquitto™ An open source MQTT broker [Электронный ресурс]. Режим доступа: https://mosquitto.org/ (дата обращения: 19.09.2022).

УДК 004.93'12

РАЗРАБОТКА ИНФОРМАЦИОННОЙ СИСТЕМЫ ЭЛЕКТРОННОГО ФОТОАРХИВА КРАЕВЕДЧЕСКОЙ ИНФОРМАЦИИ

Петухов С.А.[1] (студент)
Научный руководитель – доцент образовательно-научного кластера "Институт высоких технологий" БФУ им. И. Канта Савкин Д.А.[1]
1 – Университет БФУ им Канта
e-mail: LexusLFA477@gmail.com

Аннотация
В работе исследователь описывает процесс разработки информационной системы электронного фотоархива краеведческой информации. Работа охватывает анализ существующих аналогов и выявление проблем их использования, а также использование современных навыков обработки изображений с помощью искусственного интеллекта. В работе описывается процесс разработки структуры датасета для дообучения искусственного интеллекта (ИИ) и функционал системы, которая включает в себя функции хранения, поиска и отображения фотографий. Основной особенностью системы является наличие функционала для исключения дубликатов и поиска фотографий на базе ИИ, обученного на сформированном датасете из более 10 000 изображений, что позволяет обеспечить быстрый доступ к краеведческой информации, а также сохранение ценных фотографий для будущих поколений.
Ключевые слова
Нейронные сети, датасет, эмбеддинг, искусственный интеллект, краеведение.

При работе с краеведческой информацией исследователю становится доступным большой объем графической информации: фотографии, карты, схемы и т.п. Эта информация может попадать в руки исследователя в искаженном или неполном виде. Так, например, из различных источников могут поступить разные варианты одной и той же фотографии: при её публикации могут быть по-разному проведены кадрирование объекта съёмки, его ретуширование (исчезают или появляются разные объекты – люди, здания и др.), а также по-разному выполнена различная обрезка фото [2-3]. Также возможно получение фотографий разного размера, качества или цвета (монохромная фотография или цветная, в том числе с применением современных программ раскрашивания фото). В настоящее время существует большое количество источников фотографий: интернет, книги и фотоальбомы. Поэтому становится актуальным процесс сравнения имеющейся в датасете системы фотографий с вновь обнаруженной информацией. Все это ставит перед краеведом сложную задачу, распадающуюся на две подзадачи. Во-первых, поиск по некоторым критериям данной информации в датасете фотографий. Во-вторых, сравнительный анализ существующего там и поступившего вновь изображения с принятием решения по замене или отказу от замены фотографии в датасете.

В работе продемонстрирована задача создания фотоархива для Калининградской области с возможностью поиска фотографий по совпадениям. В результате работы была разработана информационная система, куда можно загружать фотографию с достопримечательностью города Кенигсберг, а затем система дает предположение содержимого фотографии в виде названия достопримечательности и приводит случайные фотографии этого же объекта.

Существует около 400 нейросетей для классификации изображений, такие как: VGG, DenseNET, RegNet, ResNet и т.д. При разработке функции поиска в информационной системе электронного фотоархива краеведческой информации была использована нейросеть ResNet, которая является одной из наиболее эффективных нейронных сетей для классификации изображений. На этот выбор повлияли следующие

факторы: ResNet позволяет создавать глубокие нейронные сети, которые могут эффективно обрабатывать сложные изображения. Метод ResNet, с помощью использования блоков соединений, обеспечивает более эффективное обучение сети, уменьшая проблему затухания градиента. ResNet позволяет использовать предобученные модели, что значительно ускоряет и упрощает процесс обучения [1].

Сверточные нейронные сети были разработаны специально для работы с изображениями и способны автоматически извлекать признаки из изображений, что делает их идеальным выбором для классификации.

ResNet состоит из нескольких блоков, каждый из которых включает в себя сверточные слои, батч-нормализацию и функцию активации. Сверточные слои – это тип слоев в нейронной сети, который используется для обработки изображений и других типов данных, где пространственная структура важна. Сверточные слои анализируют изображение по маленьким областям (например, 3x3 пикселя), называемым фильтрами, и находят в них определенные особенности (например, границы или текстуры). Эти особенности затем объединяются в более высокоуровневые признаки, что позволяет сети распознавать объекты на изображении. Батч-нормализация – это метод, который используется для нормализации данных на каждом слое нейронной сети. Это делается путем вычисления среднего значения и стандартного отклонения для каждого мини-батча (небольшой подмножества данных, используемого для обучения сети) и затем масштабирования данных с использованием этих значений. Это позволяет сети быстрее и более стабильно обучаться, ускоряет сходимость и может предотвращать проблемы, такие как затухание градиента.

Однако, для обучения нейросети ResNet на задаче классификации фотографий краеведческой информации, необходимо иметь большой набор размеченных данных. В случае, если таких данных нет, можно использовать метод дообучения TripletLoss.

TripletLoss – это метод дообучения нейросетей, который используется для построения эмбеддингов изображений, то есть векторов, содержащих в себе информацию об изображении. Для этого необходимо иметь набор изображений, который разбит на три части: якорную, положительную и отрицательную. Якорная часть – это изображение, для которого строится эмбеддинг. Положительная – это изображение, которое является ближайшим соседом якорной части. Отрицательная – это изображение, которое является наиболее далеким соседом якорной части.

Для обучения нейросети с использованием метода TripletLoss необходимо минимизировать расстояние между эмбеддингом якорной и положительной частей, и максимизировать расстояние между эмбеддингом якорной и отрицательной частей. Таким образом, нейросеть строит более точные эмбеддинги для изображений, что позволяет улучшить качество классификации.

Начало процесса дообучения любой сверточной нейронной сети является подготовка датасета, в моем случае - состоящего из довоенных фотографий Кенигсберга из более чем 10000 фотографий. Эти фотографии были распределены на 84 класса. Завершив подготовку датасета, была подготовлена программа на языке Python с использованием библиотеки PyTorch для дообучиния нейронной сети, на выявления достопримечательностей Кенигсберга на снимках. Пример работы программы приведен на рисунке.

В результате работы был разработан веб-сервис фотоархива для Калининградской области с возможностью поиска достопримечательностей по изображению. Для поиска была дообучена нейросеть ResNet с помощью TripletLoss. Данные для дообучения были получены благодаря тесному взаимодействию с Министерством по культуре и туризму Калининградской области, а также с Государственным архивом Калининградской области. Информационная система электронного фотоархива краеведческой информации может использоваться при организации экскурсий, проведения научных исследований и сохранения исторического наследия.

Рисунок. Пример результата работы сервиса

Литература

1. Клестов Р.А., Стоблов В.Ю. Сравнение моделей классификации на основе нейронных сетей глубокого обучения ResNet и VGG // Искусственный интеллект в решении актуальных социальных и экономических проблем XXI века. 4. 2019. С. 50–56. URL: https://permai.ru/files/sbornik_trudov_2019_1.pdf#page=51 (дата обращения: 17.01.2023).
2. Головко В.А. Глубокое обучение нейронных сетей: теория и применение // Вычислительные методы, модели и образовательные технологии. 2015. С. 4–7. URL: https://rep.bstu.by/bitstream/handle/data/28934/4-7.pdf?sequence=1 (дата обращения: 30.01.2023).
3. Ховард Д., Сильвейн Г. Глубокое обучение с fastai и PyTorch: минимум формул, минимум кода, максимум эффективности: СПб.: Питер, 2022. 624 с. URL: https://www.google.ru/books/edition/Глубокое_обучение_cfas/QKR9EAAAQBAJ?hl=ru&gbpv=1 (дата обращения: 13.12.2022).

УДК 004.09

УЯЗВИМОСТИ СМАРТ-КОНТРАКТОВ SOLIDITY

Попов И.А.[1] (студент)
Научный руководитель – к.т.н., доцент ФИКТ Капитонов А.А.[2]
1 – Университет ИТМО
e-mail: vanchespop@yandex.ru

Аннотация
За все время разработки и проверки смарт-контрактов были выявлены уязвимости, связанные с языком программирования Solidity, создающие для злоумышленников потенциальные бреши в безопасности. В данной статье будут рассмотрены и классифицированы некоторые из таких уязвимостей.
Мотивацией к написанию данной статьи послужило желание внести вклад в безопасность и надежность создаваемых смарт-контрактов, так как эти параметры являются основополагающими и именно от них зависит сохранность средств пользователей.
Ключевые слова
Ethereum, Solidity, смарт-контракт, блокчейн, одноранговая сеть, уязвимость. Распределенные технологии.

Технология блокчейна в последние годы привлекает к себе неуклонно растущее внимание как со стороны академических кругов, так и со стороны промышленного сектора. Данная технология представляет собой полностью распределенный публичный реестр и одноранговую платформу, которая использует криптографию для безопасного размещения приложений, передачи цифровых валют, сообщений и хранения данных.

На основе текущей рыночной капитализации одной из самых популярных блокчейн-платформ является Ethereum. Виталий Бутерин, главный разработчик Ethereum, описывает Ethereum как блокчейн общего назначения, что означает, что сеть Ethereum способна поддерживать исполнение алгоритмов, написанных на языках программирования общего назначения.

Это позволяет разработчикам создавать множество приложений, начиная от простых кошельков и заканчивая финансовыми системами, системами торговли энергией или другими криптовалютами. Вместо того, чтобы создавать отдельный блокчейн для каждого варианта использования или приложения, с помощью технологии, известной как смарт-контракты, все отдельно взятые планы и идеи могут быть реализованы в рамках одного Ethereum блокчейна.

Тот факт, что смарт-контракты Ethereum выполняются правильно, является необходимым условием их эффективности: в противном случае злоумышленник может вмешаться в исполнение, например, чтобы перевести часть денег от законного пользователя к себе.

Solidity является языком программирования, на котором написано большинство существующих смарт-контрактов в сети Ethereum. Возможность автономно существующих и исполняемых в сети программ работать с денежными транзакциями, а также иметь свой собственный баланс, делает смарт-контракты привлекательной целью для атак злоумышленников, заинтересованных в подрыве репутации владельца и в незаконном обогащении.

Именно поэтому написание качественных и безопасных смарт-контрактов является ключевым фактором в вопросе избежания репутационных и финансовых потерь компании, создающей децентрализованные приложения.

К сожалению, наличие брешей в безопасности смарт-контрактов Ethereum - не редкость. Есть несколько причин, которые делают реализацию смарт-контрактов

особенно подверженной ошибкам в Ethereum. Значительная часть из них связана с Solidity, языком программирования высокого уровня, поддерживаемым Ethereum.

Многие уязвимости, по-видимому, вызваны несоответствием между семантикой Solidity и ее трактовкой программистами. Одной из проблем является тот факт, что Solidity, хотя и выглядит как статически типизированный Javascript-подобный язык с присущими ему исключениями и функциями, реализует некоторые из этих функций достаточно своеобразно. Вместе с этим Solidity не вводит конструкций для решения специфичных для предметной области аспектов [1]. Таким образом, например, промежуточные этапы вычислений записываются в общедоступный блокчейн, где их очередь исполнения может быть непредсказуемо переупорядочена.

Другой важной причиной распространения небезопасных смарт-контрактов является то, что документация об известных уязвимостях разбросана по нескольким источникам, включая официальную документацию, исследовательские работы и статьи [2], а также различные форумы в сети Интернет. Всеобъемлющий, автономный и своевременно обновляемый обзор уязвимостей и атак на смарт-контракты Ethereum по-прежнему отсутствует.

В данной работе будут рассмотрены основные уязвимости смарт-контрактов, написанных на языке Solidity.

Одной из наиболее значимых уязвимостей, на которую следует обратить внимание при разработке смарт-контрактов с Solidity, являются уязвимости повторного входа или повторного вызова [3]. Основная опасность заключается в том, что злоумышленник может полностью перевести баланс уязвимого контракта к себе.

Атака повторного входа может произойти, когда в коде контракта существует функция, которая выполняет внешний вызов другого контракта до того, как будут произведены ограничивающие количество вызовов изменения. Если злоумышленник имеет контроль над внешним смарт-контрактом, он может выполнить рекурсивный обратный вызов исходной функции, повторяя взаимодействия, которые в противном случае не выполнялись бы из-за установленных ограничений.

Простейший пример на Листинге 1 представляет из себя функцию вывода средств контракта, ведущего внутренний учет с переменной баланса. Если уязвимый контракт переводит средства до того, как он установит баланс равным нулю, злоумышленник может повторно рекурсивно вызвать функцию вывода средств и вывести все доступные средства контракта.

Листинг 1 – Пример уязвимой функции

```
1. function withdraw() external {
2.     uint256 amount = balances[msg.sender];
3.     require(msg.sender.call.value(amount)());
4.     balances[msg.sender] = 0;
5. }
```

После того, как msg.sender.call.value(amount)() переведет правильную сумму средств, функция злоумышленника снова вызывает вывод средств, переводя средства снова, до того, как баланс [msg.sender] = 0 сможет остановить дальнейшие переводы.

Это будет продолжаться до тех пор, пока либо в исходном смарт-контракте не останется средств, либо стек вызовов не будет переполнен. Потенциальным решением этой уязвимости станет перенос кода, ограничивающего действия вывода средств, до вызова внешней функции. Пример такого кода представлен на Листинге 2.

Листинг 2 – Защита от рекурсивного вызова

```
1. function withdraw() external {
2.     uint256 amount = balances[msg.sender];
3.     balances[msg.sender] = 0;
4.     require(msg.sender.call.value(amount)());
5. }
```

В дополнение, хорошей практикой считается маркировка внутри кода потенциально опасных непроверенных функций, содержащий внешний вызов, предложенный на Листинге 3:

Листинг 3 – Маркировка небезопасных функций

```
function untrustedWithdraw() public {
    uint256 amount = balances[msg.sender];
    require(msg.sender.call.value(amount)());
    balances[msg.sender] = 0;
}
```

Поскольку большинство атак повторного входа связаны с функциями send, transfer или call, важно понимать разницу между ними.

Функции send и transfer считаются более безопасными, поскольку они имеют ограничение по газу в 2300 единиц [4]. Такое ограничение предотвращает многочисленные обратные вызовы внешней функции к целевому контракту.

Функция call, к сожалению, гораздо более уязвима.

Когда ожидается, что вызов внешней функции повлечет за собой выполнение сложных операции, обычно используется функция call, поскольку она имеет возможность манипулировать всем объемом доступного газа.

Для уменьшения рисков безопасности рекомендуется использовать send и transfer везде, где это позволяет сделать запланированная логика контракта [5]. Это позволит защититься от бесконечных рекурсивных вызовов со стороны злоумышленников.

Далее следует рассмотреть уязвимость отказа в обслуживании, послужившую причиной реальных атак [6] на примере частично представленных двух простых и похожих друг на друга смарт-контрактов, реализующих функционал игры King of the Ether Throne, представленный на Листинге 4.

Если кто-то хочет стать королем, он должен заплатить немного Ethereum действующему королю, оставив небольшую комиссию самому контракту. Цена получения трона постепенно увеличивается.

Листинг 4 – Контракты King of the Ether Throne и Attacker

```
contract TransferKotET {
address payable king;
uint public rewardPrice = 500;
function() external payable {
require (msg.value >= rewardPrice);
uint MCrownPrice = MfindCrownPrice();
king.transfer(MCrownPrice);
king = msg.sender;
rewardPrice = NRewardPrice();
}

contract CallKotET {
function() external payable{
require(msg.value >= rewardPrice);
uint MCrownPrice = MfindCrownPrice();
(bool success,) = king.call.value(MCrownPrice)("");
require(success);
king = msg.sender;
rewardPrice = NewRewardPrice();
}

contract Attacker {
function unseatKing(address a, uint w) payable {
a.call.value(w);
}

function () external payable {
revert();
}}
```

На первый взгляд, контракты выглядят безопасными, однако оба содержат в себе уязвимость, связанную с отправкой компенсации действующему королю.

Предположим, что Attacker смещает действующего короля, отправляя необходимую сумму контракту. После этого он останется королем навсегда, так как любые последующие попытки отправить компенсацию Attacker за трон будут неудачными из-за функции revert() в 28 строке и код контрактов игры, отвечающий за смену действующего короля, не будет исполняться.

Иными словами, функциональность контракта будет нарушена, и другие пользователи не смогут им воспользоваться.

Неправильно обработанное исключение также является часто встречающейся уязвимостью. Исключения представляю из себя разнообразные ошибки во время выполнения кода программы. Одним из хорошо известных исключений в сети Ethereum является исключение из-за нехватки газа.

В Solidity для отправки Ethereum в другой смарт-контракт используются функции send, transfer и call. Если исключение возникает в вызываемом смарт-контракте, transfer распространяет исключение в вызывающий смарт-контракт, что является безопасным. В то время как другие API, такие как call, send и delegatecall, в случае ошибки возвращают значение false, что не блокирует выполнение кода.

Таким образом, если во время разработки смарт-контракта будет пропущена необходимая проверка результата того или иного действия, выполнение продолжится, что приведет к несогласованному состоянию и непредсказуемому результату работы программы.

Проблема нехватки газа для выполнения функций смарт-контракта тоже представляется существенной проблемой. Ограничение по использованию газа для функции send установлено в 2300 единиц, чего достаточно для вызова пустой резервной функции.

Однако в случае, если резервная функция (по ошибке разработчика или из злого умысла) взаимодействует с состоянием смарт-контракта, то для ее выполнения потребуется больше газа, чем 2300 единиц. Код потенциально проблемной функции представлен на Листинге 5.

Листинг 5 – Дорогостоящая резервная функция

```
1. contract Sender {
2. function pay(uint val, address payable _recv)public {
3. if(_recv.send(val)){} else {}
4. }
5. }
6.
7. contract Receiver {
8. uint public TotalBal = 0;
9.
10.     function() external payable {
11.     TotalBal = TotalBal + msg.value;
12.     } }
```

Вызов в строке 3 повлечет за собой ошибку нехватки газа.

В случаях, когда возникшее исключение из-за нехватки газа не обрабатывается должным образом, злоумышленники могут притвориться, что они уже отправили эфир получателю, фактически сохраняя сумму отправки в кошельке.

Использование функции call вместо send поможет избежать ошибки нехватки газа, однако обнажит потенциальную уязвимость повторного входа, рассмотренную ранее.

В отдельную категорию также принято выделять внутренние уязвимости. Уязвимости такого рода в контрактах влияют на сам смарт-контракт изнутри [7]. Из-за них не только злоумышленник, но и владелец может причинить вред пользователям. Владелец может спроектировать смарт-контракт таким образом, чтобы предотвратить утечку (или передачу) любых средств, тем самым превратив смарт-контракт в черную дыру. Однако худшее случается, когда злоумышленник использует незащищенную

команду самоуничтожения смарт-контракта. Злоумышленник получает доступ к балансу, если ему удается изменить владельца смарт-контракта перед его уничтожением.

Уязвимость, носящая название «Черная дыра», характеризуется наличием у смарт-контракта функциональности, отвечающей за получение средств, однако код, который должен выводить их каким-либо образом или отсутствует, или всецело полагается на стороннюю библиотеку (импортирован из нее).

Смарт-контракты первого типа принято также называть «жадными», так как в них присутствует payable метод, однако ни один из программных путей не приводит к вызову функций вывода средств. Код такого смарт-контракта представлен на Листинге 6.

Листинг 6 – Контракт без вывода средств

```
1. contract BlackHole {
2. function () external payable { }
3. }
```

Контракты, относящиеся ко второму типу, ограничивают возможности вывода импортированными библиотеками. Но если библиотека смарт-контракта в итоге перестанет существовать, выполнив команду selfdestruct, или какой-либо другой смарт-контракт (или злоумышленник) случайно (или намеренно) удалит библиотеку, тогда это закроет все пути для извлечения средств с баланса смарт-контракта. Такой случай реализован в коде, представленном на Листинге 7.

Листинг 7 – Контракт с зависимым выводом средств

```
1. contract LibraryDependent {
2. address _nl = 0xNewLibrary;
3. function withdraw() {
4. _nl.delegatecall(msg.data);
5. }
6. }
```

Потенциальной уязвимостью считается наличие небезопасного метода самоуничтожения смарт-контракта. Метод selfdestruct позволяет уничтожить смарт-контракт при необходимости, при этом все средства с его баланса будут переведены на баланс того аккаунта, адрес которого был заложен в логике смарт-контракта. Как только контракт будет уничтожен, всем его значениям будут присвоены нулевые значения, код контракта будет удален, воспользоваться его функциональностью будет невозможно.

Данный метод используется в чрезвычайных ситуациях. Смарт-контракт считается уязвимым для самоуничтожения [8], если программный код, содержащий вызов selfdestruct не защищен должным образом (не содержит никакого условия для активации метода уничтожения). Вариант такого небезопасного кода представлен на Листинге 8.

Листинг 8 – Контракт с незащищенным самоуничтожением

```
1. contract SelfDestructExample {
2. address payable owner;
3.
4. constructor() public {
5. owner = msg.sender;
6. }
7.
8. function DeleteMe() public {
9. selfdestruct(owner);
10.     }}
```

Для удаления смарт-контракта злоумышленнику требуется доступ к блоку кода, ответственному за самоуничтожение, или обладание правами собственности на смарт-контракт.

Еще одна уязвимость, именуемая уязвимостью небезопасного баланса, основывается на определении владельца контракта вне конструктора смарт-контракта.

Это приводит к тому, что в примере, приведенном на Листинге 9, любой пользователь может переназначить владельца и воспользоваться этой брешью в безопасности для несанкционированного вывода средств с баланса контракта.

Листинг 9 – Контракт с возможностью переопределения владельца

```
contract UnsecuredBalance {
address payable owner;

function ReassignOwner(address payable _owner) public{
owner = _owner;
}

function withdraw(uint _amount) public {
if (msg.sender == owner){
    if(!owner.send(_amount)){
    }
    else {}
    }}}
```

Еще в одну группу выделяют арифметические уязвимости. Подобные уязвимости проявляются при выполнении различных математических операций и прежде всего создает проблемы в случае выхода значения за рамки минимального или максимального значения, предусмотренного конкретным целочисленным типом переменной.

Для устранения проблем неточности вычислений рекомендуется провести аудит смарт-контракта, либо прибегнуть к помощи библиотеки SafeMath или Solc-Verify [9].

Ethereum самостоятельно не взаимодействует с типами данных, доверяя всю работу по выявлению и обработке ошибок компилятору. Solidity обладает богатым набором целочисленных типов (uint8, uint16, uint24, …uint256, int8, int16, int24, ...int256), однако совсем не поддерживает значения с плавающей запятой.

Пример на Листинге 10 демонстрирует некорректность вычисления значения переменной. При попытке уменьшить uint8 равный 0 на единицу его значение станет равным 255 (после выполнения функции Decrement).

Листинг 10 – Контракт с ошибкой вычисления значения

```
contract IntegerUnderflow {
uint8 value = 0;

function Decrement(address payable _owner) public returns uint8 {
value--;
return value;
}
}
```

Хорошей практикой, помогающей держать результаты вычислений под контролем, считается использование библиотеки, оснащающей арифметические операции дополнительной предварительной проверкой, представленной на Листинге 11.

Листинг 11 – Библиотека безопасных вычислений

```
library SafeMath {
function sub(uint8 x, uint8 y) internal pure returns (uint8) {
assert(y <= x);
return x-y;
}
}

contract IntegerUnderflow {
uint8 value = 0;


    function Decrement(address payable _owner) public returns uint8
{
    value = value.sub(1);
    return value;
    }
    }
```

Помимо небезопасных математических операций, переполнение целочисленных значений может быть и неявным.

В приведенном на Листинге 12 примере для итерирования по массиву длиной 500 используется задекларированная как uint8 переменная, максимальным значением которой будет 255. Аналогичная уязвимость может быть воспроизведена и при работе с массивом динамической длины, которая в какой-то момент выполнения программы превысит максимально доступное значение итератора.

Листинг 12 – Переполнение при переборе массива

```
1. library SafeMath {
2. function sub(uint8 x, uint8 y) internal pure returns (uint8) {
3. assert(y <= x);
4. return x-y;
5. }
6. }
7.
8. contract IntegerUnderflow {
9. uint8 value = 0;
10.
11.
12.      function Decrement(address payable _owner) public returns uint8
{
13.      value = value.sub(1);
14.      return value;
15.      }
16.      }
```

Помимо арифметических уязвимостей принято классифицировать также и уязвимости смарт-контрактов, связанные с газом.

Газ предназначен для решения двух основных задач внутри виртуальной машины Ethereum [10]. Во-первых, газ служит платой для стейкеров при регистрации транзакций. Во-вторых, газ действует как своеобразное топливо для транзакций и, таким образом, позволяет регулировать очередность и скорость их выполнения.

Логически это означает, что, если пользователь не заплатит достаточную количество газа, необходимое для транзакции, транзакция завершится неудачей, выбросив исключение нехватки газа [11].

Виртуальная машина Ethereum (EVM) защищается от перегрузок и атак типа «отказ в обслуживании» путем принудительного завершения транзакций. EVM расходует газ постепенно, и, если на каком-то шаге выполнения необходимое количество газа превышает доступное, EVM завершает выполнение смарт-контракта, вызывая частичный или полный откат совершенных транзакций.

В приведенном на Листинге 13 примере злоумышленник может манипулировать размером массива неопределенной длины, добавляя в него дополнительные адреса. Таким образом, увеличивается общая стоимость исполнения всех транзакций, что может привести к истощению запасов газа и к отказу в обслуживании.

Представленная реализация позволяет сохранить текущий обрабатываемый адрес, чтобы в случае ошибки нехватки газа продолжить выполнение цикла с нужного места без повторов.

Листинг 13 – Потенциально дорогостоящий цикл

```
1. contract CostlyLoop {
2. uint constant Gas = 100000;
3. address payable addrArr;
4. function LongList(uint256 memory nextV, uint[ ] memory arr, address
payable _addr) public {
5. uint256 i= nextAddr;
6. addrArr = _addr;
7. for ( ;i < arr.length && gasleft() > Gas; i++) {
8. addrArr.send(arr[i]);
```

```
9. }
10.      nextAddr = i; }
11.      }
```

Представленная на Листинге 14 уязвимость транзакций в цикле возникает, если смарт-контракт использует цикл для отправки средств нескольким получателями. В случае, если одна из отправок завершится неудачей, все остальные транзакции также не смогут завершиться.

Листинг 14 – Транзакции в цикле

```
1. for (uint i = 0; i < worker.length ; i++) {
2. if (worker [i].paid < min_salary) {
3. if (!(worker [i].addr.send (worker [i].bonusAmount))) {
4. revert() ;
5. }
6. worker [i] = newWorker;
7. }
8. }
```

Использование устаревшего throw для обработки неудачной транзакции потребляло весь доступный газ и блокировало смарт-контракт. Использование же revert для обработки исключения создает риск бесконечного цикла и потери всего газа, если один из получателей на постоянной основе не может получить перевод (например, из-за неправильной резервной функции).

Таким образом, в данной работе были рассмотрены, описаны и классифицированы основные и наиболее известные уязвимости смарт-контрактов, написанных на Solidity.

В целом, основная проблема уязвимостей заключается в отсутствии возможности контроля выполнения и редактирования контракта после его развертывания. Это позволяет контрактам работать абсолютно автономно, однако накладывает большую ответственность на разработчика, который должен предусмотреть и разрешить все потенциальные уязвимости до публикации контракта в сети.

Имея представление о возможных угрозах, описанных в данной работе, и придерживаясь паттернов безопасной разработки [12], разработчик с большей вероятностью сможет избежать ошибок.

Литература

1. Dika, Ardit. Security Vulnerabilities in Ethereum Smart Contracts / Ardit Dika, Mariusz Nowostawski. – Norwegian University of Science and Technology. 2018/ Pp. 956-962/
2. Khan Zulfiqar Ali. A Survey on Vulnerabilities of Ethereum Smart Contracts / Zulfiqar Ali Khan, Akbar Siami Namin. – Department of Computer Science Texas Tech University, 2020. Pp. 1-13.
3. Alkhalifah A., Ng A., Watters P.A. A mechanism to detect and prevent Ethereum blockchain smart contract reentrancy attacks. / Alkhalifah A. Ng A. Watters P. A., Kayes A.S.M. Frontiers in Computer Science. 2021. Pp. 1-15.
4. Baek J. M.-g., Shin, K. Cho, Hasanova H.U. A survey on blockchain cybersecurity vulnerabilities and possible countermeasures / H. Hasanova U. J. Baek M.-g. Shin K. Cho, M.-S. Kim. International Journal of Network Management. 2019. Pp. 1-36.
5. Praitheeshan, Purathani. Security Analysis Methods on Ethereum Smart Contract Vulnerabilities: A Survey / Purathani Praitheeshan, Lei Pan. arXiv preprint. 2019. Pp.-21.
6. Prasad Baddepaka. Smart Contracts: Vulnerabilities and Real At tacks / Baddepaka Prasad, Ramachandram S. International Journal of Advanced Research in Engineering and Technology. 2020. Pp. 901-922.
7. Khan Z.A. Ethereum Smart Contracts: Vulnerabilities and their Classifications / Khan Z.A., Namin A. Siami. IEEE International Conference on Big Data (Big Data), 2020. Pp. 1-10.
8. Known attacks [Электронный ресурс]. Режим доступа: https://consensys. github.io/ smart-contract-best-practices/knownattacks/ (дата обращения: 2022.03.20).

9. Hajdu A. Solc-verify: A modular verifier for solidity smart contracts. In Working conference on verified software: theories, tools, and experiments / Hajdu´A., Jovanovi´c D. Springer Cham. 2020. Pp. 161-179.
10. Hildenbrandt, Everett. A complete formal semantics of the ethereum virtual machine / Everett Hildenbrandt. IEEE. 2018. Pp. 204–217.16.
11. Sorbo Di. Profiling gas consumption in solidity smart contracts / Sorbo Di, Laudanna A. Journal of Systems and Software. 2022. Pp. 1-18.
12. Wohrer Maximilian. Smart Contracts: Security Patterns in the Ethereum Ecosystem and Solidity / Maximilian Wohrer, Uwe Zdun. University of Vienna Faculty of Computer Science. 2018. Pp. 2-8.

УДК 004.891.2

ПРОБЛЕМЫ АНАЛИЗА КРИПТОМАЙНИНГА И ИХ РЕШЕНИЕ

Рыбкин Я.С.[1] (магистрант)
Научный руководитель – к.т.н., доцент Войтюк Т.Е.[1]
1 – Университет ИТМО
e-mail: ynchprod@gmail.com

Аннотация
В работе рассматриваются проблемы анализа криптомайнинга и предлагаются методы его оптимизации. Анализ криптомайнинга является сложным процессом, предусматривающим многокритериальную оценку его эффективности, инвестиционных рисков и рентабельности. Для упрощения анализа в качестве решения предлагается использовать методы предиктивной аналитики.

Ключевые слова
Финансы, криптовалюта, криптомайнинг, машинное обучение.

Криптовалюты – это цифровые валюты, которые используют криптографические технологии для защиты своих транзакций и контроля создания новых валютных единиц. Первой и наиболее известной криптовалютой стала Bitcoin, которая появилась в 2008 году. Затем появились Ethereum, Litecoin, Ripple и многие другие.

Общая идея этих валют заключается в том, чтобы убрать централизацию контроля над деньгами и перевести этот контроль к самим пользователям. Таким образом, криптовалюты представляют собой революционный способ передачи средств через Интернет без участия посредников [1].

В последние годы интерес к криптовалютам значительно возрос. Многие инвесторы приобретают криптовалюты в надежде на то, что их цена вырастет в будущем, как представлено на рисунке 1. Однако, рынок криптовалют нестабилен и подвержен волатильности, что требует от инвесторов внимательности и осторожности.

Рис. 1. Логарифмический график криптовалюты Биткоин

Криптовалюты не только могут быть использованы для инвестирования, но и для обмена товарами и услугами, а также для перевода денег без участия банков. Общая концепция криптовалют привлекает множество людей, которые заинтересованы в инновационных технологиях и желают внедрять их в различные области жизни.

Некоторые криптовалюты создаются с помощью процесса, называемого "майнингом". Это происходит путем решения сложных математических задач, которые связаны с проверкой и подтверждением транзакций в блокчейне криптовалюты. Когда

эти задачи решаются, новые блоки добавляются в блокчейн, которые содержат новые транзакции и сообщают всей сети о том, что эти транзакции были подтверждены.

Криптовалюты, созданные с помощью майнинга, могут быть добыты пользователем, который предоставляет вычислительную мощность своего компьютера для решения этих задач. В качестве награды за свои усилия майнеры получают определенное количество криптовалюты [2].

Основным аспектом криптомайнинга является его прибыльность в качестве бизнеса. Следовательно, проведение анализа этой деятельности нацелено на оптимизацию прибыли в контексте изменчивости рынка криптовалют.

Анализ криптомайнинга позволяет оценить рентабельность майнинга, выявить возможные проблемы и риски, связанные с этой деятельностью, а также оценить конкурентную ситуацию на рынке майнинга.

Анализ криптомайнинга усложняется большим количеством параметров, их взаимосвязью и возможными аномалиями. Также при анализе требуется производить оценку рисков и выбор подходящей стратегии инвестирования. А для оценки рентабельности майнинга необходимо учитывать доходы, расходы, прибыли и риски.

Таким образом, для анализа криптомайнинга требуется решать следующие задачи:

1. Определение оптимального значения каждого частного критерия и их взаимосвязи в многокритериальной задаче анализа криптомайнинга, что оказывает влияние на эффективность майнинга.
2. Учет аномалий в данных, связанных с возможностью изменения курса криптовалют в короткий промежуток времени. Важно иметь возможность быстро реагировать на подобные существенные изменения в данных и адаптировать свои инвестиционные стратегии.
3. Оценка инвестиционных рисков, связанных с инвестированием в оборудование и налогообложением. Анализ рисков позволяет предвидеть потенциальные убытки и принимать меры по их минимизации.
4. Выбор оптимальной стратегии инвестирования. Важно учитывать параметры, связанные с производительностью оборудования, изменением курса криптовалюты и текущей рыночной конъюнктурой.
5. Оценка рентабельности майнинга. Она предполагает оценку всех факторов, от влияния уровня сложности майнинга до технических характеристик используемого оборудования, и является первоочередной задачей.

В целом, решение задач анализа криптомайнинга представляет собой сложный процесс, требующий глубокого знания технических, экономических и математических аспектов.

В контексте экономической выгодности криптомайнинга, рекомендуется эффективно собирать, структурировать, анализировать и прогнозировать данные. Это обеспечит снижение вероятности ошибок, связанных с человеческим фактором в процессе майнинга.

Кроме того, сравнительный анализ доходности майнинга и покупки криптовалюты является полезным инструментом определения наиболее выгодного способа инвестирования капитала. Применение методов прогнозирования курса криптовалюты и доходности майнинга позволяет определить потенциальную прибыль при различных условиях, таких как курс криптовалюты и сложность сети.

Использование этих методов является важным инструментом для принятия решений в области криптомайнинга. Результаты анализа и прогнозирования данных помогают оптимизировать работу в пределах данной индустрии, что увеличивает возможность получения максимальной прибыли.

Существует множество информационных сервисов, которые предоставляют возможность получения актуальных данных по работе блокчейна, криптовалют и криптомайнинга. Сбор этих данных является важным этапом нашего исследования. Мы

рассмотрели различные способы сбора информации, такие как парсинг html, web скрапинг, боты или использование API-интерфейсов и выбрали последний в качестве наиболее подходящего для наших целей. API-интерфейс позволяет получать данные быстро и эффективно, минуя лишние данные и ограничения.

Данные, полученные через API-интерфейсы, имеют более высокую надежность и точность, по сравнению с другими методами сбора данных. Полученные таким способом данные преобразуются в текстовый формат для табличного представления CSV, которые легко читаются и обрабатываются с помощью различных программ и библиотек, что делает их удобным форматом для анализа.

Ключевыми атрибутами, которые необходимы для анализа и исследования данных, являются: тип криптовалюты, её историческая стоимость, сложность сети данной криптовалюты, сложность нахождения блока в блокчейне, мощность сети, время нахождения блока и вознаграждение за блок. Значения данных атрибутов представлены на рисунке 2.

Date	Open	High	Low	Close	Adj Close	Volume
2022-04-20	41501.746094	42126.300781	40961.097656	41374.378906	41374.378906	27819532341
2022-04-21	41371.515625	42893.582031	40063.828125	40527.363281	40527.363281	35372786395
2022-04-22	40525.863281	40777.757813	39315.417969	39740.320313	39740.320313	28011716745
2022-04-23	39738.722656	39935.859375	39352.203125	39486.730469	39486.730469	16138021249
2022-04-24	39478.375000	39845.925781	39233.539063	39469.292969	39469.292969	17964398167
2022-04-25	39472.605469	40491.753906	38338.378906	40458.308594	40458.308594	35445730570
2022-04-26	40448.421875	40713.890625	37884.984375	38117.460938	38117.460938	34569088416
2022-04-27	38120.300781	39397.917969	37997.312500	39241.121094	39241.121094	30981015184
2022-04-28	39241.429688	40269.464844	38941.421875	39773.828125	39773.828125	33903704907
2022-04-29	39768.617188	39887.269531	38235.535156	38609.824219	38609.824219	30882994649
2022-04-30	38605.859375	38771.210938	37697.941406	37714.875000	37714.875000	23895713731
2022-05-01	37713.265625	38627.859375	37585.789063	38469.093750	38469.093750	27002760110
2022-05-02	38472.187500	39074.972656	38156.562500	38529.328125	38529.328125	32922642426
2022-05-03	38528.109375	38629.996094	37585.621094	37750.453125	37750.453125	27326943244
2022-05-04	37748.011719	39902.949219	37732.058594	39698.371094	39698.371094	36754404490
2022-05-05	39695.746094	39789.281250	35856.515625	36575.140625	36575.140625	43106256317
2022-05-06	36573.183594	36624.359375	35482.132813	36040.921875	36040.921875	37795577489
2022-05-07	36042.503906	36129.925781	34940.824219	35501.953125	35501.953125	24375896406
2022-05-08	35502.941406	35502.941406	33878.964844	34059.265625	34059.265625	36763041910
2022-05-09	34060.015625	34222.074219	30296.953125	30296.953125	30296.953125	63355494961
2022-05-10	30273.654297	32596.308594	29944.802734	31022.906250	31022.906250	59811038817
2022-05-11	31016.183594	32013.402344	28170.414063	28936.355469	28936.355469	70388855818
2022-05-12	28936.734375	30032.439453	26350.490234	29047.751953	29047.751953	66989173272

Рис. 2. Таблица курса криптовалюты Биткоин

Тип криптовалюты является важным атрибутом, так как каждый тип может отличаться по техническим особенностям, и существует множество криптовалют с уникальной функциональностью и характеристиками. Историческая стоимость криптовалюты также является важным атрибутом для исследования в связи с динамикой колебания цены и влиянием экономической политики на цену.

Сложность сети криптовалюты связана с процессом оценки сложности и разнообразия криптографических алгоритмов, которые использует криптовалюта.

Сложность нахождения блока в блокчейне представляет собой математическую формулу, используемую для определения сложности нахождения нового блока криптовалюты. Это является центральным аспектом процесса майнинга криптовалюты.

Мощность сети является совокупным количеством вычислительной мощности, которую обладают участники сети. Это показатель силы и стабильности сети, и может использоваться для определения рисков инвестирования и других аспектов, связанных с использованием криптовалюты.

Время нахождения блока - время, которое требуется для нахождения нового блока криптовалюты. Этот атрибут может варьироваться и зависит от характеристик каждой криптовалюты.

Наконец, вознаграждение за блок — это количество криптовалюты, которое участник майнинга получает после нахождения нового блока. Этот атрибут также может

различаться в зависимости от типа криптовалюты и инфляционного давления на криптовалюту.

Для сравнения доходности криптомайнинга с покупкой криптовалюты используются математические модели, которые учитывают различные факторы, включая текущий курс криптовалют, стоимость оборудования для майнинга, плату за электроэнергию и другие расходы.

Анализ полученных данных показывает, что традиционный метод покупки и продажи криптовалюты на данный момент является более эффективным, чем криптомайнинг, что отражено в таблице 1, особенно учитывая значительные затраты на электроэнергию.

Таблица 1

Таблица расчета дохода с 01.02.23 по 01.03.23 при покупке Биткоина

Вложение ($)	Доход ($)
25	25.78
50	51.55
100	103.1
250	257.76
500	515.52
1000	1031.05

Однако для обладателей специальных систем Bitmain S19, а также для тех, кто не платит за электроэнергию, ситуация более благоприятна, что представлено в таблице 2.

Таблица 2

Таблица дохода с помощью ASIC с/без учета электричества с 01.02.23 по 01.03.23

Система ASIC Bitmain	Доход при оплате электричества ($)	Доход без оплаты электричества ($)
S19J PRO	29.16	198
S19 PRO	24.3	188.1
T19	7.92	174.24
S17+	-7.56	138.6

Одной из причин нерентабельности криптомайнинга является рост сложности майнинга, который приводит к увеличению количества вычислительных мощностей, необходимых для генерации блоков. Кроме того, стоимость электроэнергии оказывает значительное влияние на доходность майнинга.

Итак, выбор между криптомайнингом и покупкой криптовалюты зависит от конкретных условий и требует детального анализа. На сегодняшний день традиционный метод покупки и продажи криптовалюты является более эффективным, но использование более мощного специализированного оборудования может изменить эту ситуацию в пользу криптомайнинга.

Для прогнозирования курса криптовалюты все чаще применяются модели машинного обучения, среди которых наиболее часто используется анализ временных рядов. Методология прогнозирования основывается на анализе исторических данных о курсе криптовалюты за предшествующий период. При этом учитываются глобальный тренд на протяжении года и тенденции курса за последние два месяца.

Для обработки данных выбираются соответствующие алгоритмы, такие как дерево принятия решений, случайный лес, алгоритм с долгой краткосрочной памятью (англ. Long short-term memory - LSTM) или интегрированное скользящее среднее с

авторегрессией (англ. autoregressive integrated moving average – ARIMA), в зависимости от типа данных, которые нужно обработать. Некоторые исследователи считают [3], что оптимальным выбором являются алгоритмы LSTM и случайный лес. Эти алгоритмы обладают рядом достоинств, таких как высокая скорость работы с большими объемами данных, возможность нахождения необычных корреляций, идентификации шаблонов и трендов в данных.

Рассмотрим шаги реализации прогнозирования курса на примере алгоритма LSTM:

1. Загрузка полученной ранее выборки с помощью Pandas.
2. Выбор столбца стоимости криптовалюты для прогнозирования.
3. Разбиение выборки на тренировочную и тестовую (80/20).
4. Нормализация значений с помощью MinMaxScaler.
5. Создание временных последовательностей для тренировочной выборки с шагом look_back.
6. Создание модели LSTM с двумя слоями LSTM, Dropout и двумя полносвязными слоями.
7. Компиляция модели с оптимизатором Adam и функцией потерь Mean Squared Error.
8. Обучение модели на тренировочных данных.
9. Подготовка тестовой выборки для предсказания аналогично тренировочной.
10. Предсказание значения Close на тестовой выборке с помощью модели.
11. Обратное масштабирование предсказанных значений с помощью метода inverse_transform.
12. Наконец, строится график, на котором отображается исходная выборка, тестовые данные и предсказанные значения (рис. 3).

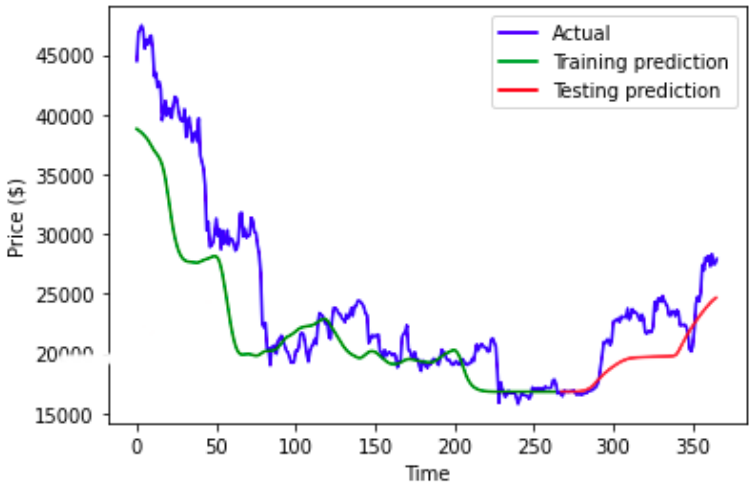


Рис. 3. Курс Биткоина за год и его предсказанная тенденция (красный график)

Использование алгоритмов машинного обучения для прогнозирования курса криптовалюты повышает точность предиктивной аналитики и способствует более эффективному принятию бизнес-решений. При этом учитываются различные факторы, включая сезонные эффекты и множество тенденций. Однако, несмотря на то что использование машинного обучения для прогнозирования курсов криптовалют может быть полезным инструментом для принятия решений, следует понимать, что это решение не является оптимальным и всегда носит рекомендательный характер. В конечном итоге, прогнозирование курса криптовалют – это сложный процесс, зависящий от многих переменных, которые не всегда можно предугадать и принять во внимание.

Предсказание доходности криптомайнинга является актуальной задачей, особенно с учетом увеличивающейся сложности сети и колебаний курсов криптовалют. Перед

исследователями стоит вопрос о разработке моделей, которые бы учитывали как основные факторы криптомайнинга, так и дополнительные, такие, как изменение сложности сети и стоимость электроэнергии. В такой задаче особенно актуальными становятся модели машинного обучения, способные работать с большим количеством входных данных и учитывать сложные зависимости между ними.

В этой области особенно выделяются нейронные сети. Они обладают высоким уровнем точности и позволяют учитывать сложные зависимости между входными данными. Например, алгоритм с долгой краткосрочной памятью (англ. Long short-term memory - LSTM) используется для обработки длительных зависимостей, что позволяет ему работать с большим количеством данных.

При анализе доходности криптомайнинга необходимо учитывать работу пулов – сервисов, объединяющих вычислительные мощности различных пользователей, что позволяет им майнить криптовалюту более эффективно и получить доход. Однако такие сервисы не всегда учитывают все важные факторы, влияющие на доходность, как на рисунке 4.

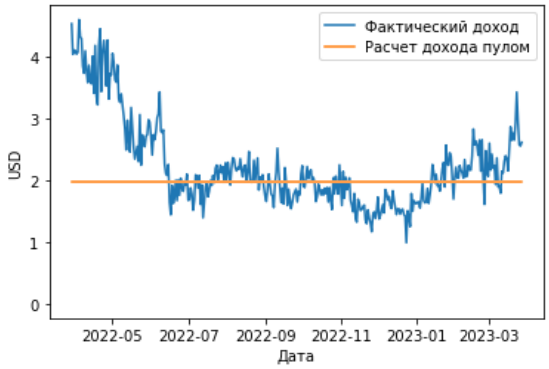


Рис. 4. График доходности криптовалюты Биткоин и ее линейный расчет пулом

В то время как пулы могут быть полезными для получения общей более-менее точной информации, исследователи должны иметь возможность предсказывать доходность на основе более комплексной модели. Использование LSTM для этой задачи позволяет учесть все важные факторы, включая сложность сети, вознаграждение за блоки, текущий курс криптовалюты и общую мощность сети. Подход, основанный на данном алгоритме, является оптимальным решением для предсказания доходности криптомайнинга в условиях высокой сложности сети и колебаний курсов криптовалют, результат которого отражен на рисунке 5.

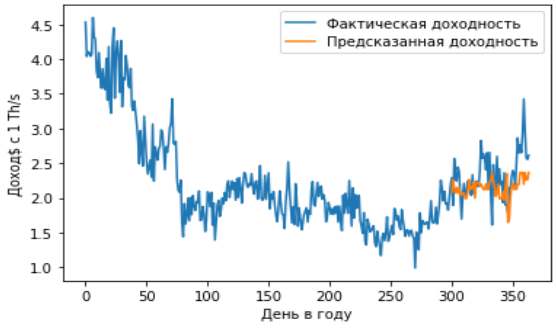


Рис. 5. График предсказания доходности криптовалюты Биткоин

Анализ криптомайнинга является сложным и многопараметрическим процессом, который включает сбор и предварительную обработку большого объема данных. Для оценки рентабельности майнинга и сравнения его с другими подходами к инвестированию необходимо учитывать все расходы и риски. Применение методов

машинного обучения, таких как регрессия, LSTM и случайный лес, позволяет предсказать курс криптовалюты и доходность криптомайнинга, учитывая различные показатели, такие как сложность сети и стоимость электроэнергии.

Развитие методов машинного обучения является перспективным направлением в области анализа криптомайнинга. Дальнейшие исследования могут быть направлены на улучшение и разработку более точных и надежных моделей прогнозирования доходности криптомайнинга на основе нейронных сетей. Также, важным направлением является оптимизация подбора оборудования для майнинга и технологии охлаждения для уменьшения издержек на электроэнергию.

Таким образом, использование методов машинного обучения является необходимым для более точного анализа криптомайнинга и принятия обоснованных инвестиционных решений. Более практические решения могут быть приняты на основе полученных результатов анализа, тем самым повышая рентабельность процесса криптомайнинга и принося больше доходов криптоинвесторам.

Литература

1. Виногоров Г.Г. Проблемы разработки методики анализа эффективности работы майнинг – фермы. // Белорусский государственный экономический университет. 2021. С. 441–445.
2. Luisanna Cocco. Modeling and Simulation of the Economics of Mining in the Bitcoin Market // Department of electrical and electronic engineering. 2016. P. 1-31.
3. Junwei Chen Analysis of Bitcoin Price Prediction Using Machine Learning // Journal of Risk and Financial Management. 2023. P. 1–9.

УДК 004.934

ОБНАРУЖЕНИЕ СЛЕДОВ АУДИОМОНТАЖА ПРИ СПУФИНГ-АТАКЕ НА СИСТЕМУ ГОЛОСОВОЙ БИОМЕТРИИ

Самородова М.Э.[1] (студент)
Научный руководитель – к.ф.-м.н., доцент, старший научный сотрудник Рыбин С.В.[1]
1 – Университет ИТМО
e-mail: samorodova@speechpro.com

Аннотация
В работе проведена оценка актуальности нестационарных спуфинг-атак на основе аудиомонтажа в традиционных системах детектирования логических спуфинг-атак, а также предложен способ подготовки данных данного типа. Рассмотренный способ обучения модели позволил добиться улучшения качества работы традиционной системы антиспуфинга в 11 раз по EER при предъявлении спуфинг-атак на основе аудиомонтажа. Тем не менее, одинарное сращивание остается сложным сценарием для детектирования, требующее дальнейшего исследования.

Ключевые слова
Голосоовой антиспуфинг, спуфинг-атака, голосовая биометрическая система, аудиомонтаж, равновероятная ошибка, partialspoof.

Расширение областей применения голосовых биометрических систем неминуемо влечет и увеличение попыток мошенничества. Под спуфинг-атакой подразумевается попытка взлома системы верификации с помощью фальсификации голосовых характеристик целевого диктора. Данные атаки стремятся сместить распределение оценок модели в сторону распределения подлинных дикторов, тем самым провоцируя увеличение частоты ложного срабатывания, то есть ошибку второго рода (False Acsept Rate, FAR). Поэтому зачастую система верификации по голосу используется в тандеме с системой антиспуфинга, задачей которой является предотвращение различных попыток мошенничества, осуществленных по средствам спуфинг-атак.

Классическая классификация спуфинг-атак подразумевает их деление на физические и логические [1]. Первые совершаются на уровне микрофона с помощью повторного воспроизведения записи целевого диктора. Вторые – на уровне передачи данных, к логическим атакам относят атаки синтезом и преобразованием голоса.

В последнее время также стали появляться работы по частичным спуфинг-атакам [2], которые осуществляются объединением подлинных высказываний целевого диктора и синтезированных фрагментов. Такие атаки можно отнести к нестационарным. Сложность их детектирования обусловлена политикой работы модели в тестовом режиме, при котором она предсказывает n оценок для одного тестового образца для повышения устойчивости к выбросам. То есть непродолжительные фрагменты синтеза будут оказывать мало влияния на итоговую оценку, что приведет к смещению её значения в сторону класса подлинной речи, провоцируя увеличение ошибки второго рода.

Однако в сообществе голосового антиспуфинга не была изучена проблема спуфинг-атак, осуществленных с помощью аудиомонтажа. Данный вид атаки также можно отнести к нестационарным, поскольку она получена на основе сращивания речевых фрагментов подлинных высказывания целевого диктора (см. рис. 1). Саму атаку будут характеризовать небольшие по продолжительности места «склеек» аудиофрагментов.

Потенциально атака на основе аудиомонтажа может представлять серьёзную угрозу для системы-антиспуфинга. В конкурсе ASVspoof2017 [3] рассматривались логические атаки повторным воспроизведением, которые очень плохо детектируются. Основной контрмерой по борьбе с данной атакой может являться использование для

верификации рандомизированной циферной последовательно или построение текстонезависимой системы с проверкой корректности высказывания. Однако данные контрмеры может помочь преодолеть аудиомонтаж, способный выполнить даже непрофессионал, используя открытые инструменты редактирования звука [4].

Рис. 1. Варианты спуфинг-атаки на основе аудиомонтажа, где *f1* – речевой фрагмент из первого аудиофайла, *f2* – речевой фрагмент из второго аудиофайла, *t*1 – общая продолжительность изготовленного поддельного образца

Таким образом, целью работы являлось исследование проблемы нестационарных спуфинг-атак, осуществленных с помощью аудиомонтажа подлинных высказываний целевого диктора.

Так как не существует наборов данных, в которых спуфинг-атаки осуществлены с помощью аудиомонтажа, первым этапом работы являлась подготовка данных. Генерация спуфинг-атак осуществлялась в трех сценариях: одинарное, двойное и многократное сращивание. Длина каждого фрагмента бралась случайным образом из нормального распределения. Потенциально наибольшую угрозу должны представлять одинарные «склейки», в то время как многократное сращивание будет являться наиболее вероятным сценарием в реальных условиях. Для повышения вариативности данных сращивание осуществлялось как высказываний, принадлежащих одному исходному аудиофайлу, так и разным. Помимо самого набора данных также была получена разметка по местам «склейки», в которой место сращивания помечалось единицами и характеризовалось небольшой продолжительностью в 200 мс.

После подготовки набора данных было обучено 5 моделей на основе WavLM [5] при различных политиках. Целью сравнения моделей было изучение влияния на качество работы моделей в тестовом режиме введения дополнительной информации на этапе обучения. В качестве дополнительной информации рассматривалось использование разметки и точных меток.

Все модели обучались не на полной продолжительности исходных высказываний, а на случайном фрагменте длиной в 3 секунды. Таким образом, при выборе случайного сегмента, могло так произойти, что сегмент поддельного образца не содержал мест сращивания. Наиболее вероятно это было в случае одинарной «склейки». В таком случае схожие образцы были отнесены к разным классам, что приводило к запутыванию модели.

Применение разметки на этапе обучения гарантировало, что поддельный образец содержал хотя бы одно место сращивания. Осуществлялось это следующим образом: из *n* мест сращивания принадлежащих одному аудиообразцу выбиралось одно случайным образом. Относительно этой точки, рассчитывалось, также случайным образом, отношение продолжительности до и после неё. Это позволяло повысить вариативность расположение места сращивания в итоговом сегменте общей продолжительностью в 3 секунды.

Под точными метками здесь подразумеваются метки, сопоставленные с каждым бином исходной фонограммы (см. рис. 2). То есть в случае образцов, принадлежащих классу подлинной речи, каждому бину была присвоена метка равная 0. В то время как для образцов, принадлежащих классу поддельной речи, местам сращивания аудиофрагментов была присвоена метка 1, остальным фрагментам – 0. Точные метки для спуфинг-атак были получены на основе разметки по местам сращивания. В итоге каждому образцу соответствовало *n* меток, по которым рассчитывалась отдельная функция потерь. Итоговые потери равнялись сумме потерь по точным и агрегированным меткам, которые соответствовали всему образцу, т.е. подлинный класс был равен 0, а поддельный – 1. Точные метки при этом могли были как мягкими, так и жесткими. Под жесткими здесь подразумеваются метки, которые могут принимать только два значения: 0 или 1. Мягкие метки изменялись плавно от 0 до 1.

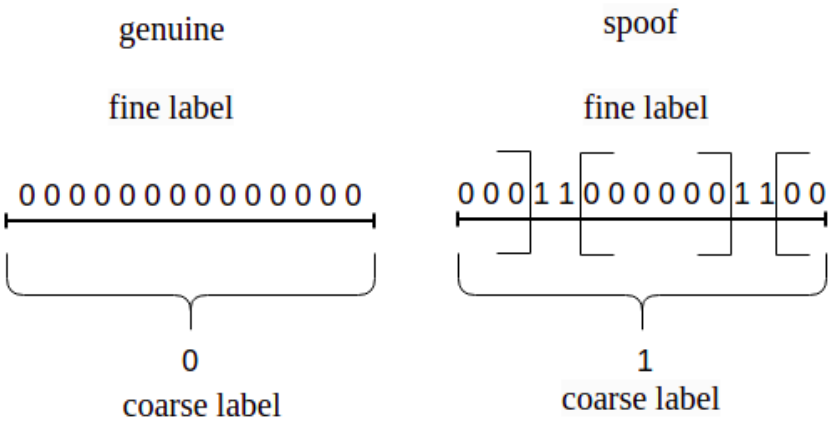


Рис. 2. Схематичное изображение присвоения точных меток аудиозаписи

В качестве базового решения была взята модель, обученная на подготовленных данных без дополнительной информации. Качество работы по EER обученных моделей на тестовом множестве подготовленных данных представлено в таблице 1. Здесь EER_1 рассчитан по агрегированным меткам, EER_2 – по точным меткам.

Таблица 1

Качество работы по EER на тестовом множестве подготовленных данных

модель	EER_1, %			EER_2, %		
	одинарное	двойное	много-кратное	одинарное	двойное	много-кратное
без разметки	5,5	2,8	1,9	27,7	17,7	8,2
разметка	4,1	2,6	1,3	17,2	10,7	5,4
разметка + точные мягкие метки	3,6	2,2	1,0	12,1	8,2	4,1
разметка + жесткие точные метки	3,1	1,7	1,0	12,1	8,9	4,4

Как видно из таблицы 1, добавление при обучении как разметки, так и точных меток позволило повысить качество работы систем в тестовом режиме относительно базового решения. Наименьшее значение EER продемонстрировала модель, обученная с

привлечением разметки и жестких точных меток. В среднем ее качество выше на 55 процентов относительно модели, обученной без дополнительной информации. Как и ожидалось ранее, самым сложным сценарием оказалось одинарное сращивание.

Для изучения проблемы масштабируемости системы антиспуфинга, к имеющимся подготовленным «склеечным» данным, были добавлены данные, содержащие логические спуфинг-атаки, полученные с помощью синтеза речи. Здесь было произведено в дополнение к рассмотренным ранее обучение 2 моделей. Целью данного ряда экспериментов являлось изучение эффективности детектирования «склеечных» атак традиционной системой антиспуфинга, обученной на логических атаках. Мотивация данного исследования обусловлена тем, что модель антиспуфинга должна быть универсальным детектором, способным справляться с различного рода атаками, поскольку в реальных условиях заведомо неизвестно, какой тип атаки применит злоумышленник. Качество работы по EER на агрегированных метках на двух тестовых множествах представлено в таблице 2. Здесь в качестве первого тестового множества выступают спуфинг-атаки на основе адуиомонтажа, а в качестве второго – на основе синтеза речи. Значения EER представлены по агрегированным меткам.

Таблица 2

Качество работы по EER на разных доменах

модель	EER_1, %				
	одинарное	двойное	многократное	аудиомонтаж	синтез речи
Синтез речи	56,7	55,9	61,6	58	0,1
Синтез + аудиомонтаж с разметкой и жесткими метками	7,7	6,1	2,0	5,2	0,13
Синтез + аудиомонтаж с разметкой и жесткими метками	3,1	1,7	1,0	1,9	43

Из таблицы 2 видно, что модели, обученные на одном домене, не способны детектировать данные из другого распределения. Совместное обучение, то есть добавление в обучение спуфинг-атак на основе аудиомонтажа и синтеза речи, позволило распознавать оба домена. При этом на спуфинг-атаках на основе аудиомонтажа наблюдается деградация качества относительно модели, обученной только на этом домене. Тем не менее, качество традиционной системы детектирования логических спуфинг-атак в среднем возросло в 11 раз при предъявлении нестационарных поддельных образцов.

Таким образом, в данном исследовании было показано, что спуфинг-атаки на основе сращивания подлинных высказываний целевого диктора представляют серьезную угрозу для традиционной системы детектирования логических спуфинг-атак. Добавление подобных данных в обучение позволяет добиться улучшения качества, в основном на частых «склейках», однако одиночное сращивание аудиофрагментов остается сложным сценарием. Для решения этой проблемы были предложены способы применения разметки о местоположении «склеек» на фонограмме при обучении системы, что в частных случаях привело к улучшению качества. Общая проблема

детектирования одиночных «склеек», а также кросс-доменного детектирования «склеек» не была решена и требует дальнейшей разработки.

Литература

1. Z. et al. Spoofing and countermeasures for speaker verification: A survey //speech communication. 2015. Т. 66. С. 130-153.
2. Zhang L. et al. The partialspoof database and countermeasures for the detection of short generated audio segments embedded in a speech utterance //arXiv preprint arXiv:2204.05177. 2022.
3. Kinnunen T. et al. The ASVspoof 2017 challenge: Assessing the limits of replay spoofing attack detection. 2017.
4. Audacity: Audacity. Free, open source, cross-platform audio software for multitrack recording and editing [Электронный ресурс]. URL: https://www.audacityteam.org/ (Дата обращения: 10.03.23)
5. Chen S. et al. Wavlm: Large-scale self-supervised pre-training for full stack speech processing //IEEE Journal of Selected Topics in Signal Processing. 2022. Т. 16. №. 6. С. 1505-1518.

УДК 004.8

ИСПОЛЬЗОВАНИЕ НЕЙРОСЕТЕВОГО МОДЕЛИРОВАНИЯ В СИСТЕМАХ КОНТРОЛЯ И УПРАВЛЕНИЯ ДОСТУПОМ НА ОСНОВЕ ЛИЦЕВОЙ БИОМЕТРИИ

Холькина А.А.[1] (студент)
Научный руководитель – к.э.н., доцент, зав. кафедрой ММЭУ Писарева О.М.[1]
1 – ФГБОУ ВО «Государственный университет управления»
e-mail: kholkina.a2021@gmail.com

Аннотация
При решении задачи повышения комплексной безопасности в настоящее время широко востребованы системы контроля и управления доступом на основе распознавания лиц. В работе представлен один из возможных подходов к разработке программного обеспечения подобной системы с применением нейросетевого моделирования. Результатом работы являются программные модули детекции и распознавания лиц, имеющие точность 0.98 и 0.87 соответственно.

Ключевые слова
Система контроля, управление доступом, биометрия, сверточные нейронные сети, распознавание образов, нейросетевое моделирование.

Проблема обеспечения контроля и безопасности на территориях с режимом ограниченного доступа, в частности, образовательных учреждений, всегда находится в фокусе интересов как администрации и клиентов подобных организаций, так и разработчиков систем контроля и управления доступом (далее - СКУД). В частности, за период с 2014 по 2021 год следственный комитет Российской Федерации расследовал 33 факта предотвращенных или реализованных вооруженных нападений на школы, колледжи и университеты. Очевидно, что менее трагичных событий, связанных с несанкционированным проникновением на объекты, существенно больше, также, как и причинённого от этого совокупного ущерба. В этой связи встает вопрос о повышении технико-эксплуатационных характеристик СКУД, и прежде всего их надежности, непосредственно направленных на рост вероятности предотвращения попыток проникновения на объекты ограниченного доступа лиц, не имеющих соответствующих прав. Одним из возможных подходов к решению этой задачи является установление систем контроля и управления доступом, проводящих идентификацию личности на основании данных биометрии.

По прогнозам компании RecFaces, в ближайший год отечественный рынок биометрических систем продолжит активный рост, а его среднегодовое значение превысит 60%, при этом особым спросом пользуются системы на основе лицевой биометрии: они более надежны по сравнению с аналогами, использующими в качестве ключей доступа отпечатки пальцев, голос и радужную оболочку глаза. Вероятность несанкционированного допуска (FAR) систем на основе лицевой биометрии составляет 0,1%, что является самым низким показателем [1]. Лидерами российского рынка в области разработки таких систем являются компании ISS, Sigur, ITV, Macroscop. Причем средняя удельная стоимость их лицензии на модуль распознавания лиц на одну камеру составляет 378 440 рублей при базе данных в 1 000 человек, т.е. существенность стоимости программного обеспечения может стать ощутимой проблемой для государственных образовательных учреждений.

Представим наиболее распространенные методы, используемые в настоящее время для распознавания лиц по их геометрическим признакам.

Метод сравнения с эталоном является одним из самых первых методов распознавания образов. Его суть заключается в выделении областей лица на

изображениях и их сравнении [2]. Для сравнения используются примитивные алгоритмы по типу попиксельного сравнения ключевых областей (глаз, губ и т.д.). Недостаток этого метода заключается в том, что он требует много ресурсов как для хранения участков областей, так и для их сравнения. В виду того, что используется простейший алгоритм сравнения, изображения должны быть сняты в строго установленных условиях: не допускается заметных изменений ракурса, освещения, эмоционального выражения и пр.

Метод анализа геометрических характеристик лица заключается в выделении набора ключевых точек (областей) лица и последующем выделении набора признаков. Каждый признак является либо расстоянием между ключевыми точками, либо отношением таких расстояний [2]. Ключевыми точками могут быть уголки глаз, губ, кончик носа, центр глаза и т.д. В качестве ключевых областей могут быть прямоугольные области, включающие в себя: глаза, нос, рот. Нахождение ключевых точек на изображении является трудоемкой задачей в процессе распознавания, и правильное определение этих точек играет ключевую роль в успешном распознавании. Этот метод требует строгого соблюдения условий съемки, а также надежного механизма для нахождения ключевых точек в общем случае. Для решения этой задачи необходимо применение более продвинутых методов классификации или создания модели изменений.

Нейронные сети являются универсальным алгоритмом решения самого широкого класса задач [2]. Они могут выполнять функции как извлечения признаков из изображений на локальном и глобальном уровнях, так и для их классификации. Нейронные сети можно использовать как в качестве классификаторов, выделяя признаки с помощью других методов, так и наоборот. Возможна также комбинация разных сетей для решения этих задач: одна для извлечения признаков, а другая для их классификации.

Пример использования нейронных сетей для разработки ПО для СКУД, можно найти в работе [1]. Здесь автор использует две предобученные на обнаружение лиц нейронных сети на архитектуре ResNet. Первая детектирует лицо, а вторая преобразует обнаруженную область в набор числовых данных, который сравнивается с наборами данных изображений лиц, содержащихся в базе. Преимуществами таких решения являются быстрота метода, облегчение модели по сравнению с аналогами (CNN, Haar Cascade) и хорошие результаты работы для фронтальных граней; к недостаткам можно отнести несовершенство алгоритма выделения области лица (могут вырезаться значимые части области лица) и возможность обмана алгоритма при помощи фотографий (печатного формата и с экрана телефона).

Также хочется выделить работу [3]. В ней описан алгоритм детектирования лиц на основе каскадного классификатора Хаара, а распознавания – на основе сиамской нейронной сети. Сиамская сеть является модифицированной версией предобученной сети MobilenetV3. Ключевая особенность таких архитектур заключается в отсутствии необходимости формирования большого датасета, поскольку принципом её работы является вычисление расстояния между наборами числовых признаков обнаруженной области лица и изображениями лиц из базы. Отмечено [3], что реализованная нейронная сеть достигла точности работы 90%.

Изучив работы по данной тематике, можно сделать вывод, что на практике нейронные сети набирают популярность, увеличивая точность распознавания. С учетом опыта коллег, описанного в [1, 3], автор работы предлагает некоторые доработки и альтернативные решения. В частности, дополнительное внимание уделялось обработке случаев, когда систему обманывают при помощи фотографий (печатных или с экрана телефона), а также проведению экспериментов по повышению точности распознавания.

Процедуру создания биометрической системы контроля доступа можно разбить на следующие этапы:

1) детектирование лица (процесс, обнаружения камерой лица в кадре);

2) предобработка изображения и анализ лица (коррекция изображения, определение геометрии лица на 2D-изображении и др.);

3) распознавание лица или поиск соответствия (сравнение преобразованного биометрического набора данных с эталонными в базе данных).

Иллюстрация процесса идентификации субъектов представлена на рисунке 1.

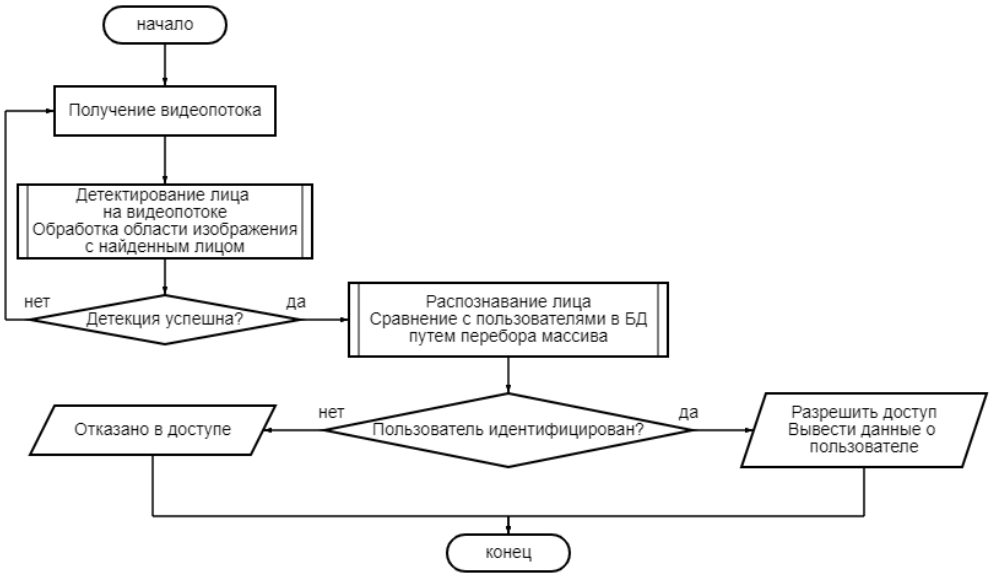


Рис. 1. Представление процесса идентификации субъектов

Согласно схеме, в ходе работы необходимо создать два основных программных модуля – детектор лиц и модуль для их дальнейшего распознавания.

При анализе подходов к решению задач, связанных с лицевой биометрией, перспективным направлением является нейросетевое моделирование: обладая высокой точностью, нейросетевые алгоритмы сохраняют быстродействие и способность адаптироваться к изменениям окружающей среды. В такой сложной задаче, как детекция лиц, нейронные сети также показывают отличные результаты, сохраняя при этом высокую скорость работы. Скорость работы сети зачастую зависит от типа ее архитектуры: она может быть одноэтапной (single-shot, proposal-free) или двухэтапной (two-stage, region-based). Одноэтапные архитектуры нейронных сетей работают по принципу передачи входных сигналов через несколько слоев нейронов, где каждый слой производит определенные математические операции с полученными данными, при этом данная архитектура не требует предобработки. На выходе получается результат классификации и выделения искомых прямоугольников. Двухэтапные нейронные сети состоят из двух частей: первая часть выделяет полезные признаки (feature extraction), а вторая часть на основе полученных признаков проводит финальную классификацию и уточнение искомых прямоугольников. Такая архитектура позволяет улучшить качество решения задачи, т.к. признаки, выделенные первой частью сети, являются более информативными и учитывают особенности входных данных. Одними из наиболее известных нейросетей детекции на основе одноэтапного метода являются YOLO и SSD, а на основе двухэтапного – Faster-RCNN, R-CNN, Fast-RCNN и др. В экспериментах, описанных в [4], приводится сравнение нейросетей детекции YOLO и Faster-RCNN. Обладая высокой скоростью работы, одноэтапная YOLO незначительно уступает своему двухэтапному конкуренту. Поскольку скорость работы является приоритетом при построении эффективной СКУД, было принято решение использовать именно одноэтапный алгоритм для создания детектора лиц.

Вторым шагом для решения поставленной задачи является разработка программного модуля для распознавания лиц (поиска соответствия в базе данных эталонов). Данный модуль должен быстро и качественно производить сравнение поступающих с модуля детекции 2D-изображений лица с эталонными изображениями, которые заносятся в базу данных при добавлении в нее новых пользователей. Перспективным решением для поиска соответствия являются сиамские нейронные сети.

Сиамская нейронная сеть – это тип сети, который используется для решения задач сравнения и классификации объектов. Она состоит из двух идентичных подсетей, каждая из которых обрабатывает одно изображение и извлекает из него векторы признаков. Далее эти векторы сравниваются и рассчитывается мера сходства объектов. Для этого часто используется косинусное расстояние, евклидово расстояние или манхэттенское расстояние.

Несмотря на потенциальную полезность, сиамские нейронные сети не получили широкого распространения по нескольким причинам:

1) реализация сиамских нейронных сетей может быть сложной и требует определенных знаний и умений;
2) необходимость большого количества данных для обучения;
3) ограниченная область применения.

На данный момент готовые решения найти крайне сложно, и, в связи с этим, автором предпринята попытка разработать собственную сиамскую нейронную сеть. Остановимся подробнее на описании этапов реализации поставленной цели исследования. Для создания модуля детекции для СКУД были определены три класса, которые необходимо классифицировать и определять их местоположение:

1) лицо;
2) лицо в маске;
3) фейк.

При обнаружении класса «лицо», обработанное изображение передается непосредственно модулю распознавания. Класс «лицо в маске» будет уведомлять входящего человека о том, что для корректного распознавания маску нужно снять. Класс «фейк» необходим для тех случаев, когда входящий человек показывает на камеру фотографию человека, имеющего доступ на объект, пытаясь обмануть систему. Борьба с такими случаями называется антиспуфинг и является важным направлением обеспечения безопасности для систем распознавания. При обнаружении этого класса сотрудникам охраны будет поступать сообщение о возможном проникновении.

Нейронная сеть детекции в контексте решаемой задачи должна обладать высокой скоростью работы, поэтому приоритет был отдан одноэтапной нейронной сети YOLO (You Only Look Once). Эта сеть зарекомендовала себя хорошими показателями по точности детекции, множеством вариаций архитектуры, влияющими на скорость работы, режимом детекции в реальном времени, удобным программным инструментарием переобучения под разные спектры задач. Архитектуру YOLO можно разделить на две части:

1) backbone (основа) – часть сети, извлекающая признаки из изображения, используя сверточные слои;
2) head (голова) – часть сети, использующая извлеченные признаки для обнаружения объектов и их классификации.

Принцип работы YOLO заключается в том, что на каждое изображение накладывается детектирующая сетка, каждая ячейка которой отвечает за обнаружение объектов. Каждая ячейка предсказывает несколько рамок (bounding boxes), которые могут содержать объекты, а также вероятность того, что объект присутствует в каждой рамке. Каждая рамка также предсказывает вероятность принадлежности объекта к каждому из заранее определенных классов.

Функция потерь, используемая для обучения YOLO, представляет собой комбинацию трех компонентов:

1) box loss (потеря локализации) – измеряет ошибку между предсказанными ограничительными рамками и истинными ограничительными рамками;

2) objective loss (потеря уверенности) – измеряет ошибку в предсказанной оценке уверенности ограничивающих ячеек;

3) classification loss (потеря классификации) – измеряет ошибку в предсказанной вероятности класса для каждого ограничивающего поля.

Окончательная функция потерь представляет собой взвешенную сумму всех трех компонентов.

Для эффективной работы нейросети детекции необходимо иметь обучающую выборку, которая должна соответствовать определенным требованиям, в частности:

1) достаточное количество примеров каждого класса объектов и их сбалансированность (обеспечивает нейросети возможность четко различать объекты на изображениях);

2) разнообразие примеров каждого класса (позволяет нейросети обучаться на различных вариантах представления объектов, таких как размеры, положение на изображении, освещение и т.д.);

3) наличие аннотаций для каждого объекта на изображении (позволяет нейросети научиться распознавать и отделять объекты друг от друга);

4) высокое качество представленных аннотаций (гарантирует правильность обнаружения нейросетью объектов на изображениях);

5) достаточно большой объем выборки (гарантия высокой точности детекции сети).

Основой формирования обучающей выборки послужили данные (изображения) платформы RoboFlow, содержащей большое количество открытых датасетов под разные типы задачи. Автором был собран набор из 12 000 изображений, где 80% - тренировочные, 20% - валидационные. На рисунке 2 представлены изображения из датасета, где рисунок 2. а представляет класс “лицо”, 2. б – класс “лицо в маске”, 2. в – класс “фейк”.

а *б* *в*

Рис. 2. Иллюстрация изображений классов в датасете

В качестве архитектуры была выбрана YOLOv5s. Это довольно компактная модель, позволяющая детектировать объекты в режиме реального времени. В качестве начальных весов были взяты веса модели, предобученной на наборе данных COCO. Дообучение модели на новой обучающей выборке длилось 40 эпох, при этом изменение значений функции ошибки на тренировочных и валидационных данных представлено на рисунке 3.

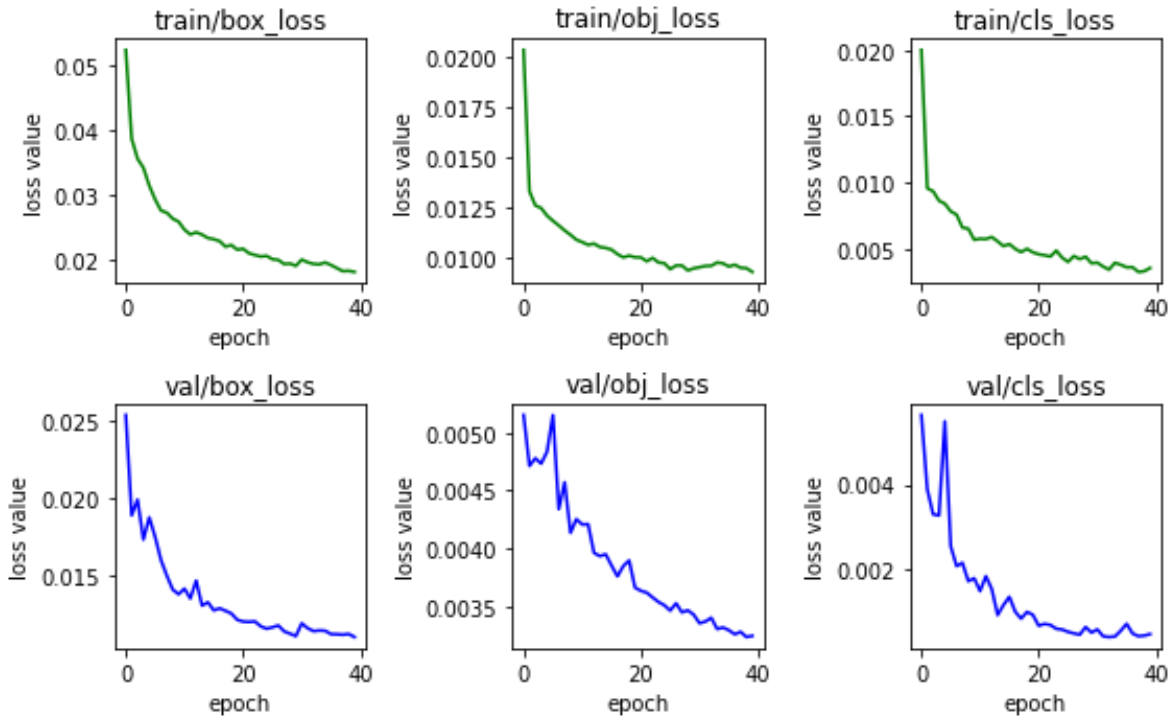


Рис. 3. Графики изменения значения функции ошибки в процессе дообучения

Кроме того, в ходе исследования рассчитывались следующие дополнительные метрики качества распознавания (см. рис. 4):

1) precision (точность) – доля объектов, выделенных классификатором как положительные и при этом действительно являющимися положительными;

2) recall (полнота) – отношение верно угаданных объектов класса ко всем представителям этого класса;

3) mAP_0.5 – средняя точность (mAP) при пороге IoU (пересечение через объединение) равном 0,5;

4) mAP_0.5:0.95 – средняя точность по различным пороговым значениям IoU, варьирующимся от 0,5 до 0,95. В частности, различие метрик mAP_0.5 и mAP_0.5:0.95 состоит в том, что последняя предъявляет более жесткие требования к расположению ограничивающих рамок обнаруженных объектов).

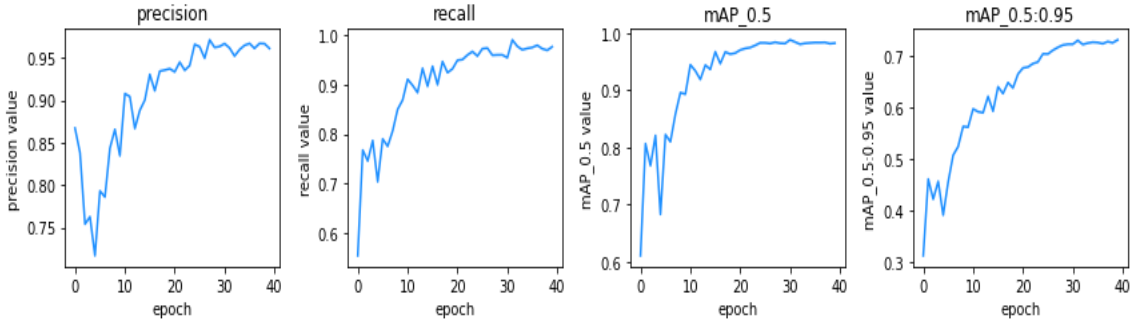


Рис. 4. Графики изменения метрик качества в процессе обучения

Также рассмотрим несколько примеров работы детектора на тестовых изображениях, взятых из сети Интернет (см. рис. 5).

Рис. 5. Иллюстрации процедуры детектирования на тестовых изображениях

Как можно заметить, модель правильно классифицировала и локализовала классы, которым обучалась.

Базовый модуль этапа распознавания лиц предложено разрабатывать на основе сиамских нейронных сетей. Как говорилось ранее, сиамская нейронная сеть – это искусственная нейронная сеть, которая использует одинаковые веса, работая в тандеме с двумя входными изображениями для вычисления сравнимых выходных векторов [5]. Затем полученные векторы признаков сравниваются путем вычисления расстояния между ними. Если на изображениях достаточно схожие объекты, то сеть должна выучить параметры, т.е. веса и смещения таким образом, чтобы получить меньшее расстояние между двумя изображениями, а если объекты разные, то расстояние должно быть больше. В отличие от обычной сверточной нейронной сети, сиамская сеть не классифицирует изображения по определенным категориям или меткам, а только определяет расстояние между любыми двумя изображениями.

В ходе работы было разработано три архитектуры нейронной сети. Первая нейронная сеть была довольно простой и состояла из 13 слоев. В них входили свертки, макспуллинг, линейные преобразования и функции активации ReLU.

Вторая и третья архитектуры базировались на более сложном подходе, который заключается в использовании ранее обученных нейронных сетях классификации, к выходам которых добавляется несколько полносвязных слоев. Использование сложных предобученных моделей позволяет извлекать из изображений более полезные векторы признаков, а также сократить время обучения и размер набора данных. Для второй модификации архитектуры в качестве предобученной модели использовалась глубокая архитектура ResNet50, для третьей - более компактная Xception.

Полученные векторы признаков подавались небольшой нейронной полносвязной сети, состоящей из двух слоев линейных преобразований с функцией активации ReLU, батч-нормализацией и L2-нормализацией.

Между окончательными векторами признаков во всех архитектурах вычисляется евклидово расстояние, по которому оценивается функция потерь.

Для первой модификации архитектуры в качестве функции ошибки была выбрана контрастная функция ошибки (contrastive loss, L). Её цель - минимизация расстояния между векторами признаков двух изображений одного класса (positive pairs) и максимизация расстояния между векторами признаков изображений двух разных классов (negative pairs). Она оценивается следующим образом:

$$L = (1 - Y) * \left\|x_i - x_j\right\|^2 + Y * \max\left(0, m - \left\|x_i - x_j\right\|^2\right),$$

где m – гиперпараметр, определяющий нижнее граничное расстояние между различающимися изображениями; Y – истинный ответ.

Для второй и третьей архитектур использовалась триплетная функция ошибки (triplet loss). Она используется для обучения модели, которая может находить различия между тремя изображениями: якорным, положительным и отрицательным, в частности:

1) якорное изображение — то, к которому надо найти похожие изображения;
2) положительное изображение — то, которое считается похожим на якорное;
3) отрицательное изображение — то, которое не похоже на якорное изображение.

Тройка изображений (якорное, положительное и отрицательное) называется триплетом. Каждое изображение триплета проходит через нейронную сеть, на выходе которой оно преобразуется в вектор признаков. Далее вычисляется оценка ошибки по следующей формуле:

$$L = \max(0, \|x - x^+\|^2 - \|x - x^-\|^2 + m),$$

где x^+ - вектор признаков положительного изображения; x^- - вектор признаков отрицательного изображения; m - гиперпараметр, определяющий нижнее граничное расстояние между различающимися изображениями.

Цель триплетной функции ошибки - минимизировать расстояние между якорным и положительным изображениями и максимизировать расстояние между якорным и отрицательным изображениями.

Обучающая выборка для модуля распознавания лиц была найдена в открытом доступе сети Интернет. Она содержит в себе фотографии знаменитостей и является частью набора данных «Labeled Faces in the Wild». Оригинальные изображения имеют размер 250 x 250 пикселей. Набор данных содержит 1324 каталога, каждый из которых представляет одну знаменитость. Каждый каталог содержит 2-50 изображений.

На рисунке 6 представлены несколько фотографий из набора данных в виде триплета, который использовался для обучения второй и третьей архитектур, где первое изображение – якорное, второе – положительное, а третье – отрицательное.

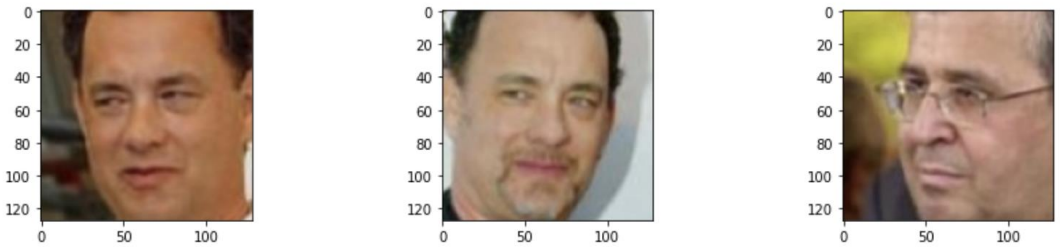


Рис. 6. Иллюстрация примера изображений в триплете

Обучающая и валидационная выборки были поделены автором в соотношении 9:1. Обучение каждой из трех нейронных сетей длилось 50 эпох. В таблице представлены метрика Accuracy по каждой архитектуре в конце обучения.

Таблица

Результаты обучения сиамских сетей

Архитектура	Accuracy
13-ти слойная архитектура	0,63
Resnet50 + полносвязная сеть	0,77
Xception + полносвязная сеть	0,94

Наилучшие результаты во время обучения показала третья нейронная сеть. На рисунке 7 представлена схема её обучения и информация об архитектуре.

Рис. 7. Укрупненная схема обучения модели

Наилучшие результаты во время обучения показала третья нейронная сеть, ее графики изменения функции ошибки и метрики Accucary во время обучения представлены на рисунке 8.

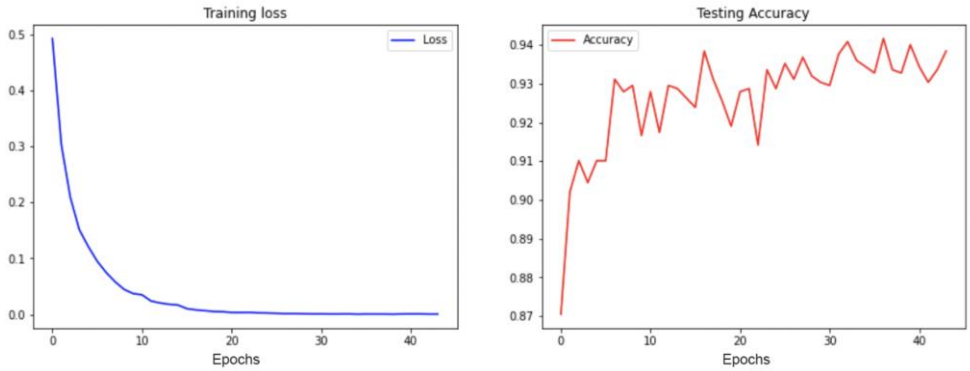


Рис. 8. Графики изменения функции ошибки и метрики Accuracy

Поскольку данная нейронная сеть не выдает на выходе вероятность принадлежности к определенному классу, а лишь вычисляет евклидово расстояние между двумя векторами признаков, для использования модели необходимо определить порог отсечения. Так, если расстояние между векторами меньше порога отсечения, можно считать, что на двух изображениях один человек, иначе – два разных. После варьирования порога отсечения было установлено, что наилучшим порогом отсечения для получившейся нейронной сети является значение 1,3. При таком пороге отсечения матрица ошибок на валидационном наборе данных выглядит так как она представлена на рисунке 9.

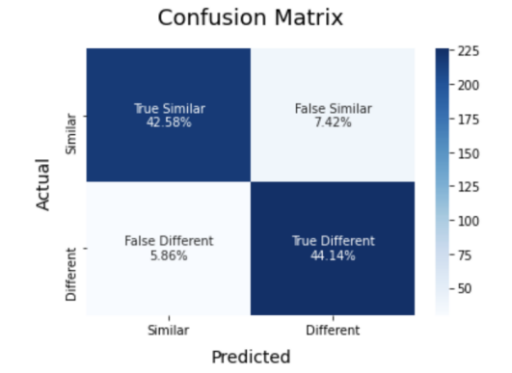


Рис. 9. Матрица ошибок сиамской сети на архитектуре Xception
и полносвязной сети на валидационном наборе данных

Из матрицы ошибок видно, что ложноположительно нейронная сеть классифицировала 7,42% субъектов, ложноотрицательно 5,86%, а 86,72% классифицировала верно. Из этого можно сделать вывод, что нейронная сеть с приемлемым качеством обучилась выполнять задачу распознавания лиц. Для дальнейшего улучшения качества работы сети необходимо продолжить обучения и расширить набор данных.

В ходе решения задачи повышения эффективности работы СКУД на основе методологии нейросетевого моделирования было создано два программных модуля детекции и распознавания лиц.

Таким образом, а процессе создания модуля детекции автором были проанализированы основные архитектуры сверточных нейронных сетей, был подобран набор данных для корректной работы при ношении медицинских масок и предотвращения случаев спуфинга, использовался метод трансферного обучения для снижения временных затрат. Результатом разработки данного модуля стала модель, способная выполнять задачу детекции лиц в режиме реального времени. За 40 эпох обучения метрики precision, recall и mAP@50 достигли значений 0.96, 0.97 и 0.98 соответственно. Модуль распознавания лиц был создан на основе сиамской нейронной сети. Были проведены эксперименты с различными архитектурами и функциями ошибки (триплетной и контрастной), по итогам которых лучшей стала модель на основе совмещения предобученной нейросети Xception и полносвязных слоев, точность ее работы составляет 0,87.

Исходный код программы и модели размещены в приватном репозитории по ссылке https://github.com/AnnaKholkina/detect_and_ident, доступ к которому предоставляется по запросу.

Автор статьи выражает искреннюю благодарность своему научному руководителю Писаревой Ольге Михайловне за ценные советы при планировании исследования и рекомендации по оформлению статьи.

Литература

1. Разработка интеллектуальной системы биометрической идентификации пользователя / Девицына С.Н., Елецкая Т.А., Балабанова Т.Н., Н.Н. Гахова // Экономика. Информатика. 2019. Т. 46. № 1. С. 148-160.
2. Программный комплекс распознавания лица человека / Ахметзянов К.Р., Сазонов В.И., Липин Ю.Н., Южаков А.А. // Вестник Уфимского государственного авиационного технического университета. 2017. Т. 21. № 3. С. 79-86.
3. Антипова С.А. Разработка системы контроля доступа на основе распознавания лиц / Антипова С.А. // Программные продукты и системы. 2021. Т. 34. № 2. С. 245-256.
4. Тимошкин М.С. Сравнение YOLOv5 и Faster R-CNN для обнаружения людей на изображении в потоковом режиме / Тимошкин М.С., Миронов А.Н., Леонтьев А.С. // Международный научно-исследовательский журнал. 2022. № 6 (120). URL: https:// research- journal.org/archive/ 6-120-2022-june/sravnenie-yolov5-i -fasterr-cnn-dlya-obnaruzheniya- lyudej-na- izobrazhenii-v-potokovom-rezhime (дата обращения: 12.12.2022).
5. Chicco Davide Siamese Neural Networks: An Overview / Davide Chicco // Methods in molecular biology: электронный журнал. URL: https: // pubmed. ncbi. nlm. nih. gov/ 32804361/. (дата обращения: 12.12.2022).

УДК 004.622

СИСТЕМА ГЕОПРОСТРАНСТВЕННОЙ ПРИВЯЗКИ И ВИЗУАЛИЗАЦИИ ДАННЫХ ИЗ ИНФОРМАЦИОННО-НОВОСТНЫХ ИСТОЧНИКОВ

Чубур С.В.[1] (студент), **Сабиров Т.Р.**[1] (старший преподаватель)
Научный руководитель – к.т.н. Сабиров Т.Р.[1]
1 – Военно-космическая академия имени А.Ф. Можайского
e-mail: chsemenv@mail.ru

Аннотация
В работе рассмотрен подход к разработке программного комплекса, обеспечивающего повышение эффективности деятельности информационно-аналитических служб, способствующего повышению уровня ситуационной осведомленности и предоставляющего возможность выявления закономерностей и сокращения времени на принятие правильных решений.

Ключевые слова
Геопространственная привязка, обработка естественного языка, аналитическая деятельность, ситуационная осведомленность, выявление закономерностей, улучшение экономических показателей.

Объем цифровой информации в современном мире увеличивается в стремительно ускоряющемся темпе. До 95% поток цифровой информации состоит из неструктурированных данных и лишь 5% составляют различные базы данных, представляющие собой тем или иным образом структурированную информацию [1]. Особенности работы аналитика в настоящих реалиях обязывают его обрабатывать большой объем информации в короткие сроки. Традиционные прикладные аналитические инструменты в настоящее время являются недостаточно эффективными. Разработка программного комплекса, обеспечивающего автоматизированный сбор, хранение, фильтрацию, геопространственную привязку и визуализацию данных из информационно-новостных источников, позволит повысить эффективность работы множества служб и компаний. Создание инструмента для улучшения качества работы информационно-аналитических служб актуально в связи с необходимостью повышения уровня обеспечения безопасности граждан, окружающей среды, улучшения экономических показателей компаний и предприятий, совершенствования механизма изучения обстановки и принятия решений.

Программный комплекс должен решать следующие типы задач:

1. Автоматизированный сбор и накопление большого объема информации из различных информационно-новостных источников.
2. Агрегация информационно-новостных сообщений по географическому положению и характерным особенностям.
3. Фильтрация информационно-новостных сообщений с целью уменьшения общего объема данных без потери смысловой нагрузки.
4. Осуществление геопространственной привязки и визуализации данных, полученных в результате выполнения предшествующих мероприятий.

Рассмотрим способы решения поставленных задач.

Задача автоматизированного сбора и накопления большого объема информации из различных информационно-новостных источников.

Хранение данных за прошедший период позволит использовать имеющуюся информацию для выявления закономерностей и прогнозирования вариантов дальнейшего развития событий.

В качестве источников данных могут быть использованы информационно-новостные сайты, социальные сети и мессенджеры. Использования большого количества

различных источников предоставляет возможность полного и всестороннего изучения обстановки.

Для автоматизированного сбора информации можно использовать технологию парсинга (синтаксического анализа).

Синтаксический анализ (англ. parsing) — процесс сопоставления линейной последовательности лексем (слов, токенов) естественного или формального языка с его формальной грамматикой. Результатом будет являться дерево разбора (синтаксическое дерево). Принцип работы парсера представлен на рисунке 1.

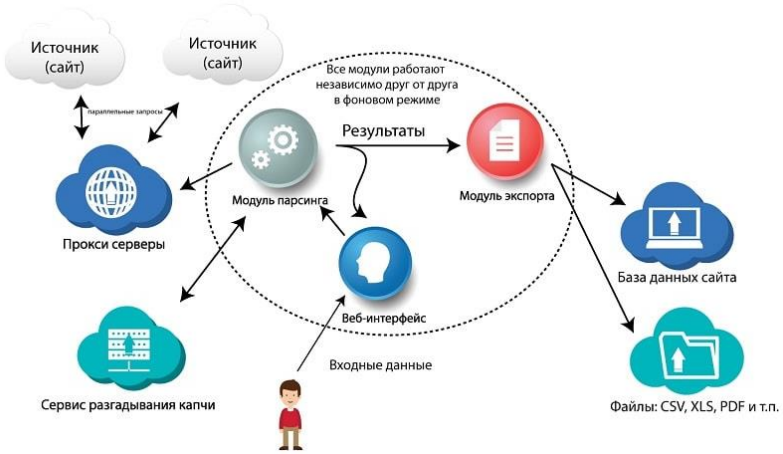


Рис. 1. Схема работы парсера

Для хранения данных целесообразно использовать компактную встраиваемую систему управления базами данных (СУБД) SQLite, представляющую собой библиотеку, с которой компонуется основная программа. В качестве протокола обмена целесообразно использовать вызовы функций библиотеки SQLite, это уменьшает время отклика и ускоряет работу программы. Данная СУБД хранит все данные, таблицы, определения и индексы в виде единственного файла на используемом компьютере.

Структура используемой таблицы базы данных будет иметь следующий вид (рис. 2).

tg_data_id	SSAGE_	SENDER	CHAT_TITLE	DATE	MESSAGE	ADV_MESSAGE
Фильтр	Филь...	Фильтр	Фильтр	Фильтр	Фильтр	Фильтр
1733	533401	https://t.me/avto999999555	RU АВТОРЫНОК ...	1679229883	Форд фокус 2 2009г ...	["\u0444\u043e\u0440\u0434", ...
1734	533390	https://t.me/avto999999555	RU АВТОРЫНОК ...	1679228247	Москва-Химки ...	["\u043c\u043e\u0441\u043a\u0432...
1735	533380	https://t.me/avto999999555	RU АВТОРЫНОК ...	1679227934	Mercedes-Benz GLE ...	["mercedesbez", "gle", "coupe", ...
1736	533375	https://t.me/avto999999555	RU АВТОРЫНОК ...	1679227843	Volkswagen поло 2019...	["volkswage", ...
1737	533364	https://t.me/avto999999555	RU АВТОРЫНОК ...	1679226327	Москва ...	["\u043c\u043e\u0441\u043a\u0432...

Рис. 2. Структура хранения данных

Список, полученный из основного текста информационно-новостного сообщения в результате использования алгоритмов обработки естественного языка, представлен в седьмом столбце таблицы. Так как SQLite поддерживает ограниченное количество типов данных, полученный список сериализуется и в виде строкового объекта.

Принцип взаимодействия приложения с базой данных представлен на рисунке 3.

В зависимости от интересов пользователя возможно использование клиент-серверной СУБД PostgreSQL. В таком случае все клиентские запросы на обработку данных обрабатываются клиент-серверной СУБД централизованно.

Задача агрегации информационно-новостных сообщений по географическому положению и характерным особенностям.

Компоновка по определенным признакам неструктурированных данных большого

объема значительно снижает уровень энтропии ситуационной обстановки, тем самым обеспечивает возможность более детального изучения условий изменения настоящего положения [2].

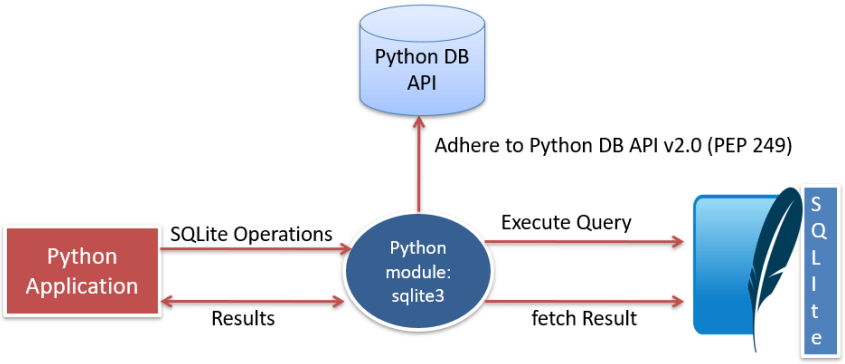


Рис. 3. Схема взаимодействия приложения с базой данных

В программном комплексе необходимо реализовать возможность агрегации информационно-новостных сообщений по географическому положению, а также по ключевым словам.

Для осуществления группировки информационно-новостных сообщений по географическому положению используется JSON-файл, содержащий информацию о названиях населенных пунктов и их координатах. В результате десериализации данного файла для каждого населенного пункта формируется кортеж с его координатами, что делает возможным осуществление агрегации данных по географическому положению, что продемонстрировано на рисунке 4.

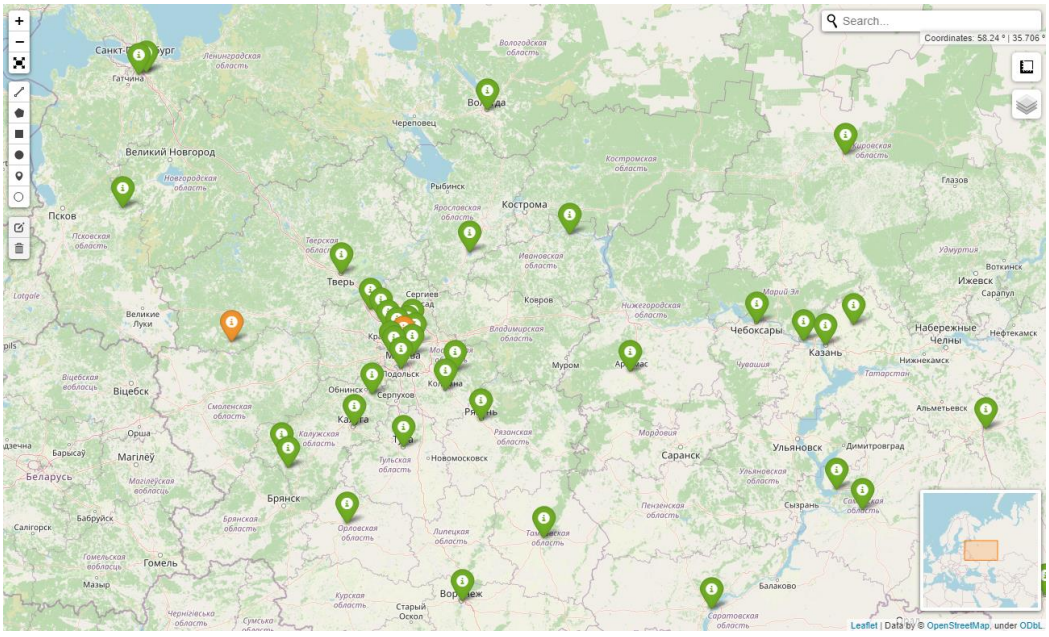


Рис. 4. Компоновка данных по географическому положению

Задача фильтрации информационно-новостных сообщений с целью уменьшения общего объема данных без потери смысловой нагрузки.

Использование алгоритмов нечеткого сравнения информационно-новостных сообщений по семантическому признаку позволяет уменьшить количество схожих и повторяющихся данных, тем самым повысить уникальность отображаемой информации [3].

В проекте для фильтрации набора информационных сообщений возможно использование модуля fuzz библиотеки FuzzyWuzzy. С помощью данной библиотеки осуществляется поочередное нечеткое сравнение информационных сообщений.

В основе метода нечеткого сравнения лежит определение расстояния Левенштейна. Расстояние Левенштейна (редакционное расстояние, дистанция редактирования) – метрика, измеряющая по модулю разность между двумя последовательностями символов. Она определяется как минимальное количество односимвольных операций (а именно вставки, удаления, замены), необходимых для превращения одной последовательности символов в другую. С использованием нескольких методов библиотеки FuzzyWuzzy осуществляется комплексный анализ и сопоставление переменных строкового типа.

Если значение, полученное в результате применения функции fuzz.token_set_ratio выше, чем параметр уникальности, указанный пользователем, то данное сообщение отмечается на интерактивной карте и попадает в текстовый отчет.

Задача осуществления геопространственной привязки и визуализации данных, полученных в результате выполнения предшествующих мероприятий.

Геопространственная привязка агрегированных по определенному признаку и отфильтрованных данных делает возможным визуализацию сложившейся обстановки. Сопоставление конкретных событий с их положением на местности открывает широкий спектр возможностей в аналитической деятельности.

Визуализация данных осуществляется путем создания интерактивной карты, что способствует наглядному представлению ситуационной обстановки за конкретный временной интервал.

Создание интерактивной карты можно осуществить с использованием библиотеки folium. Данная библиотека позволяет создать карту местности по координатам. Имеется возможность создания одиночных маркеров, кластеров маркеров, прокладывания маршрутов, наложения на карту геометрических фигур, масштабирования карты. Библиотека имеет ряд встроенных наборов фрагментов из OpenStreetMap, Mapbox и Stamen и поддерживает настраиваемые наборы фрагментов с ключами API Mapbox или Cloudmade. Folium поддерживает наложения как GeoJSON, так и TopoJSON, а также привязку данных к этим наложениям для создания карт с цветовыми схемами. Имеется возможность использования множества дополнительных плагинов для измерения расстояний и площадей, добавления новых меток и геометрических объектов. На рисунке 5 продемонстрировано основное окно приложения.

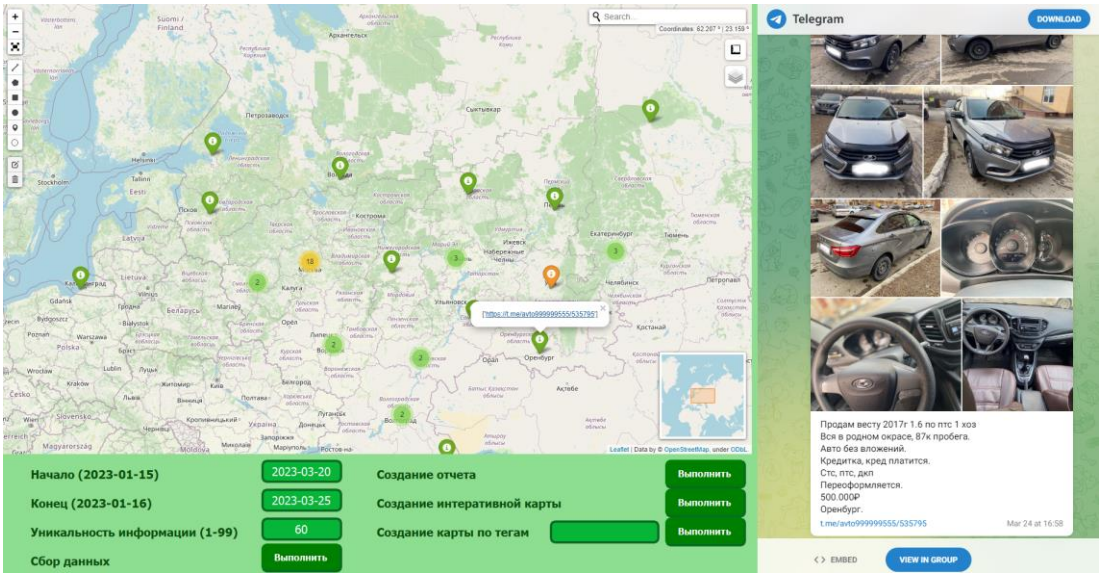


Рис. 5. Пользовательский интерфейс

Осуществление геопространственной привязки информационно-новостных сообщений к точкам на местности осуществляется с применением технологий обработки естественного языка NLP (Natural language processing).

Основной текст информационно-новостного сообщения подлежит следующим поэтапным преобразованиям (рис. 6):

1) удаление всех не цифробуквенных символов;
2) токенизация текста, то есть разделение на отдельные слова;
3) удаление стоп-слов, к которым относятся артикли, союзы, междометия и т.д.;
4) перевод всех символов токенизированных слов в нижний регистр;
5) лематизация токенов – преобразование каждого слова к его начальной форме с использованием морфологического анализатора PyMorphy2.

Использование морфологического анализатора позволяет осуществить синтаксический анализ данных для сопоставления информационно-новостного сообщения с его географическим положением [4].

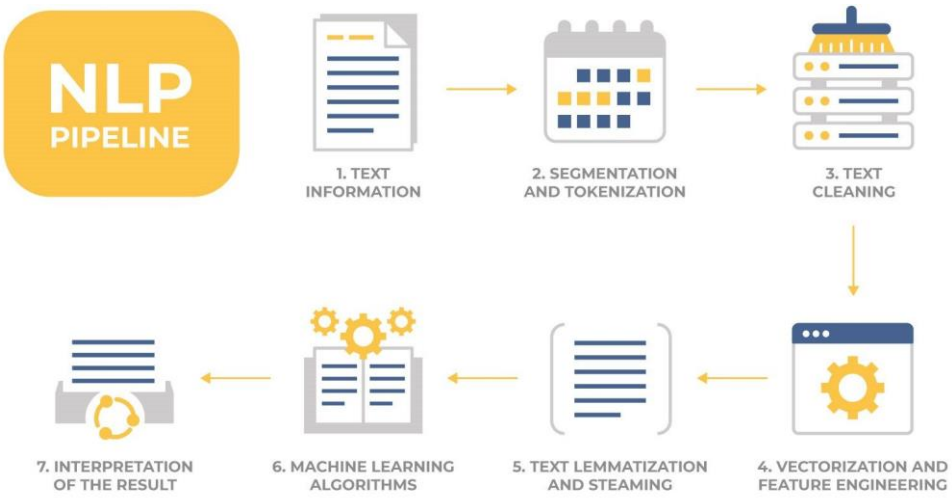


Рис. 6. Этапы обработки естественного языка

Введение в эксплуатацию разработанного программного комплекса обеспечит повышение эффективности деятельности информационно-аналитических служб, существенно повысит уровень ситуационной осведомленности, предоставит возможность выявления закономерностей и сократит время на принятие правильных решений. В зависимости от требований пользователя и преследуемых интересов система может быть адаптирована для использования как в разных информационных сферах, так и на разных по площади участках местности (район, город, регион, страна).

Предложенное решение может быть применено в различных областях здравоохранения, обеспечения безопасности граждан, повышения уровня правопорядка, улучшения экономических показателей деятельности компаний и предприятий разной направленности.

Дальнейшее развитие программного комплекса предусматривает совершенствование алгоритмов сбора, обработки, фильтрации информационно-новостных сообщений, а также развитие инструмента визуализации данных.

Литература

1. Информационно технический взрыв и его основные этапы, информационная перегрузка [Электронный ресурс]. Режим доступа: https: //leally.ru/how-to-open-file/ informacionno- tehnicheskii- vzryv- i-ego- osnovnye -etapy/ (дата обращения: 10.02.2023).

2. Методы ситуационного анализа и графической визуализации потоков больших данных [Электронный ресурс]. Режим доступа: https:// cyberleninka.ru/ article/n/ metody- situatsionnogo- analiza-i-graficheskoy- vizualizatsii- potokov-bolshih-dannyh (дата обращения: 10.02.2023).
3. Подход к устранению избыточности семантически близкой текстовой информации в хранилищах и базах данных [Электронный ресурс]. Режим доступа: https:// cyberleninka.ru/ article/n/podhod-k- ustraneniyu- izbytochnosti- semanticheski-blizkoy-tekstovoy-informatsii-v-hranilischah-i-bazah-dannyh (дата обращения: 10.02.2023).
4. Обработка естественного языка [Электронный ресурс]. Режим доступа: https:// neerc. ifmo.ru/wiki/ index.php? title=Обработка_естественного_языка (дата обращения: 10.02.2023).

УДК 004.93

УСКОРЕНИЕ ИНФЕРЕНСА НЕЙРОННЫХ СЕТЕЙ ПРЯМОГО РАСПРОСТРАНЕНИЯ

Шеметов Ф.А.[1] (студент)
Научный руководитель – к.ф.-м.н., доцент Рыбин С.В.[1]
1 – Университет ИТМО
e-mail: philippitemail@gmail.com

Аннотация
В данной работе исследованы методы и подходы к оптимизации времени прямого прохода на CPU при использовании одного и нескольких потоков для различных фреймворков (PyTorch, ONNX Runtime, OpenVino, Tensorflow) глубокого обучения. Были проведены эксперименты с различными настройками сборки, алгоритмами и методами изменения размеров многомерных матриц, обрабатываемых процессором.

Ключевые слова
Нейронные сети прямого распространения, оптимизация, прямой проход, фреймворки, ONNX Runtime, LibTorch, Tensorflow, многопоточность.

Прямой проход (англ. inference) – действие, когда на обученную модель нейронной сети подается набор неизвестных данных, происходит их вычисление и в результате выдается прогноз, основанный на точности предсказания нейронной сети. Фреймворк (англ. Framework) – это набор инструментов и библиотек, который предоставляет возможность разработки, обучения и валидации нейронной сети с помощью высокоуровневого интерфейса. Неоптимизированный прямой проход нейронной сети может занимать не мало времени на большом наборе данных чтобы спрогнозировать результат, из-за чего конечный результат не является эффективным. Как пример можно взять камеры с нейронными сетями в различных областях, где на вход в одну минуту подаётся огромное количество кадров и требуется быстрый прямой проход для получения прогноза в каждом кадре. Целью данной работы является оптимизация времени прямого прохода нейронных сетей. Чтобы выявить потенциальные методы оптимизации прямого прохода нейронных сетей прямого распространения, важно изучить существующие научные работы и исследования, посвященные данной тематике. Анализ сторонних исследований позволит определить наиболее перспективные и эффективные подходы, а также учесть преимущества и ограничения каждого метода. В работе [1] автор описывает оптимизацию времени прямого прохода при использовании фреймворка OpenVino. Преимущества этого подхода включают ускорение за счет пакетной обработки, параллелизма на уровне сети, упрощения структуры сети с использованием инструмента OpenVino Model Optimizer и уменьшения разрядности вычислений. В работе [2] авторы исследуют оптимизацию времени прямого прохода для сверточных нейронных сетей, используя последние возможности процессоров и оптимизации на уровне вычислительного графа. Подход обеспечивает лучшую производительность при выполнении сверточных операций в одном потоке и создает собственный пул потоков для эффективного распараллеливания операций. Основным недостатком может быть сложность реализации данного подхода и потребность в специальных механизмах для эффективного параллелизма. В работе [3] авторы описывают возможные оптимизации для MXNet, PyTorch и Tensorflow. Для MXNet используется опция сборки MKL, для PyTorch используется поддержка LAPACK установкой библиотеки magma-cuda90, а для Tensorflow используется пакет Python, подготовленный TinyMind. Недостатком подхода для Tensorflow является то, что оптимизация пакета происходит к обучению модели, а не к прямому проходу. Все три

подхода имеют свои особенности и выбор оптимального подхода будет зависеть от используемого фреймворка, аппаратных возможностей и конкретных задач.

Для исследования скорости работы по времени каждого фреймворка была взята предобученная PyTorch модель IResNet-50 [4] на архитектуре IResNet [5], так как бейзлайн (англ. baseline) на момент начала исследований являлся популярным для распознавания лиц. Чтобы получить метрики в виде среднего значения и медианы времени работы прямого прохода, был взят набор данных ImageNet на 1000 изображений для подачи на модель. Эксперимент повторяется 10 раз для получения более точных метрик.

Основной схемой работы системы является использование SDK с предобученными моделями и фреймворками. При сборке SDK, фреймворки настраиваются в соответствии с заданными параметрами для оптимизированной производительности прямого прохода. Система также включает сервер, который осуществляет прямой проход, используя предобученную модель из SDK и собранный фреймворк. Данные для прямого прохода поступают с внешних устройств, а оптимизированный процесс прямого прохода позволяет получать результаты предсказаний быстрее и более эффективно. Пример можно увидеть на рисунке 1.

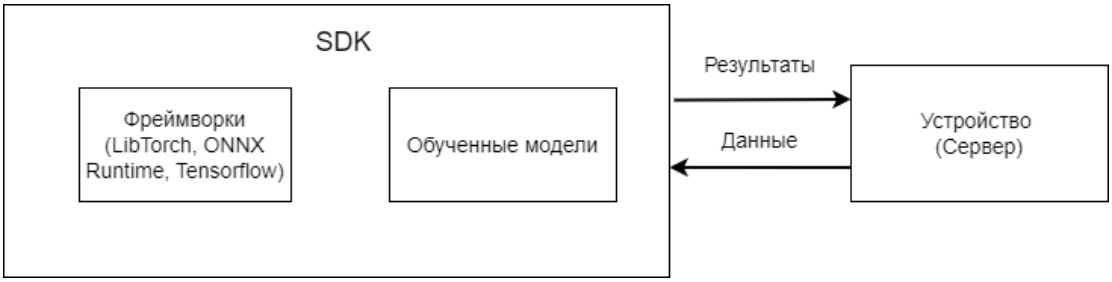


Рис. 1. Схема работы системы

Структура каждой главы представляет собой следующую последовательность подразделов: Сборка фреймворка, конвертация модели, выполнение фреймворка, методы оптимизации (дополнительные главы), выводы. Некоторые из указанных подразделов могут быть опущены в зависимости от характера исследования и наличия соответствующих данных.

Выбор фреймворков основан на результатах статьи “Оптимизация нейронных сетей прямого распространения” [6]

Для оптимизации времени прямого прохода модели, был взят LibTorch (версии 1.12.0) - C++ интерфейс фреймворка PyTorch. Libtorch предоставляет широкий выбор настроек и параметров для оптимизации прямого прохода.

Для того, чтобы начать исследование оптимизации, нужно узнать показатели производительности LibTorch без оптимизации. Запускаем прямой проход модели IResNet-50 и получаем результаты времени работы: среднее значение – 99 мс, медиана – 99 мс.

Основные этапы сборки Libtorch:

1. Подготовка окружения:

Сначала устанавливаем все зависимости, такие как компиляторы (например, GCC, Clang), системы сборки (CMake), библиотеки линейной алгебры (BLAS, LAPACK) и другие. CMake - инструмент для автоматизации сборки проекта, который генерирует конфигурационные файлы для сборки.

2. Скачивание исходного кода LibTorch:

Клонируем репозиторий с исходным кодом PyTorch или скачиваем архив с исходным кодом с официального сайта.

3. Конфигурация и сборка:

Создаем директорий для сборки и выполняем конфигурацию сборки с помощью CMake, указываем необходимые параметры и пути к зависимостям. Дальше запускаем процесс сборки с помощью системы сборки (например через команду “cmake --build”) и устанавливаем библиотеку LibTorch в систему.

4. Интеграция Libtorch в проект:

Линкуем библиотеку LibTorch к проекту на языке C++, указываем пути к заголовочным файлам и библиотекам и используем функциональность LibTorch.

Для экспериментов были выбраны две опции сборки, так как мы оптимизируем прямой проход одного потока при CPU вычислениях:

– использование опции CMake *-DUSE_FAKELOWP=ON*, которая включает сборку Libtorch с использованием низкоточных типов данных таких как 8 битные целые числа и числа с половинной точностью. Результат: среднее значение – 99 мс, медиана – 99 мс;

– использование опции CMake *-DUSE_FBGEMM=OFF*, которая включает низкоточное, высокопроизводительное умножение матриц на матрицы и использование своей библиотеки сверточных операций. Результат: среднее значение – 99 мс, медиана – 99 мс.

Для того, чтобы подать модель на прямой проход фреймворка Libtorch, нам нужно преобразовать модель в формат TorchScript (.pt). Для этого сначала загружаем предобученную модель и ее веса в Python. Затем убеждаемся, что модель находится в режиме оценки (англ. evaluation mode)

Далее выполняем компиляцию модели с помощью PyTorch JIT. Это позволяет конвертировать её в набор TorchScript операций и сохранить результат в файле с расширением “.pt”. Дальше уже подаем файл в LibTorch для выполнения прямого прохода.

Для эксперимента мы используем класс c10::InferenceMode, который отключает градиент модели, а также дополнительные операции, связанные с обучением. Это позволяет снизить нагрузку на вычислительные ресурсы и сфокусироваться на прямом проходе. Во время выполнения происходит вызов конструктора c10::InferenceMode. Под конец выполнения результаты показывают улучшение производительности: среднее значение времени прямого прохода составляет 91 мс, медианное значение – 91 мс.

В данном разделе были рассмотрены различные этапы оптимизации и ускорения прямого прохода предобученной модели с использованием Libtorch. Описан процесс сборки Libtorch с различными опциями для оптимизации производительности, а также подробно разобраны методы конвертации модели из формата PyTorch в формат TorchScript для использования в LibTorch.

Проведенные эксперименты включали использование класса c10::InferenceMode для отключения градиента модели и опций сборки. Эксперименты показали улучшение производительности, время прямого прохода было уменьшено до 91 мс.

Результаты оптимизаций представлены в таблице 1.

Таблица 1

Результаты оптимизации времени прямого прохода фреймворка Libtorch

Оптимизация	Время прямого прохода (средние значение / медиана) в мс
Без оптимизации	99 / 99
c10::InferenceMode	91 / 91
USE_FBGEMM = OFF	99 / 99
USE_FAKELOWP = ON	99 / 99

Остальные опции CMake используются для двух и больше потоков или для других вычислителей.

Для оптимизации времени прямого прохода модели, был взят фреймворк ONNX Runtime версии 1.12.1. ONNX Runtime — это кроссплатформенный фреймворк для обучения и прямого прохода нейронных сетей. Для конфигурации и сборки используется CMake.

Чтобы начать исследование оптимизации, нужно узнать показатели производительности ONNX Runtime без оптимизации. Запускаем прямой проход модели IResNet-50 и получаем результаты времени работы: среднее значение – 80 мс, медиана – 79 мс.

Дальше объясняется сборка и работа ONNX Runtime с проведением экспериментов по оптимизации. На рисунке 2 показана схема работы ONNX Runtime.

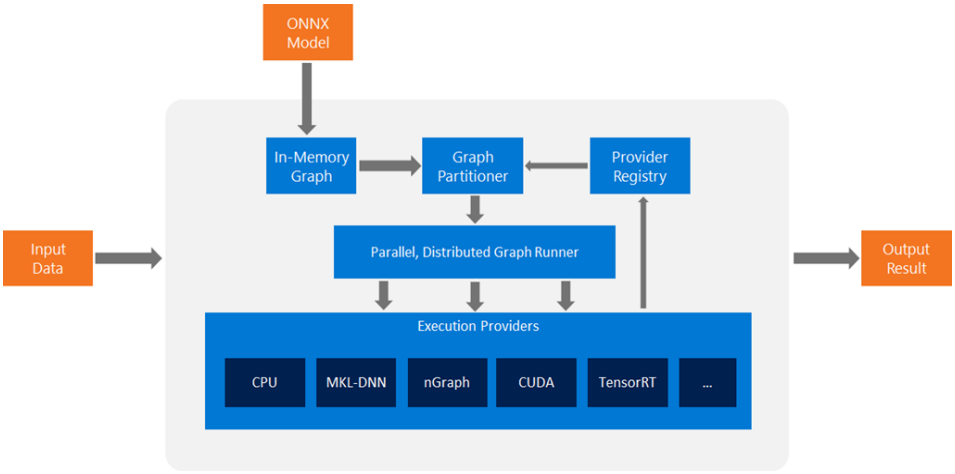


Рис. 2. Схема работы ONNX Runtime

Во время сборки фреймворка ONNX Runtime через инструмент CMake одной из возможных оптимизаций является изменение опций сборки и провайдера выполнения.

Провайдер выполнения (англ. Execution Provider) – это дополнительный модуль для ONNX Runtime, который предоставляет собственный интерфейс и позволяет исполнять код прямого прохода модели на различной аппаратной платформе (CPU, GPU и т.п.).

Для экспериментов были выбраны две опции сборки, так как мы оптимизируем прямой проход одного потока при CPU вычислениях:

— использование опции CMake *-Donnxruntime_USE_LLVM=ON*. Результат: среднее значение – 79 мс, медиана – 79 мс;

— использование опции CMake *-Donnxruntime_BUILD_FOR_NATIVE_MACHINE=ON*, которая включает все доступные оптимизации компиляции на устройстве на котором происходит сборка. Результат: среднее значение – 79 мс, медиана – 79 мс.

Для того, чтобы подать модель на прямой проход фреймворка ONNX Runtime, требуется модель формата ONNX. Из-за того, что не была найдена ONNX модель IResNet-50, было решено сконвертировать PyTorch модель в ONNX формат. Для этого можно использовать из PyTorch библиотеки метод torch.onnx.export, который позволяет сконвертировать предобученную модель из представления PyTorch в ONNX представление.

На первом этапе ONNX Runtime конвертирует граф модели ONNX в собственное представление графа в памяти. На этом этапе не происходит никаких оптимизаций.

На втором этапе происходит выполнение набора независимых от провайдера оптимизаций. Один из них - метод SetGraphOptimizationLevel – этот метод позволяет определять набор из 4-х уровней оптимизации. Каждый уровень оптимизации включает в себя все уровни оптимизации, которые находятся ниже по уровню. Результаты каждого уровня оптимизации описываются дальше в работе.

На третьем этапе ONNX Runtime разбивает граф на множество подграфов. На этом этапе не происходит никаких оптимизаций.

На последнем этапе каждый подграф назначается провайдеру выполнения. ONNX гарантирует, что подграф может быть выполнен провайдером выполнения. По умолчанию, подграфам назначается провайдер выполнения от ONNX Runtime для CPU вычислителей. Для исследования были взяты сторонние провайдеры TVM (Tensor Virtual Machine), DNNL (Deep Neural Network Library), OpenVino.

TVM – это провайдер выполнения, который построен на базе Apache TVM. В настоящее время не компилируется для языка программирования C++.

DNNL – это провайдер выполнения, который создан компанией Intel и Microsoft. В настоящий момент называется oneDNN. Содержит собственные векторные и потоковые блоки. В настоящий момент имеет проблемы с производительностью. Результат: среднее значение – 99 мс, медиана – 99 мс.

OpenVino - это провайдер выполнения, который построен на базе OpenVino toolkit. Результат: среднее значение – 70 мс, медиана – 69 мс.

Также на сторонние провайдеры распространяется первый уровень оптимизации графа. Эксперименты с первым уровнем оптимизации показали те же результаты для сторонних провайдеров выполнения, поэтому было решено не приводить повторно результаты.

Результаты работы при использовании определенного уровня оптимизации:

– атрибут GraphOptimizationLevel::ORT_DISABLE_ALL (0 уровень) отключает любую оптимизацию графа. Результат: среднее значение – 80 мс, медиана – 79 мс;

– атрибут GraphOptimizationLevel::ORT_ENABLE_BASIC (1 уровень) оптимизирует граф путём удаления ненужных узлов и уменьшением лишних вычислений. Реализуется в 3 шага:

– статически вычисляет часть графа, которое зависит от константных инициализаторов;

– удаляет все ненужные узлы без изменения структуры графа. Узлы, поддерживающие оптимизацию: Identity, Slice, Unsqueeze, Dropout;

– объединяет (англ. Fusion) нескольких узлов в один узел. Поддерживаемые слияние узлов: Conv Add, Conv Mul, Conv BatchNorm, ReLu Clip, Reshape,

Результат: среднее значение – 80 мс, медиана – 80 мс.

– атрибут GraphOptimisation:ORT_ENABLE_EXTENDED (2 уровень) оптимизирует граф путём объединения сложных узлов графа. Результат: среднее значение – 79 мс., медиана – 79 мс;

– атрибут GraphOptimizationLevel::ORT_ENABLE_ALL (3 уровень) изменят тип хранения многомерной матрицы NCHW на NCHWc для определенных узлов. Результат: среднее значение – 64 мс., медиана – 64 мс.

Если посмотреть на результаты, то видно, что 3 уровень оптимизации, то есть передача атрибута GraphOptimizationLevel::ORT_ENABLE_ALL в метод SetGraphOptimizationLevel, позволяет нам уменьшить время работы прямого прохода. Оптимизация применяется только для CPU вычислителей архитектуры x86.

Формат NCHW представляет собой описание размерности многомерной матрицы, обрабатываемое процессором, где N — число образцов, C - число каналов, H - высота, W – ширина. Так как процессор может иметь разные разрядности векторных инструкций, авторы фреймворка ONNX предложили новый формат описания размерности многомерной матрицы для оптимизации – NCHWc. В этом формате добавляется новая размерность с равная разрядности векторных инструкций, а размерность каналов делится на размерность c. Происходит разделение и переупорядочивание циклов для обработки многомерных векторов. Подобная трансформация позволяет процессору определенной разрядности выполнять быстрее

операции той же разрядности в зависимости от расширений системы. Пример расположение данных можно увидеть на рисунке 3.

Преобразование NCHW в NCHWc происходит для слоёв:

- свёрточный (англ. Convolution);
- пулинг (англ. Pool);
- функция активации (англ. Activation);
- пакетная нормализация (англ. Batch Normalization);
- изменение размера (англ. Resize);
- объединение (англ. Concat);
- бинарный (англ. Binary).

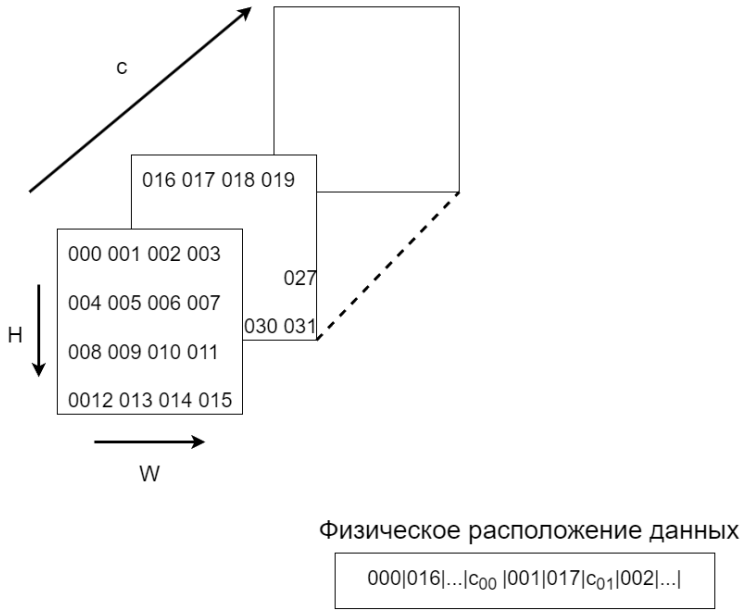


Рис. 3. Представление в памяти многомерной матрицы NCHWc.
Размерность N и C – были опущены

OpenVino - это фреймворк с набором инструментов, разработанный Intel, для оптимизации и прямого прохода предобученных моделей. OpenVino может использоваться не только в качестве провайдера выполнения в ONNX Runtime, но и как отдельный фреймворк. Благодаря этому его можно использовать для исследования оптимизации времени прямого прохода модели, а также для сравнения производительности с другими фреймворками и решениями. Одним из преимуществ OpenVino является то, что он предлагает оптимизированные реализации для различных аппаратных платформ, что может помочь улучшить производительность и эффективность применения предобученных моделей в реальных сценариях.

Сборка выполняется аналогично сборке LibTorch, однако перед интеграцией в проект необходимо настроить окружение и выполнить скрипт OpenVINO для получения всех путей к библиотекам.

Фреймворк OpenVino принимает предобученную модель как в формате ONNX, так и в формате IR (анг. Intermediate Representation). Так как модель IResNet-50 в формате IR не была найдена, то для получения IR формата, был использован инструмент Model Optimizer из библиотеки OpenVino, которая позволяет нам конвертировать предобученную модель в формате ONNX в формат IR, при этом инструмент применяет свои собственные настройки оптимизации для модели.

Фреймворк Tensorflow_CC версии 2.11.0 представляет собой C++ интерфейс фреймворка Tensorflow. Так как название Tensorflow_CC мало используется, то дальше заменяется на название Tensorflow.

Основные этапы сборки Tensorflow:

1. Подготовка окружения:

Устанавливаем систему сборки Bazel. Bazel – это система сборки созданная Google для конфигурации и компиляции кодовых баз.

2. Скачивание исходного кода Tensorflow:

Клонируем репозиторий с исходным кодом Tensorflow или скачиваем архив.

3. Конфигурация и сборка:

Создаем директорий для сборки и выполняем конфигурацию сборки с помощью Bazel, указываем необходимые параметры и пути к зависимостям. Дальше запускаем процесс сборки с помощью системы сборки (например через команду “bazel build --config=opt tensorflow:libtensorflow_cc.so”) и устанавливаем библиотеку Tensorflow в систему.

4. Интеграция Tensorflow в проект:

Линкуем библиотеки Tensorflow к проекту на языке C++, указываем пути к заголовочным файлам и библиотекам и используем функциональность Tensorflow

Так как фреймворк Tensorflow работает только с моделями формата SavedModel, а модель IResNet-50 не была найдена в этом формате, то конвертация модели происходит путём использования инструмента ONNX-Tensorflow, где модель IResNet-50 в формате ONNX конвертируется в SavedModel. SavedModel – формат хранения Tensorflow модели, который содержит структуру нейронной сети с обученными параметрами.

Для исследования оптимизации времени прямого прохода было проведено несколько экспериментов с опциями сборками:

– Использовании по умолчании опции -mkl, которая включает поддержку библиотеки MKL (Math Kernel Library). Результат: среднее значение – 90 мс, медиана – 90 мс;

– Использование опции -mfma, которая применяет оптимизацию для инструкций умножения-сложения над числами в формате с плавающей запятой (англ. Fused Multiply Add, FMA). Результат: среднее значение – 90 мс, медиана – 90 мс;

– Использование опции -mavx2, которая применяет оптимизацию для набора AVX2-инструкций (Advanced Vector Extensions 2). Результат: среднее значение – 90 мс, медиана – 90 мс.

После получения результатов оптимизации времени прямого прохода для каждого фреймворка на CPU архитектуре при одном потоке, были проведены эксперименты с использованием 1, 2, 4, 6 и 12 потоков и двумя типами параллелизма: inter-op и intra-op. Intra-op означает распараллеливание одной операции на несколько потоков для более быстрого выполнения. Inter-op означает выполнение нескольких операций одновременно. Количество потоков было выбрано оптимально относительно характеристики устройства.

При прямом проходе модели IResNet-50 без оптимизации LibTorch, получаем результаты времени работы представленные в таблице 2:

Таблица 2

Результаты прямого прохода фреймворка Libtorch без оптимизации

inter/intra	1	2	4	6	12
1	106/106 мс	63/64 мс	41/42 мс	35/35 мс	38/37 мс
2	108/108 мс	63/63 мс	41/41 мс	35/35 мс	38/38 мс
4	107/108 мс	62/62 мс	41/41 мс	35/35 мс	38/37 мс
6	108/108 мс	63/64 мс	41/41 мс	35/35 мс	37/37 мс
12	108/108 мс	62/63 мс	40/41 мс	36/36 мс	37/37 мс

Дальше проводим эксперименты с использованием класса c10::InferenceMode для оптимизации времени прямого прохода. Результаты времени работы представлены в таблице 3.

Таблица 3

Результаты оптимизации времени прямого прохода фреймворка Libtorch с использованием класса c10::InferenceMode

inter/intra	1	2	4	6	12
1	99/99	57/57	36/36	30/31	32/33
2	99/99	57/57	36/36	30/30	33/40
4	99/100	56/57	36/36	31/30	32/32
6	99/99	57/57	35/35	30/30	32/32
12	99/100	57/57	35/35	30/30	32/33

Также в рамках проведенных экспериментов была использована последняя версия LibTorch (1.13.1). Результаты прямого прохода без оптимизации и с применением класса c10::InferenceMode представлены в таблицах 4 и 5 соответственно. Проведя исследования с inter-op, было принято решение не включать результаты inter-op для различных потоков, так как они не дали различимых результатов.

Таблица 4

Результаты прямого прохода фреймворка Libtorch без оптимизации

inter/intra	1	2	4	6	12
По умолчанию	107/107	63/64	40/41	35/35	38/37

Таблица 5

Результаты оптимизации времени прямого прохода фреймворка Libtorch с использованием класса c10::InferenceMode

inter/intra	1	2	4	6	12
По умолчанию	99/99	57/57	36/36	30/30	32/32

При прямом проходе модели IResNet-50 без оптимизации ONNX Runtime для набора операторов (Opset) версий 10 и 13, получаем результаты времени работы, показанные в таблице 6. Opset (англ. operator set) - это набор операторов и их версий, которые поддерживаются в определенной версии ONNX.

Таблица 6

Результаты прямого прохода фреймворка ONNX Runtime без оптимизации

inter/intra	1	2	4	6	12
1	81/81	45/45	26/26	22/23	25/24
2	82/82	45/45	26/26	22/22	25/24
4	82/82	45/45	26/26	22/22	25/24
6	82/82	45/45	26/26	22/23	25/24
12	82/82	45/45	26/26	22/24	25/24

Дальше проводим несколько экспериментов для оптимизации времени прямого прохода:

– использование метода SetGraphOptimizationLevel с 3 уровнем оптимизации. Результаты времени работы представлены в таблице 7;

– использование провайдера выполнения DNNL. Результаты времени работы представлены в таблице 8;

– использование провайдера выполнения OpenVino. Результаты времени работы представлены в таблице 9.

Таблица 7

Результаты оптимизации времени прямого прохода фреймворка ONNX Runtime с 3 уровнем оптимизации

inter/intra	1	2	4	6	12
1	67/67	34/35	18/18	14/14	15/17
2	67/67	34/35	18/18	14/14	15/16
4	66/67	34/34	18/18	14/14	15/15
6	67/67	34/35	18/18	14/14	15/14
12	67/67	34/34	18/18	14/14	15/15

Таблица 8

Результаты оптимизации времени прямого прохода фреймворка ONNX Runtime с использованием провайдера выполнения DNNL

inter/intra	1	2	4	6	12
1	98/99	54/54	32/32	27/27	28/27
2	99/99	54/54	32/32	27/27	28/27
4	99/98	54/54	32/33	27/27	28/28
6	99/99	54/54	32/32	27/27	28/28
12	99/100	54/54	32/32	27/27	28/27

Проведя исследования с inter-op, было принято решение не включать результаты inter-op для различных потоков при использовании фреймворка ONNX Runtime, так как они не дали различимых результатов.

Таблица 9

Результаты оптимизации времени прямого прохода фреймворка ONNX Runtime с использованием провайдера выполнения OpenVino

inter/intra	1	2	4	6	12
1	72/72	38/38	21/20	16/16	17/17

Также в рамках проведенных экспериментов была использована последняя версия ONNX Runtime (1.14.0). Результаты прямого прохода без оптимизации и с 3 уровнем оптимизации представлены в таблицах 10 и 11 соответственно.

Таблица 10

Результаты прямого прохода фреймворка ONNX Runtime без оптимизации

inter/intra	1	2	4	6	12
1	82/82	45/45	26/26	22/22	24/24

Таблица 11

Результаты прямого прохода фреймворка ONNX Runtime с 3 уровнем оптимизации

inter/intra	1	2	4	6	12
1	67/67	34/34	18/18	14/14	15/15

При прямом проходе модели IResNet-50 без оптимизации Tensorflow, получаем результаты времени работы, показанные в таблице 12.

Таблица 12

Результаты прямого прохода фреймворка Tensorflow без оптимизации

inter/intra	1	2	4	6	12
1	95/93	56/56	36/35	29/29	32/32
2	95/94	56/56	36/35	28/29	33/32
4	95/93	56/56	36/36	29/29	32/32
6	94/93	56/56	36/35	29/30	33/32
12	95/93	56/56	36/35	29/29	32/32

Следующие эксперименты по оптимизации с использованием опций *-mfma* и *-mavx2* не дали различимых результатов, поэтому их результаты не были включены.

Результатами моей работы являются исследование и проведение экспериментов по оптимизации фреймворков для улучшения времени работы прямого прохода нейронной сети и получение итоговых метрик. По итоговым метрикам оптимизированный фреймворк ONNX Runtime на CPU вычислителе с одним потоком выполняет прямой проход на 27 миллисекунд (29.7%) быстрее фреймворка Libtorch с методом c10::InferenceMode, на 26 миллисекунд (28.9%) быстрее фреймворка Tenosrflow и на 6 миллисекунд (8.6 %) быстрее фреймворка OpenVino. На CPU вычислителе с несколькими потоками время прямого прохода на 37 (67.2%) миллисекунд быстрее фреймворка Libtorch с методом c10::InferenceMode и на 18 миллисекунд (50%) быстрее фреймворка Tenosrflow. Тем самым использование ONNX Runtime является оптимальным для оптимизации времени прямого прохода на CPU архитектуре.

Литература

1. Демидовский А.В. Особенности программных и аппаратных средств для оптимизации инференса обученных нейронных сетей в рамках среды OpenVINO //XXI Международная научно-техническая конференция "нейроинформатика-2019". 2019. С. 204-212.
2. Liu Y. et al. Optimizing CNN model inference on CPUs. 2019.
3. Пчелинцев С.Ю. Сравнительный анализ фреймворков глубокого обучения //Информационно-экономические аспекты стандартизации и технического регулирования. 2020. №. 1. С. 41-51.
4. IResNet-50 [Электронный ресурс]. Режим доступа: https://github.com/iduta/iresnet (дата обращения: 27.02.2023).
5. Duta I. C. et al. Improved residual networks for image and video recognition //2020 25th International Conference on Pattern Recognition (ICPR). IEEE. 2021. P. 9415-9422.
6. Шеметов Ф.А. Оптимизация нейронных сетей прямого распространения //XI Конгресс молодых учёных. 2022. С. 192-196.

ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ

УДК 004.8

СПЕЦИФИКА СБОРА ДАТАСЕТА ПРИ СОЗДАНИИ ЦИФРОВОГО ПОМОЩНИКА ДЛЯ СЛАБОВИДЯЩИХ ЛЮДЕЙ

Абрегова З.Х.[1](студент)
Научный руководитель – к.э.н., доцент Тимченко О.В.[1]
1 – Пятигорский государственный университет
e-mail: zalinaabregova@mail.ru

Аннотация
Цифровые помощники для людей с ограниченными возможностями зрения стали истинным спасением. Однако, чтобы создать такого помощника, нужен объемный и точный датасет, который будет являться базой для работы алгоритмов и нейронных сетей. В работе рассматривается проблема сбора датасета для создания цифрового помощника людям, имеющим проблемы со зрением.
Ключевые слова
Датасет, цифровой помощник, слабовидящие, нейросеть,тифлокомментарий.

В современном мире технологии становятся все более важными для облегчения повседневной жизни людей. Однако, не всем группам населения доступны эти новые возможности. Среди таких групп можно выделить слабовидящих людей, которым необходимы специальные инструменты для облегчения их жизни. В этой связи, создание цифровых помощников для слабовидящих становится актуальной проблемой.

Сегодня, каждый второй житель России имеет какое-либо нарушения зрения. Количество слабовидящих и полностью слепых людей составляет около миллиона трехсот тысяч по всей России. Слабовидящие по-прежнему испытывают трудности в жизни, несмотря на наличие технологий или умных устройств, которые им помогают [1]. Чтобы решить эту проблему существуют цифровые помощники, призванные помочь слабовидящим людям распознавать текст на реальных объектах, сами объекты, предоставлять звуковую обратную связь в режиме реального времени [4]. Но, и тем не менее, существующие интеллектуальные голосовые ассистенты не предназначены для озвучения объектов культуры и искусства, и не в состоянии в режиме реального времени описать и передать содержание фильма, спектакля, арт-объекта.

Для решения данной проблемы необходимо осуществлять тифлокомментирование, которое может быть создано либо автоматически, либо напрямую человеком. Тифлокомментарии, созданные человеком, бесспорно, остаются более точными и ёмкими, но они также являются более дорогими и требуют больших временных затрат. В то время как, аудиодескрипцию можно генерировать автоматически с помощью нейронных сетей. На данный момент нейронные сети прекрасно справляются с реферированием, генерацией и озвучкой текста, т.е. с задачами Text-To-Speech. Но при этом тексты, которые генерируются, могут содержать слова, конструкции или обороты, которые недопустимы в тифлокомментировании. Тут и возникает проблема сбора датасета для такого рода цифрового помощника.

Правильно составленный датасет является фундаментом для разработки эффективных алгоритмов машинного обучения и успешного функционирования приложений для слабовидящих. Во-первых, необходимо учитывать, что слабовидящие люди могут иметь разные степени нарушения зрения и потребности в поддержке. Поэтому важно определить типы пользователей и их потребности для составления соответствующего датасета.

Во-вторых, при сборе данных необходимо использовать различные типы контента: текст, изображения и аудио. Текст должен быть легко читаемым, а изображения должны быть доступными для скринридера или иметь описательный текст. Аудиоконтент также

должен быть четким и понятным. Поэтому, при сборе датасета рекомендуется воспользоваться методами аудиодескрипции, а именно следует избегать метафор, многозначных слов, сложных сочетаний и оборотов, избыточной информации, не имеющей смысла, а также учитывать стилевую принадлежность слов, их экспрессивность [3, 5]. Важность подчеркивается еще и тем, что данные должны содержать только реальные, точные, конкретные характеристики, поскольку цифровой помощник предназначен для озвучивания объектов культуры в реальном времени [2].

Наконец, необходимо привлечение экспертов-слабовидящих людей для получения обратной связи. Это поможет улучшить качество датасета и сделать его более эффективным для пользователей.

Важно помнить, что сбор датасета должен быть непрерывным процессом, который позволяет адаптироваться к изменениям в потребностях пользователей.

После того, как был проведен сбор данных для создания цифрового помощника для слабовидящих людей, необходимо произвести анализ и обработку полученных данных. Этот процесс является важным этапом разработки такого приложения. Во-первых, необходимо провести анализ качества данных. Это включает проверку наличия ошибок и пропусков в полученных данных. Если присутствуют ошибки или отсутствуют данные, то это может повлиять на точность работы помощника.

Далее следует произвести обработку данных. Она может включать в себя фильтрацию, группировку и преобразование данных. Например, можно объединить данные из разных источников или конвертировать форматы файлов для последующего использования.

Кроме того, нужно учитывать этические аспекты при работе с данными. Данные должны быть защищены от несанкционированного доступа и использования.

Наконец, после анализа и обработки данных создается датасет для дальнейшего использования его в цифровом помощнике для слабовидящих людей. Качество этого датасета напрямую влияет на работу приложения, поэтому необходимо уделить этому процессу достаточно внимания и времени.

Реализацию цифрового помощника – экскурсовода по объектам искусства и культуры рассмотрим подробнее на экспозиции домика – музея М.Ю. Лермонтова в г. Пятигорске.

На первом этапе разработки было принято решение об архитектуре, предполагалось использование GPT и голосового инструмента Google Text-To-Speech.

Голосовой инструмент Google (GTTS) – это API-интерфейс для преобразования текста в речь с использованием голосовых моделей Google. Он обеспечивает простой способ добавления речи в приложения и программное обеспечение на разных языках и голосовых акцентах.

GTTS может быть полезен для создания приложений, которые поддерживают голосовое взаимодействие с пользователем, а также для автоматизированной или голосовой транскрипции текста.

На втором этапе предполагалось обучать нейросеть на уже имеющихся описаниях экспонатов. Тестирование показало, что нейронная сеть вполне приемлемо генерирует описание экспонатов, но полученные описания не удовлетворяли требованиям тифлокомментирования.

Следующий этап состоял в эмпирическом исследовании вопроса улучшения аудиосопровождения. Для этого было принято решение о реферировании текстов. Таким образом, возникла необходимость применения модели – трансформера BertSumExt. BertSumExt метод извлекающего реферирования на основе BERT. BertSumExt использует BERT для извлечения важных признаков из входного текста и затем применяет сложную модель, которая максимизирует совместную вероятность появления всего краткого описания. Модель использует механизм выборки ключевых слов, который позволяет ей выбрать наиболее значимые слова из текста, которые

впоследствии используются при генерации краткого описания. BertSumExt обучалась на 137 текстовых описаниях.

В отличие от других алгоритмов автоматической генерации кратких описаний текста, BertSumExt полагается на извлечение важной информации из текста, а затем генерирует описание на основе этой информации. Это позволяет алгоритму генерировать более информативные, точные и связные описания. BertSumExt является одним из наиболее эффективных и точных алгоритмов автоматической генерации кратких описаний текста (на сегодняшний день и используется во многих приложениях).

Прежде всего, следует отметить, что создание тифлокомментариев – это ответственный процесс и требует точности и понимания того, как люди с ограниченными возможностями зрения получают информацию. Для создания тифлокомментария на основе сгенерированного текста можно использовать модель генерации языка GPT (Generative Pre-trained Transformer).

Шаг 1: Подготовка данных

Первым шагом необходимо подготовить данные для обучения модели. На этом этапе происходит сбор набора текстовых данных, которые будут использоваться для обучения и генерации текста (мы используем результаты, полученные BertSumExt).

Шаг 2: Обучение модели

Модель обучается на выбранном наборе данных. Обучение может занять длительное время, в зависимости от размера набора данных и сложности архитектуры модели.

Шаг 3: Генерация текста и создание тифлокомментария

По результатам обучения генерируется текст, который впоследствии и становится тифлокомментарием (с учетом соблюдения методов аудиодескрипции).

Следует отметить, что создание тифлокомментариев – это сложная задача, которая требует профессиональных навыков и знаний.

Для оценки качества генерации текста использовалась метрика BLEURT (BilingualEvaluationUnderstudywithRepresentationsfromTransformers), основанный на представлениях из Transformer. Модель обучается на парах шаблонных ответов (reference texts) и генерируемых ответов (candidate texts), полученных от генеративной модели. Модель выполняет операции над полученными представлениями, извлекая признаки, такие как сходство с шаблоном, лексические и смысловые ошибки, стилистическое соответствие и другие.

BLEURT позволяет учитывать лингвистические феномены, как перефразирование. В целом, BLEURT представляет собой мощный и эффективный инструмент для оценки качества генерации текста, который может быть использован в различных задачах генерации текста и во многих языковых парах.

Построенная модель генерации тифлокомментария была протестирована сначала на полных описаниях, затем на рефератах полученных от BertSumExt. В качестве оценки были использованы метрики BLUERT, полученная автоматически и метрика precision, полученная как отклик от эксперта по тифлокомментированию, поскольку ни одна автоматическая метрика не может оценить правильность тифлокомментария. Сравнение метрик приведено в таблице.

Таблица

Сравнение метрик

	Фактическое описание экспоната		Реферат от BertSumExt	
Эпоха обучения	BLUERT	Точность генерации	BLUERT	Точность генерации
25	0,8867	0,4238	0,8757	0,6112
50	0,9054	0,4556	0,9136	0,6227
75	0,9361	0,4901	0,9402	0,6354

Анализ полученных результатов показывает, что использование предобработки текста в виде реферирования существенно улучшает генерацию текста и придает ему свойства тифлокомментария арт-объекта, но необходимо продолжить процесс обучения для увеличения точности генерации. Таким образом, тестирование показало необходимость дальнейшей работы по расширению датасета.

Создание цифрового помощника для слабовидящих людей является важной задачей современной технологической индустрии. Подобный цифровой помощник должен отличаться лаконичностью, и в то же время ёмкостью данных. В таком случае использование нейросетевых технологий не только возможно, но и необходимо. При создании цифрового ассистента на основе нейросетей важно понимать, что не все данные могут быть использованы для разработки системы. Поэтому грамотно составленный датасет – это важная составляющая для успешной работы цифрового помощника.

В перспективе проект разработки цифрового помощника для слабовидящих людей может иметь большое значение для социальной интеграции этой группы населения. Ведь такая система позволит им получить доступ к новой информации и ресурсам, что повысит их возможности в образовании, работе и личной жизни.

Литература

1. Плотников С.О., Сметанин Д.Ю., Басова А.В., Львутин И.А., Белоусова М.Н. Использование нейронной сети для создания цифрового помощника слабовидящим людям // E-Management. Москва: 2022. С. 73–82.
2. Аудиодескрипция [Электронный ресурс]. Режим доступа: https://24wiki.ru/%D0%90%D1%83%D0%B4%D0%B8%D0%BE%D0%B4%D0%B5%D1%81%D0 %BA%D1%8 0%D0%B8% D0%BF%D1% 86%D0%B8%D1%8F (дата обращения: 12.02.2023).
3. Тифлокомментирование (Аудиодескрипция) [Электронный ресурс]. Режим доступа: https:// dcpcinema.ru/ tiflokommentirovanie/ ?ysclid=le7h4c3e38308561377 (дата обращения: 17.02.2023).
4. Нейросеть описывает мир незрячим людям [Электронный ресурс]. Режим доступа: https://www.reg.ru/blog/nejroset-opisyvaet-mir-nezryachim-lyudyam/ (дата обращения: 17.02.2023).
5. Стандарты создания аудиодескрипции и субтитров [Электронный ресурс]. Режим доступа: https:// center1.by/ standarty-sozdaniya- audiodeskriptsii-i- subtitrov/ (дата обращения: 18.02.2023).

УДК 004.89

СБОР ГЕОЗАВИСИМЫХ ДАННЫХ ИЗ НОВОСТНЫХ РЕСУРСОВ

Авдюшина А.Е.[1], Крисанов Р.В.[1]
Научный руководитель – д.т.н., профессор Бессмертный И.А.[1]
1 – Университет ИТМО
e-mail: avdushina98@gmail.com

Аннотация
Статья предлагает алгоритм комплексного использования тематического моделирования и геоинформационной рекомендательной системы для анализа новостных данных и определения наиболее актуальных тем в различных районах Санкт-Петербурга. Алгоритм включает сбор, предварительную обработку и анализ новостных данных, определение местоположения и извлечение именованных сущностей. Результаты могут быть полезны для принятия решений о направлении развития города и его проблемах.
Ключевые слова
NLP, нейронные сети, информационный поиск, тематическое моделирование, географические-информационные системы.

В современном мире информация играет ключевую роль во всех сферах жизни. Особенно важна информация, связанная с географическим расположением объектов на планете, так как она может быть использована для множества приложений, начиная от научных исследований и заканчивая коммерческими целями. Для хранения и управления структурированными данными, связанными с географическими объектами, часто используются географические информационные системы (ГИС), которые позволяют эффективно хранить, обрабатывать и анализировать данные.

Однако существует множество неструктурированных данных, которые также имеют географические компоненты, такие как тексты, фотографии и видео, и которые не могут быть эффективно обработаны с помощью классических ГИС. Именно в этой области возникает необходимость в разработке новых методов и технологий для обработки и анализа неструктурированных географических данных.

В литературе существуют исследования, которые рассматривают географический информационный поиск и используют методы тематического моделирования и метаданных изображений в социальных сетях для формирования контекста территории. Также были проведены исследования основных проблем информационного поиска на примере новостных ресурсов.

Цель данной статьи - представить новый подход к географическому информационному поиску, основанный на использовании неструктурированных данных, таких как социальные медиа и новостные ресурсы. Мы рассмотрим проблемы и вызовы, связанные с обработкой таких данных, и предложим решения, которые могут помочь в реализации нашего подхода. Наша работа внесет вклад в развитие географических информационных систем и информационного поиска в целом, а также может быть полезна для различных областей, таких как научные исследования, туризм, маркетинг и другие.

В GIR существуют различные методы поиска информации, такие как поиск по ключевым словам, поиск по географическим координатам и комбинированный поиск, который сочетает оба подхода [1]. В качестве инструментов для работы с географической информацией могут использоваться геоинформационные системы (ГИС), которые обладают функциональностью по работе с пространственными данными. Для определения географической информации из новостных статей часто используются методы машинного обучения и анализа текста. В процессе анализа текста осуществляется выделение ключевых слов и фраз, а также извлечение смысла из текста.

Одним из методов анализа текста является тематическое моделирование, которое позволяет выявить скрытые темы в тексте и определить их веса [4]. Тематическое моделирование может использоваться в GIR для определения территориального контекста поискового запроса, что позволяет улучшить результаты поиска и повысить качество извлечения географической информации [2]. Другим методом извлечения географической информации является анализ изображений и видео. В GIR часто используется обработка изображений и видео, осуществляемая с помощью компьютерного зрения и глубоких нейронных сетей [1]. Например, системы GIR могут использовать анализ фотографий и видео, чтобы определить географическое положение объектов на них, что может быть полезно в таких областях, как картография и геология [3]. Однако при использовании методов извлечения географической информации возникают проблемы, связанные с качеством данных и точностью определения координат. Также могут возникать проблемы с обработкой больших объемов данных и быстродействием систем. Для решения этих проблем часто используются методы оптимизации и параллельных вычислений [1].

Также существует множество других методов для извлечения географической информации из неструктурированных данных, таких как анализ тональности, распознавание именованных сущностей и машинное обучение. Например, анализ тональности может использоваться для определения эмоциональной окраски текста, связанного с определенным местоположением, что позволяет лучше понимать общее настроение сообщества в этом регионе. Распознавание именованных сущностей может быть использовано для извлечения имени конкретного места из текста. Машинное обучение может помочь определить связи между текстом и географическими объектами на основе их сходства в разных аспектах, например, похожести в описании или использовании тех же ключевых слов.

Таким образом, поиск географической информации является актуальной задачей, требующей использования различных методов и подходов. GIR-системы позволяют извлекать информацию из различных источников, а использование методов машинного обучения и анализа данных помогает улучшить качество поиска и сделать его более точным и эффективным.

Исследование проводится с целью разработки эффективного алгоритма для обработки новостных данных и выявления тематик, особенностей и событий в разных районах Санкт-Петербурга. Предлагается использовать методы геоинформационного анализа (GIR) и тематического моделирования, чтобы объединить данные из разных новостных порталов в кластерах, которые будут позволять извлекать список тем.

Алгоритм обработки данных включает несколько шагов. Сначала необходимо собрать и предварительно обработать набор новостных данных для удаления стоп-слов, знаков препинания и специальных символов, а также провести нормализацию текста методами лемматизации. Затем необходимо определить местоположение каждой новостной статьи с помощью методов распознавания именованных объектов, геокодирования и картографических сервисов, что представлено на рисунке. Автоматически было выделено некоторое количество кластеров на основе плотности объектов на карте после геокодирования. Каждый кластер представляет собой подвыборку новостей. Нахождение распределение тем по подвыборке, объединенной одним кластером позволит определить контекст местности и в дальнейшем определять эмоциональную окраску и извлекать семантические отношения и наборов тем. Таким образом, объединение методов машинного обучения без учителя (кластеризация и тематическое моделирования) позволят полноценно дополнить имеющуюся классификацию тем новостных сайтов.

Далее, для обработки данных новостных порталов используется метод тематического моделирования LDA, который позволяет выделить наиболее значимые темы в наборе данных. Полученные темы сохраняются для дальнейшего использования

и анализа. Например, можно определить наиболее распространенные темы для конкретных географических местоположений.

Использование GIR и тематического моделирования позволяет выявить связи между темами и местоположениями, что может быть полезно для анализа новостных событий и разработки мер по управлению рисками. Например, можно определить, какие темы наиболее актуальны для определенных районов города и какие меры можно предпринять для предотвращения или уменьшения рисков, связанных с этими темами (рисунок).

Рисунок. Пример кластеризации картографических объектов

Полученный алгоритм позволяет получать более точные и релевантные данные о событиях и темах, происходящих в различных районах города. Например, представители власти могут использовать эти данные для принятия решений о выделении средств на улучшение инфраструктуры в конкретном районе, основываясь на данных о темах, которые наиболее актуальны для жителей этого района.

Также полученный алгоритм может быть полезен журналистам и редакторам новостных порталов, позволяя им быстро получать информацию о темах, которые наиболее интересны для конкретных географических областей. Это позволит им более точно и эффективно освещать важные события и темы, которые наиболее актуальны для конкретных местоположений.

Кроме того, полученный алгоритм может быть использован в различных сферах бизнеса. Например, компании, занимающиеся маркетингом и продажами, могут использовать его для более точного и эффективного определения целевой аудитории и рекламных кампаний в разных географических областях. Также полученный алгоритм может быть использован для анализа социальной ситуации в разных регионах, что может помочь при принятии решений о социальной политике.

Таким образом, полученный алгоритм является эффективным инструментом для анализа тем и событий в различных географических областях и может быть использован в различных сферах, таких как политика, журналистика, бизнес и социальная сфера.

Литература

1. Jones C., Purves R.S. Geographical information retrieval // International Journal of Geographical Information Science. 2008. V.22. P. 219-228. [Электронный ресурс]. Режим доступа: DOI: 10.1080/13658810701626343 (дата обращения: 09.04.2023).
2. Funkner A.A. Geographical Topic Modelling on Spatial Social Network Data / Funkner A.A., Elkhovskaya L.O., Lenivtceva I.D., Egorov M.P., Kshenin A.D., Khrulkov A.A.//

Procedia computer science. 2021. V. 193. P. 22–31.
3. Zharikov A. Information Retrieval System for News Articles in Russian / Zharikov A., Kristalovsky K., Pivovarov V. // Web of Data: The joint RuSSIR/EDBT 2011 Summer School, August 15–19. 2011. Proceedings of the Fifth Russian Young Scientists Conference in Information Retrieval / B. Novikov, P. Braslavsky (Eds.). St. Petersburg. 2011. P. 5–14.
4. Olieman A., Kamps J., Merino Claros R. LocLinkVis: A Geographic Information Retrieval-Based System for Large-Scale Exploratory Search // 2015. [Электронный ресурс]. Режим доступа: https://arxiv.org/pdf/1509.02010.pdf (дата обращения: 09.04.20230.

УДК 004.942

RCPSP RESOURCE-CONSTRAINED PROJECT SCHEDULING PROBLEM (ЗАДАЧА ПЛАНИРОВАНИЯ ПРОЕКТОВ В УСЛОВИЯХ ОГРАНИЧЕННОСТИ РЕСУРСОВ)

Алексеев А.С.[1] (студент)
Научный руководитель – Ковальчук М.А.[1] (аспирант)**, Филатова А.А.**[1] (аспирант)
1 – Университет ИТМО
e-mail: aalexeev@itmo.ru

Аннотация
Решая задачу построения оптимального расписания выполнения работ проекта с ограниченными ресурсами (RCPSP), где требуется учитывать отношения предшествования между работами и доступность ресурсов, исследователи в течение многих лет используют преимущественно один и тот же набор синтетических данных, который не универсален. В данной статье рассматриваются альтернативные наборы данных, а также исследуются способы генерации новых наборов данных для RCPSP, которые лучше соответствуют особенностям, присутствующим в реальных проектах.
Ключевые слова
RCPSP, теория расписаний, проект, синтетические данные, генераторы данных, графы.

Задача построения оптимального расписания выполнения работ проекта с учетом отношений предшествования между работами и с учетом необходимых и (или) доступных ресурсов (resource-constrained project scheduling problem, далее – RCPSP) – одна из наиболее изучаемых проблем в теории расписаний, которая ставит целью составление расписания с наименьшим временем выполнения проекта, где расписание удовлетворяет условиям ограничения предшествования работ и ограничениям на возобновляемые ресурсы. Рассматриваемая проблема относится к NP-трудным, поэтому было предложено много алгоритмов, находящих решение близкое к оптимальному. Применение решений этой задачи можно найти в различных сферах (благодаря чему RCPSP имеет множество вариантов постановки задачи) и в каждой из сфер проекты могут иметь свои уникальные особенности, поэтому для получения объективной оценки применимости алгоритма решения RCPSP в конкретной сфере необходимы наборы данных с такими особенностями. Для изучения возможности генерации наборов данных для RCPSP, похожих на имеющиеся и/или удовлетворяющих определенным условиям были проанализированы существующие наборы данных, используемые для оценки методов решения RCPSP, алгоритмы генерации синтетических наборов данных для этой задачи, а также подходы для генерации графов, которые используются в других областях.

Постановка задачи RCPSP формулируется следующим образом. Проект состоит из n работ, которые необходимо выполнить. Проект можно представить в виде направленного ациклического графа $G(V, E)$, где вершинами будут работы, а ребра графа будут задавать порядок предшествования работ. Таким образом, множество вершин графа $V = \{A_0, A_1, \ldots, A_n, A_{n+1}\}$ состоит из работ A_i, где работы A_0 и A_{n+1} обозначают начало и конец проекта соответственно. Множество пар $(A_i, A_j) \in E$ будет задавать порядок предшествования работ, где каждая пара обозначает, что работа A_i должна быть выполнена прежде работы A_j. А работа A_0 будет предшествовать всем остальным работам, и работа A_{n+1} будет идти после всех работ.

Для выполнения работ в каждый момент доступны возобновляемые ресурсы $R = \{R_1, R_2, \ldots, R_q\}$. И для каждой работы A_i требуется неотрицательный ресурс r_{ij}, причем $r_{ij} \leq R_j$. После выполнения работы ресурс, затраченный на выполнение работы становится снова доступным для других работ. Длительность проекта поделена на равные

интервалы и каждой работе A_i соответствует положительная продолжительность d_i, кроме начальной и конечной работы, где продолжительность $d_0=d_{n+1}=0$.

Основой проекта является граф работ – наиболее популярный способ визуализации проектов в теории расписаний. Именно графу работ уделяется большое внимание при создании синтетических наборов данных, так как он в первую очередь влияет на сложность проекта.

Первую попытку стандартизировать подход к сравнению алгоритмов для решения RCPSP предпринял Паттерсон в 1984 году. Для сравнения методов решения RCPSP Паттерсон использовал сборку из 110 проектов, доступных в литературе того времени. Но благодаря развитию технологий и совершенствованию методов решения RCPSP со временем этот набор данных теряет актуальность и появляется необходимость создания более сложных наборов данных. Первые генераторы синтетических проектов RCPSP встречаются в статьях 1993 года, в которых предлагается два метода генерации графа работ полностью случайным образом (одинаковая вероятность генерации любого графа): метод удаления и метод добавления. Для генерации этими методами пользователь задает только два параметра: количество работ и количество взаимосвязей.

В 1996 году был представлен набор проблем PSPLIB, который впоследствии будет использован для оценки алгоритмов решения RCPSP в большом количестве статей и который не потерял популярность до сегодняшнего дня. Набор PSPLIB был создан с помощью генератора ProGen [1], который имеет больше настраиваемых параметров (список параметров приведён в таблице), чем ранее предложенные методы. На первом этапе в граф работ добавляются начальные и конечные работы, затем к ним добавляются связанные работы и процесс повторяется до достижения заданной сложности графа.

В том же году был предложен генератор DAGEN, в работе которого предлагается использовать индекс сложности (complexity index, далее - CI) (мера несоответствия графа последовательной/параллельной топологии) в качестве одного из задаваемых параметров. При этом методе генерации в начале создается основа графа (каждая вершина соединена с вершинами, у которых более высокий номер), а затем случайно выбираются две вершины и к ним применяется одно из пяти правил по удалению/добавлению вершин и ребер до тех пор, пока не будет достигнут требуемый CI.

Ни ProGen, ни DAGEN не гарантируют, что сгенерированные графы будут случайными из всего возможного множества графов. Данный факт и послужил причиной для разработки следующего генератора – RanGen [2]. Алгоритм начинает работу с полносвязного графа и случайно удаляет ребра до достижения заданного порядка прочности (order strength, далее – OS) (отношение количества соединений к максимально возможному числу соединений). Процедура повторяется многократно до достижения ограничения по времени, таким образом формируется набор графов, из которых затем случайно выбирается граф, удовлетворяющая заданному CI. В итоге для генерации графов этим методом используется всего три параметра: количество работ, CI и OS. Наиболее популярные наборы данных, полученные с помощью RanGen, это RG300 и DC2.

В статье [3] приводится сравнение имеющихся на тот момент генераторов в разрезе характеристик получаемых этими генераторами структур графов. Авторы включают в сравнение новую версию генератора RanGen. RanGen2 работает по похожему алгоритму, удаляя ребра из начального графа, но с двумя отличиями. В то время как RanGen начинает с одного и того же графа, RanGen2 начинает с набора графов, индикатор последовательности (serial or parallel indicator) (близость графа к последовательному графу) которых больше заданного. Во-вторых, удаление ребер происходит до достижения заданного индикатора последовательности. Кроме того, авторы в статье показали, что набор проблем PSPLIB покрывает лишь малую часть значений характеристик графов. С помощью RanGen2 был получен набор данных RG30.

Сети работ с одинаковым количеством работ могут быть как легкими, так и невероятно трудными для алгоритмов, решающих RCPSP (например, это было показано в статьях [1-3]). Поэтому в исследовании 2020 года изучаются самые трудные случаи для алгоритмов, решающих RCPSP, и по итогам исследования предлагается набор из 623 проектов под названием CV. Для отбора сложных случаев из имеющихся наборов данных авторы используют индикатор сложности, который учитывает время, затраченное на поиск решения, и качество решения.

Как было показано выше развитие генераторов графа работ для RCPSP шло по пути генерации наборов, которые могли бы наиболее репрезентативно представить все множество графов работ. Как видно из таблицы у генераторов имеются регулируемые параметры (индекс сложности, порядок прочности и индикатор последовательности), отвечающие только за настройку параллельности/последовательности работ. В то время как для тестирования алгоритмов решения RCPSP в конкретных областях требуются наборы данных, имеющие специфические особенности случаев из этих областей, и этих параметров недостаточно для настройки генератора под определенную задачу.

Таблица

Существующие генераторы графов работ для RCPSP

Генератор	Год	Регулируемые параметры графа работ
RANG	1993	− количество работ, − количество соединений
ProGen [1]	1995	− интервал (минимальное и максимальное число) количества работ, − интервал количества начальных работ, − интервал количества конечных работ, − максимальное количество предшествующих работ, − сложность графа
DAGEN	1996	− интервал количества работ, − интервал количества соединений, − индекс сложности (CI)
RanGen [2]	2003	− количество работ, − индекс сложности (CI), − порядок прочности (OS)
RanGen2 [3]	2008	− количество работ, − индикатор последовательности

Как отмечалось выше основной частью проектов RCPSP является граф, а в области генерации графов на сегодняшний день разработано большое количество методов. Генерация графов востребована в социальной, инженерной и научных сферах. В виде графов можно представить социальные сети, базы знаний, структуры молекул и т. д.

Традиционные подходы к генерации графов основаны на статистических и структурных предположениях. В большинстве таких подходов генерация вершин и/или ребер происходит независимо, как и в выше рассмотренных методах. В отличие от традиционных подходов на сегодняшний день получили широкое распространение генераторы графов на основе нейронных сетей [4, 5]. Эти подходы способны извлекать закономерности построения графа, основываясь на наборе графов, представленных для обучения, и генерировать новые графы, опираясь на эти закономерности.

Нейросетевые подходы генерации графов предоставляют большое разнообразие методов, имеющих свои достоинства и недостатки. Наиболее распространенным способом генерации является авторегрессионный (последовательный) подход, при котором на каждом шаге генерации добавляются новые вершины и ребра к ранее

сгенерированному графу. Альтернативой авторегрессионному подходу являются методы генерации графов целиком, результатом генерации которых является матрица смежности. К большинству нейросетевых методов генерации графом можно добавить условную генерацию, при которой вместе с набором графов для обучения передается дополнительная информация об этих графах, и таким образом сгенерированный граф может иметь требуемые особенности, свойственные графам из части обучающего набора. Большинство подходов обучаются на наборе графов, поэтому среди остальных методов можно выделить метод генерации больших графов, для обучения которого используется один большой граф.

В работе рассмотрены существующие методы генерации синтетических данных для RCPSP и нейросетевые генераторы графов, используемые в других областях. Методы генерации, применяемые для RCPSP, имеют мало параметров для настройки генерации данных похожих на реальные данные. С другой стороны, нейросетевые генераторы графов способны при обучении самостоятельно улавливать закономерности и особенности данных, которые используются для обучения. Таким образом, нейросетевые генераторы графов хорошо подходят для генерации синтетических данных похожих на реальные данные и в дальнейшем планируется разработка генератора синтетических данных для RCPSP на основе нейронных сетей.

Литература

1. Kolisch R., Sprecher A., Drexl A. Characterization and Generation of a General Class of Resource-Constrained Project Scheduling Problems // https://doi.org/10.1287/mnsc.41.10.1693. INFORMS , 1995. Vol. 41. № 10. P. 1693–1703.
2. Demeulemeester E., Vanhoucke M., Herroelen W. RanGen: A Random Network Generator for Activity-on-the-Node Networks // Journal of Scheduling 2003 6:1. Springer. 2003. Vol. 6, № 1. P. 17–38.
3. Vanhoucke M. et al. An evaluation of the adequacy of project network generators with systematically sampled networks // Eur J Oper Res. North-Holland. 2008. Vol. 187. № 2. P. 511–524.
4. Faez F. et al. Deep Graph Generators: A Survey // IEEE Access. Institute of Electrical and Electronics Engineers Inc. 2021. Vol. 9. P. 106675–106702.
5. Guo X., Zhao L. A Systematic Survey on Deep Generative Models for Graph Generation // IEEE Trans Pattern Anal Mach Intell. IEEE Computer Society. 2022.

УДК 004.089

УСТРОЙСТВО ДЛЯ ПОМОЩИ СЛАБОВИДЯЩИМ ЛЮДЯМ НА БАЗЕ ТЕХНОЛОГИИ EDGE AI

Бетеня Д.С.[1] (студент)
Научный руководитель – к.т.н., доцент Муромцев Д.И.[1]
1 – Университет ИТМО
e-mail: dashabetenia@gmail.com

Аннотация
В данной работе предлагается новое решение для помощи слабовидящим людям на базе технологии периферийного искусственного интеллекта Edge AI. Спроектированная система представляет из себя программно-аппаратный комплекс для распознавания изображений с помощью сверточной нейронной сети YOLOv5 с аппаратной поддержкой в виде микроконтроллера со встроенным нейроускорителем Kneron KL520. На базе этой же технологии был создан опытный образец устройства для распознавания голосовых команд пользователя.
Ключевые слова
Периферийный искусственный интеллект, сверточная нейронная сеть, микроконтроллер, нейроускоритель, программно-аппаратный комплекс.

Разработка устройств для улучшения качества жизни слабовидящих и незрячих людей является актуальным направлением современных исследований. Наиболее продвинутые технологии ассистирования, использующие искусственный интеллект, включают такие решения, как PeopleLens, Seeing AI и Biped.ai (рис. 1).

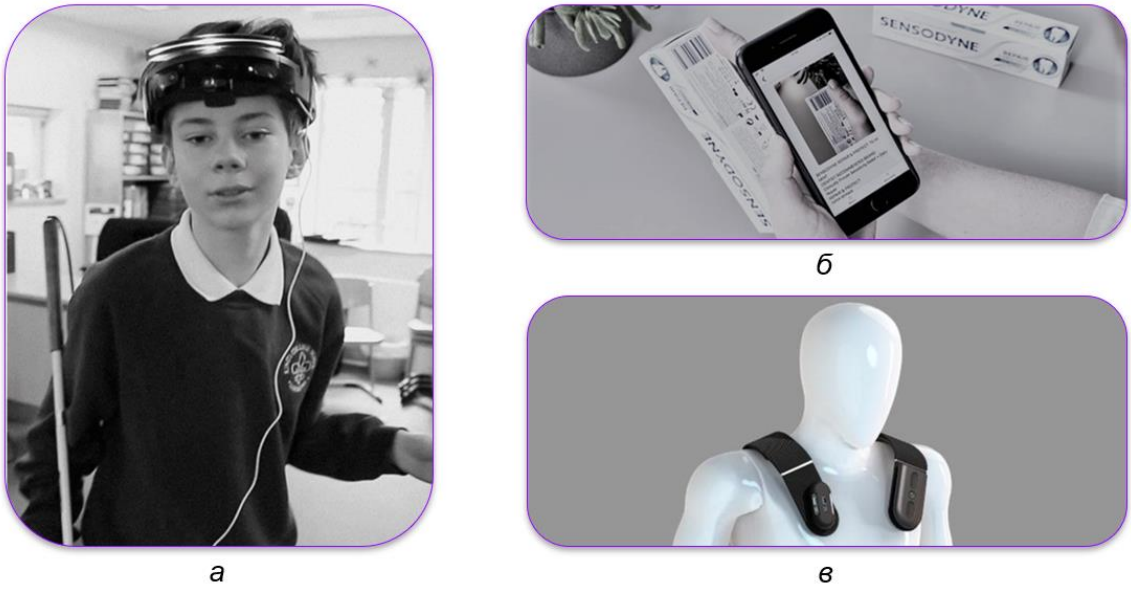


Рис. 1. Технологии для помощи слабовидящим людям:
а) PeopleLens [1]; б) Seeing AI [2]; в) Biped.ai [3]

PeopleLens компании Microsoft – платформа для помощи слепым людям взаимодействовать с их социальным окружением, состоящая из устройства в виде двух камер, закрепленных на голове пользователя, и облачного сервиса. Основное внимание данной технологии сосредоточено на идентификации людей. Seeing AI – это мобильное приложение, которое использует технологию искусственного интеллекта для распознавания и чтения коротких текстовых отрывков, документов, этикеток продуктов, штрихкодов, а также для описания людей и пейзажей, снятых камерой мобильного

телефона. Biped.ai – устройство, которое носится на плечах как рюкзак, оснащенный сверхширокоугольными 3D-камерами слева от груди, батареей за шеей и небольшим компьютером справа от груди. Biped анализирует окружающую среду, чтобы обнаруживать и прогнозировать препятствия на несколько секунд вперед. Затем он фильтрует соответствующую информацию и предупреждает владельца с помощью иммерсивных 3D-звуков, передаваемых через наушники. Устройство использует технологию автономного вождения от научно-исследовательского института Honda. Система также предоставляет инструкции GPS.

Все вышеперечисленные устройства используют подключение к облачным сервисам для обработки изображений. Предлагаемое новое решение состоит в применении технологии периферийного искусственного интеллекта Edge AI.

Устройство на базе технологии Edge AI включает в себя очки с камерой, распознающей видео в режиме реального времени, динамик, преобразующий визуальную информацию в аудио для передачи данных об окружающих объектах пользователю, а также микрофон для считывания команд пользователя. Благодаря использованию периферийных вычислений система не нуждается в связке с интернетом, подключении к удаленным серверам. Таким образом, огромные объемы данных пользователя не передаются во внешний мир, всё сохраняется внутри, что обеспечивает большую безопасность и конфиденциальность, экономию трафика и быстрые вычисления.

Дополнительным преимуществом является независимость устройства от качества связи и передачи данных, что предоставляет эффективный контроль над непрерывностью операций. В результате повышается производительность устройства и снижаются эксплуатационные расходы. Еще одним достоинством чипов на базе технологии Edge AI является небольшой размер, что позволит создать компактное устройство, удобное в эксплуатации [4].

В качестве аппаратного решения технологии Edge AI были выбраны нейрочипы KL520 производителя Kneron (рис. 2).

Рис. 2. Внешний вид нейрочипов KL520

Программное обеспечение Kneron включает в себя средства для эмуляции работы чипов с помощью виртуальных контейнеров, заранее созданные и оптимизированные реализации нейронных сетей и алгоритмов машинного обучения, а также библиотеки для взаимодействия с ними, что упрощает и ускоряет разработку системы [5]. Для обработки входного потока данных была выбрана легковесная нейросетевая модель YOLOv5.

С целью распознавания видео в реальном времени со встроенной камеры в прошивке необходимо настроить режим, в котором каждый кадр исходного потока является обособленным набором данных. Все кадры складываются в FIFO буфер. Для каждого из них формируется специальный дескриптор. В этом дескрипторе содержатся

параметры картинки, такие как размер и формат; адреса для чтения исходных данных и записи результата; адреса буферов и параметры для предпроцессинга и постобработки; идентификатор нейросети; флаг режима параллельной работы.

Распознавание каждого кадра изображения является независимым процессом, а режим параллельной работы позволяет запускать обработку нового кадра, не дожидаясь окончания работы NPU для предыдущего кадра. Поэтому для сопоставления исходных данных и результатов используется специальный идентификатор. Все результаты складываются в свой FIFO буфер, откуда происходит их считывание в другом процессе операционной системы.

На вход нейросети должно подаваться изображение размером 416×416 пикселей в формате RGB, однако этот размер изображения не соответствует ни одному из стандартных разрешений, которое можно получить со встроенной камеры. Для устранения этой проблемы были внедрены алгоритмы предобработки сырых данных с камеры, а также постобработки результатов YOLO в прошивку. Сам алгоритм постобработки уже применяется в наборе инструментов при квантизации, его можно найти в составе docker контейнера.

Для параллельного запуска процессов в KL520 предусмотрено два отдельных ядра Cortex-M4. Помимо этого, и сам нейроускоритель работает независимо, фактически как еще одно ядро. Все три компонента используют единое пространство быстродействующей памяти, интегрированной в чип.

Основное ядро (SCPU) занято получением и преобразованием изображений с камеры и отрисовкой обрамляющих прямоугольников. Вспомогательное ядро (NCPU) выполняет вычислительные задачи постобработки результатов YOLO. Нейроускоритель (NPU) занят прогоном нейросети.

В рамках задачи детектирования множества объектов на городских улицах используется набор данных MOT17 [6]. В работе устройства предполагаются два режима использования:

1. Автоматическое озвучивание предметов в кадре, важных для перемещения пользователя: светофоров, переходов, ям, лестниц, дверей, деревьев, столбов и других препятствий, а также пешеходов и транспортных средств. Приоритет при этом отдается объектам, находящимся ближе к пользователю, то есть с большей площадью внутри определительной рамки. Распознанные объекты в формате текста через нейроускоритель преобразуются в соответствующие аудио-инструкции, передающиеся в наушники пользователя.

2. Озвучивание по соответствующей команде нахождения конкретного предмета в окружающей обстановке.

Распознавание голосовых команд пользователя можно также организовать на нейроускорителе KL520. Для записи звука используется микрофон с интерфейсом I2S (рис. 3), подобный тому, что встроен в мобильные телефоны, и вход для него присутствует в чипе KL520.

Рис. 3. Микрофон с интерфейсом I2S

При включении питания микрофон передает данные автоматически в формате 24 бит с частотой дискретизации 16 кГц. Длительность интервалов сигнала для обработки составляет 4 секунды.

Сырой звуковой сигнал, записанный с частотой дискретизации 16 КГц, не подходит для анализа и обработки в исходном виде, поэтому его необходимо преобразовать. Для этой цели используется метод представления звука в виде набора мел-кепстральных коэффициентов. В результате сигнал длительностью 4 секунды, записанный с частотой дискретизации 16 КГц, преобразуется в набор мел-кепстральных коэффициентов в виде матрицы размером 40x320 чисел в формате float32, где 40 – количество коэффициентов, а 320 – количество участков (рис. 4).

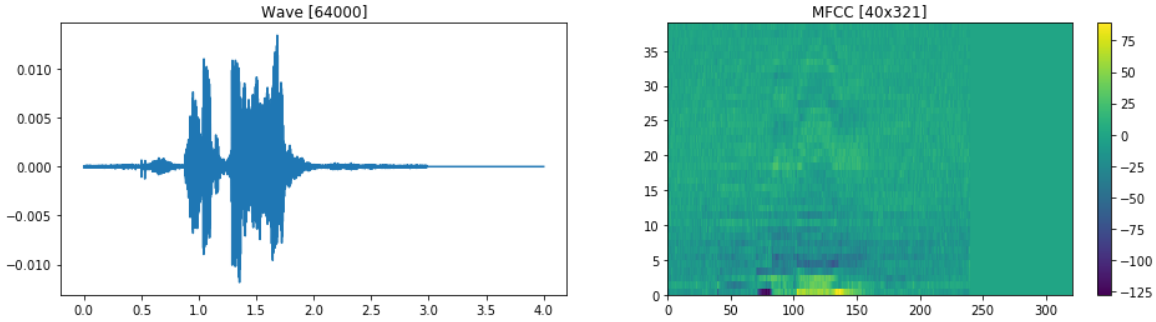


Рис. 4. Представление исходного сигнала и его преобразования
в набор мел-кепстральных коэффициентов

Если звук представлен в 24-битном формате, то его исходный размер составляет 192000 байт, а размер матрицы коэффициентов уже 51200 байт. То есть, данные сжимаются почти в 4 раза с незначительными потерями. Для работы на KL520 используются квантизованные данные в формате int8, что позволяет дополнительно сжать их до размера входа в 12800 байт, что уже соответствует сжатию в 15 раз.

Полученный набор мел-кепстральных коэффициентов подвергается нормализации, результат которой в виде картинки будет использоваться в качестве основы для распознавания нейросетью. В данной работе была использована модифицированная нейросеть ResNet18, где последний полносвязный слой был заменен на новый, выход которого представлен одномерным вектором с длиной, равной количеству команд.

На рисунке 5 представлен полученный опытный образец устройства. На экране выведена одна из 38 возможных голосовых команд пользователя.

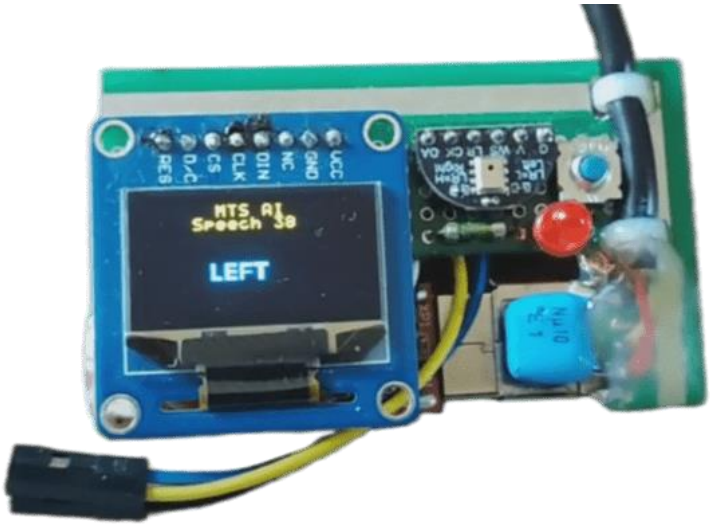


Рис. 5. Устройство для распознавания голосовых команд

В результате работы была спроектирована система, представляющая из себя программно-аппаратный комплекс для распознавания изображений с помощью сверточной нейронной сети YOLOv5 с аппаратной поддержкой в виде микроконтроллера со встроенным нейроускорителем Kneron KL520. При помощи настройки взаимодействия с прошивкой чипа и отладки процесса обработки изображений нейронной сетью был получен прототип программы, работающий как с записанными видео, так и в режиме реального времени. Также на базе нейрочипов KL520 и нейросети ResNet18 был создан опытный образец устройства для распознавания голосовых команд пользователя.

Литература

1. Morrison C., Cutrell E., Grayson M. PeopleLens // Interactions. 2021. V. 28. № 3. P. 10–13.
2. Description of talking camera app for those with a visual impairment [Электронный ресурс]. Режим доступа: https://www.microsoft.com/en-gb/ai/seeing-ai (дата обращения: 23.01.2023).
3. Description of smart copilot for blind and visually impaired people [Электронный ресурс]. Режим доступа: https://www.biped.ai (дата обращения: 23.01.2023).
4. Wang X., Han Y., Victor C.M. Convergence of Edge Computing and Deep Learning: A Comprehensive Survey // IEEE Communications Surveys & Tutorials. 2020. V. 22. № 2. P. 869–904.
5. Toolchain Manual Overview [Электронный ресурс]. Режим доступа: https://doc.kneron.com/docs/#toolchain (дата обращения: 10.09.2022).
6. Multiple Object Tracking Benchmark [Электронный ресурс]. Режим доступа: https://motchallenge.net/data/MOT17 (дата обращения: 15.11.2022).

УДК 004.896

ПОХОД К СОЗДАНИЮ ИНТЕЛЛЕКТУАЛЬНОЙ СИСТЕМЫ ПОСТРОЕНИЯ 3Д МОДЕЛИ ПОМЕЩЕНИЙ ПРИ ПОМОЩИ БЕСПИЛОТНЫХ ЛЕТАТЕЛЬНЫХ АППАРАТОВ

Вандышев К.А.[1], Горшков Р.Д.[1], Коныхов В.С.[1]
Научный руководитель – Дудкин А.С.[1]
1 – Военно-космическая академия имени А.Ф. Можайского
e-mail: gorshkov-romka@inbox.ru

Аннотация
В статье рассматривается применение триангуляции, навигации по меткам и SLAM навигации, а также совместном применении указанных методов для построения моделей зданий и сооружений за счёт использования автономного роя беспилотных летательных аппаратов, управляемых искусственным интеллектом, что обосновано необходимостью в повышении оперативности сбора и обработки информации.
Ключевые слова
SLAM, RRT, навигация, система рой, искусственный интеллект, самоорганизующийся рой, адаптивные системы.

Введение

В современных условиях необходимость своевременного обеспечения информацией без людских потерь многократно увеличилась. Существует множество видов получения информации при помощи навигации автономных беспилотных аппаратов и роботов.

Роевое управление при помощи искусственного интеллекта представляет собой новый подход в управлении беспилотным летательным аппаратом. Данное решение позволяет быстро получать целостное представление об исследуемом объекте за счет поступления обработанной информации с нескольких аппаратов, в то время как использование одного управляемого аппарата требует постоянного контроля со стороны оператора [1-4].

Основа решения

Основной задачей проекта является возможность быстрого и автономного, а именно без участия человека, построения 3D модели здания как снаружи, так и изнутри. Оператору необходимо лишь запустить необходимое количество дронов на исследуемый объект (рис. 1).

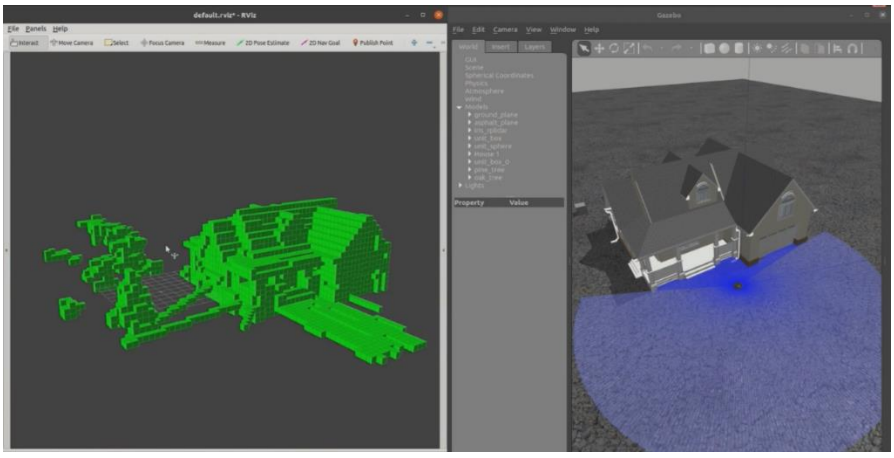


Рис. 1. Построение 3D модели здания.

На сегодняшний день реализован проект, который на базе построенной карты с помощью алгоритма SLAM производит разведку зданий, исследует труднодоступные места с одновременным построением карты пройденного места.

Основными результатами проекта является возможность быстрого и автономного, а именно без участия человека, построения 3D модели здания как снаружи, так и изнутри. Оператору необходимо лишь запустить необходимое количество дронов на исследуемый объект. Благодаря искусственному интеллекту между дронами равномерно распределиться нагрузка (система Рой). В случае исследования здания изнутри дроны располагаются так, чтобы всегда была возможность передачи всей собранной информации, так как само здание может не пропускать связь до главного центра управления. Вся информация, которую собрал один дрон, дублируется на все остальные, для исключения потери собранных сведений в случае выхода из строя.

Для повышения стабильности передачи полученной информации в центр управления также используется автономный ДРОН* (рис. 2), находящийся над исследуемым объектом на расстоянии 1 км.

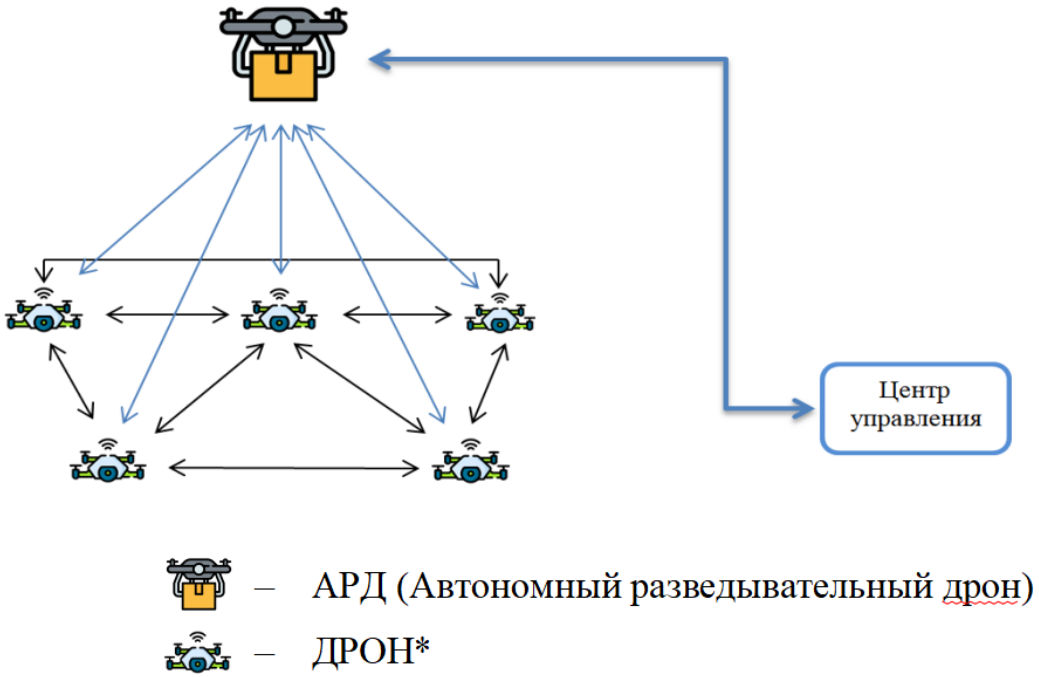


Рис. 2. Схема передачи информации на центр управления

Данные БЛА могут быть также использованы и для корректировки уже имеющихся схем или моделей. Например, есть эталонная схема здания и вследствие аварии произошел завал этажей. При подаче системе эталонной схемы дрон уже будет знать, как исследовать здание и налету будет производить коррекцию схемы при несостыковке реальных данных с эталонными.

Для передачи данных между БЛА используется модуль WIFI, в дальнейшем будет использована другая сеть. Обеспечение безопасной передачи данных осуществляется за счёт использования криптографического алгоритма «КУЗНЕЧИК».

В качестве основы АРД взят отечественный микропроцессор СКИФ, позволяющий избавиться от недекларированных возможностей и значительно снизить зависимость от иностранного производства.

На борту АРД находится множество интеллектуальных систем анализа, защиты и синхронизации, позволяющие наиболее эффективно производить разведку зданий и сооружений, исключая угрозы перехвата, захвата и ложного изменения получаемой информации, а также физического проникновения в программную часть БЛА.

В данной системе используется машинное обучение с подкреплением, на основе метода, при котором происходит обучение модели, которая не имеет сведений о системе, но имеет возможность производить какие-либо действия в ней и получать состояния среды. На каждое действие модели возвращается состояние среды и вознаграждение, на основании которых агент учится, пытаясь максимизировать общее вознаграждение. Обучение производилось в среде Unity, где каждый раз генерировалось случайное здание и модели необходимо было пройти его полностью. В случае столкновения дрона с препятствиями симуляция прекращалась и модели присуждалась отрицательная награда. Также в целях ускорения построения карты после каждого выбранного действия у модели отнимались очки вознаграждения.

В виртуальной среде реализован агент, способный управлять дроном для достижения поставленной цели. Также реализован алгоритм автоматизированного построения пути RRT для построения маршрута движения дрона после выполнения задачи. На основании проведенных экспериментов по обучению модели ИИ в виртуальной среде можно с уверенность сказать, что при переходе из симуляции в реальную среду агент успешно справляется с чуть измененными условиями и достигает хороших результатов уже в настоящем мире.

Для осуществления безопасной передачи данных используется алгоритм «КУЗНЕЧИК».

Результаты

1. Реализованы алгоритмы RRT и SLAM.
2. Создан прототип беспилотного летательного аппарата.
3. Реализован механизм синхронизации «рой» для повышения эффективности выполнения задачи.
4. Реализовано построение 3D модели зданий и сооружений как снаружи, так и изнутри.

Заключение

Для данной разработки имеется большой спектр задач. С её помощью можно исследовать сооружения, подвергшиеся разрушению, объекты, о который ничего не известно и куда при экстремальных условиях человек не может добраться, либо затраченное время не оправдает затраченных усилий и средств. Поэтому данный комплекс активно развивается и тестируется в стенах нашей академии.

Литература

1. Krul S. et al. Visual SLAM for indoor livestock and farming using a small drone with a monocular camera: A feasibility study //Drones. 2021. Т. 5. №. 2. С. 41.
2. Heukels F.R. Simultaneous Localization and Mapping (SLAM): towards an autonomous search and rescue aiding drone: дис. University of Twente. 2015.
3. Park J.K., Chung T.M. Boundary-RRT* algorithm for drone collision avoidance and interleaved path re-planning //Journal of Information Processing Systems. 2020. Т. 16. №6. С. 1324-1342.
4. Killian L., Backhaus J. Utilizing the RRT*-Algorithm for Collision Avoidance in UAV Photogrammetry Missions //arXiv preprint arXiv:2108.03863. 2021.

УДК 004.8

ГЕНЕРАЦИЯ ИЗОБРАЖЕНИЙ НА ОСНОВЕ СЕМАНТИЧЕСКИ НАГРУЖЕННЫХ БОЛЬШИХ ХУДОЖЕСТВЕННЫХ ТЕКСТОВ

Вологина Е.А.[1] (студент)
Научный руководитель – к.э.н., доцент Тимченко О.В.[2]
1 – Пятигорский государственный университет
e-mail: sweeetdude@yandex.ru

Аннотация
Данная работа включает в себя исследование проблемы извлечения смысла из большого объема текста, которая является важной задачей в области обработки естественного языка, а также сравнительный анализ широко известных технологий генерации изображений по расширенному текстовому описанию, и выявлены проблемы, связанные с недостаточным учетом семантической структуры текста и средств выразительности.
Ключевые слова
Нейронная сеть, датасет, семантика, библиотеки Python, архитектура.

На сегодняшний день существует огромное количество различных нейронных сетей, которые используются для создания изображений на основе введенного текста. Однако до сих пор остается проблема полного и осмысленного понимания больших объемов текста, осложненных различными средствами выразительности и стилистическими приемами, что необходимо для успешной генерации изображений.

Нейронные сети - это компьютерные модели, которые используются для решения задач, таких как распознавание образов, классификация, генерация текста и других. Был проведен анализ наиболее популярных генераторов изображений, основанных на текстовых данных. Несмотря на то, что нейронные сети достигли значительных успехов во многих областях, они все еще недостаточно учитывают семантическую структуру текста и не могут полностью понимать средства выразительности. Например, одно и то же слово может иметь разные значения в разных контекстах, и это может быть понятно только человеку, который обладает культурным и лингвистическим багажом.

Кроме того, нейронные сети работают на основе обучения на больших объемах данных. Они могут изучать паттерны и связи в данных, но они не могут понимать эмоциональную окраску и интонацию, которые могут изменить значение фразы или слова. Также следует отметить, что нейронные сети не обладают общим знанием о мире и культуре, которые необходимы для полного понимания языка и выражения. Например, метафоры и идиомы могут быть непонятны нейронной сети, если она не обучалась на достаточно широком диапазоне текстов, чтобы понимать эти явления [1].

Одной из основных проблем при создании этой нейросети является связь между текстом и изображением. Для того чтобы нейросеть могла генерировать изображения на основе текстового описания, ей необходимо понимать смысл текста и уметь переводить его в соответствующие изображения. Для этого можно использовать технологии обработки естественного языка, такие как модели глубокого обучения на основе рекуррентных нейронных сетей (RNN), сверточных или трансформеров.

Другой важной задачей является создание базы данных изображений, соответствующих семантическому содержанию текстов. Для этого можно использовать различные методы, например, автоматический сбор изображений из интернета или ручной отбор изображений на основе семантической близости с текстом.

Этот подход позволяет создавать высококачественные изображения, которые соответствуют содержанию текста, используя технологию глубокого обучения.

Для того, чтобы научить нейросеть различать различные фигуры речи и средства выразительности, нужно подготовить набор данных с образцами текстов, которые

содержат их. Эти данные могут быть собраны из разных источников, например, из литературных произведений.

Затем нужно разметить этот набор данных, указав, какие литературные приемы присутствуют в каждом образце текста. Эта разметка должна быть выполнена вручную, так как на данный момент не существует специализированных автоматических инструментов разметки семантически нагруженного текста.

После этого, используя нейронные сети и алгоритмы машинного обучения, можно обучить модель классификации, которая будет способна определять, какие фигуры речи и средства выразительности содержатся в новых текстах. Для этого нужно выбрать соответствующую архитектуру нейронной сети и определить параметры обучения, такие как функцию потерь, оптимизатор и гиперпараметры.

В процессе обучения модель будет получать обратную связь от набора данных и настраиваться на распознавание различных фигур речи и средств выразительности. Когда модель будет достаточно точной на обучающем наборе данных, ее можно протестировать на новых текстах, чтобы оценить ее точность и эффективность.

Для разработки данной нейросети было было решено использовать синергию двух ведущих нейросетей – Stable Diffusion и GPT-3.

Для создания изображений с использованием GPT-3 и Stable Diffusion, необходимо сначала обучить модель на большом наборе художественных текстов. Эти тексты могут быть различных жанров, от романов до поэзии, и должны содержать сильную семантическую нагрузку.

Для того, чтобы GPT-3 могла сгенерировать текст сказки, на ее вход необходимо подать начальный текст, который будет являться контекстом истории. Этот текст может быть каким-то прологом, описанием героя, сценой, описанием местности, в которой происходят события и т.д.

На основе этого начального текста GPT-3 использует свою внутреннюю модель, основанную на технологии глубокого обучения, чтобы понимать, какие слова, фразы и предложения могут логически следовать за данным контекстом и какие события могут происходить в истории. Затем она генерирует текст, используя предыдущий контекст и свою внутреннюю модель, которая учитывает грамматику, стиль и семантику текста, чтобы создать новую часть истории.

Затем, этот текст подается на вход Stable Diffusion, чтобы создать иллюстрации, которые соответствуют сюжету и содержанию сказки. Использование Stable Diffusion позволяет генерировать иллюстрации, которые имеют художественное качество, и которые визуально соответствуют содержанию сказки.

А теперь поподробнее рассмотрим архитектуру данных нейросетей.

Архитектура GPT-3 состоит из стека трансформер-блоков. GPT-3 также использует специальные механизмы для управления памятью и обработки последовательностей произвольной длины. В отличие от предыдущих версий GPT, GPT-3 использует более сложный механизм многоголового внимания, который позволяет модели "учиться" на основе более широкого контекста.

Одной из главных особенностей GPT-3 является ее способность генерировать тексты с высокой степенью связности и разнообразия, что делает ее полезной для многих задач обработки естественного языка, включая автогенерацию текстов, ответы на вопросы, перевод и другие.

Архитектура Stable Diffusion (SD) – это метод для моделирования вероятностных распределений, который основан на применении оператора диффузии к начальной плотности распределения.SD представляет собой генеративную модель, которая может использоваться для сэмплирования из заданного распределения. В отличие от классических методов, которые требуют аналитической формулы для плотности распределения, SD может использоваться для моделирования произвольного распределения [2].

Разработка данной нейронной сети производится на языке Python.

Так как Python обладает широкой функциональностью и богатой экосистемой библиотек, которые упрощают и ускоряют процесс разработки нейронной сети.

PyTorch – это библиотека для глубокого обучения, которая предоставляет инструменты для создания и обучения нейросетей. Она широко используется в области компьютерного зрения и обработки естественного языка, и поддерживает модели генеративной моделирования, включая GPT-3 и Stable Diffusion.

Transformers – это библиотека, основанная на PyTorch, для создания и обучения моделей обработки естественного языка, включая модели генеративной моделирования, такие как GPT-3.

DALL-E API – это библиотека для генерации изображений на основе текстовых описаний. Она использует технологии глубокого обучения, такие как GPT-3 и Stable Diffusion, для создания реалистичных изображений на основе вводных текстовых описаний.

NumPy – это библиотека для работы с массивами данных и математическими операциями над ними. Она широко используется в машинном обучении и обработке изображений для обработки данных и предварительной обработки изображений.

OpenCV – это библиотека для обработки изображений, которая предоставляет инструменты для чтения, записи, обработки и анализа изображений. Она может использоваться для предварительной обработки и анализа изображений, созданных нейросетью [4].

Что касается сбора датасета, то к этому нужно подходить с осторожностью, так как большая часть работы нейронной сети зависит от качества набора данных, а также от его количества и разнообразия [3].

Сбор датасета для нейросети, генерирующей изображения на основе семантически нагруженных художественных текстов, имеет свои особенности.

Для того, чтобы нейросеть могла создавать разнообразные изображения, необходимо предоставить ей достаточно большой датасет. Рекомендуется использовать не менее 10,000 художественных текстов, которые содержат семантически нагруженные описания сцен и персонажей:

1. При сборе датасета был определен стиль изображений, который был использован в нейросети. Например, можно использовать реалистичные или стилизованные изображения, цветовую палитру и стиль рисования.
2. Произведена фильтрация и очистка данных. Это включает удаление повторяющихся изображений, проверку качества изображений и удаление ненужных данных.
3. Необходимо учитывать авторские права на художественные тексты и изображения. Лучше всего использовать материалы, которые являются свободными от авторских прав, либо имеют соответствующие разрешения на использование.
4. Важно было разметить каждое изображение с соответствующим текстовым описанием, чтобы модель могла понимать связь между текстом и изображением. Для этого можно использовать различные инструменты разметки, такие как VGG Image Annotator (VIA), LabelImg и т.д.
5. Для удобства работы с датасетом были созданы аннотации и метаданные, которые помогают быстро находить нужные художественные тексты и соответствующие им изображения по определенным критериям, таким как персонажи, сюжетные линии, место действия и т.д.
6. Для создания более разнообразного датасета можно использовать художественные тексты на разных языках. Также можно использовать разные жанры художественной литературы, чтобы получить более широкий спектр сцен и персонажей.
7. Для улучшения качества датасета были использованы техники data augmentation, такие как повороты, масштабирование и изменение яркости и

контрастности изображений. Это позволило получить более разнообразные изображения на основе описаний из художественных текстов.

8. И последнее, что было сделано это разделение датасета на обучающую, проверочную и тестовую выборки.

Для оценки качества работы нейросети были выбранны данные метрики оценки качества сгенерированного изображения:

1. Inception Score (IS): Эта метрика оценивает качество генерации изображений, используя Inceptionv3, обученную на наборе данных ImageNet, для измерения разнообразия и качества изображений. Она оценивает сходство распределения классов входных изображений с распределением классов сгенерированных изображений.
2. Structural Similarity Index (SSIM) - метрика, которая измеряет сходство между двумя изображениями на основе яркости, контрастности и структуры и контента. Она может быть полезным инструментом для оценки качества изображений, сгенерированных моделью [5].

Кроме автоматических метрик качество генерации изображений по сказкам народов мира оценивали эксперты-лингвисты.

Разработанный прототип получил экспериментальную оценку (таблица). Тестовая выборка составила 357 сказок и 4284 иллюстрации к ним.

Таблица

Результаты метрик

Эпоха обучения	Inception Score (IS)	Structural Similarity Index (SSIM)	Экспертная оценка точности генерации текста	Экспертная оценка точности генерации изображения
15	0,7412	0,7412	0,5205	0,4115
30	0,7937	0,7855	0,5439	0,4533
45	0,8391	0,8267	0,5802	0,4808
60	0,8716	0,8698	0,6101	0,5007

Однако, важно понимать, что результаты эксперимента могут быть неоднозначными и требуют дополнительного анализа. Кроме того, стоит также учитывать, что результаты эксперимента могут зависеть от выбранной метрики и критерия оценки качества.

Таким образом, можно сделать вывод, что к настоящему времени генеративные нейронные сети пока что не в состоянии корректно обрабатывать и воспринимать большие специфические тексты. Одним из вариантов решений видится не только обучение нейронных сетей на больших объемах текста с использованием различных стилистических приемов, чтобы улучшить их способность к распознаванию и пониманию различных форм и стилей текста, но и использование комбинированных моделей, объединяющих различные типы нейронных сетей, чтобы создавать более точные и выразительные изображения на основе текстовых описаний.

Литература

1. Батура Т.В. Семантический анализ и способы представления смысла текста в компьютерной лингвистике // Программные продукты и системы. 2016. №4 (116). [Электронный ресурс]. Режим доступа: https://cyberleninka.ru/article/n/semanticheskiy-analiz-i-sposoby-predstavleniya-smysla-teksta-v-kompyuternoy-lingvistike (дата обращения: 18.03.2023).
2. Understand the tech: Stable diffusion vs GPT-3 vs Dall-E [Электронный ресурс]. Режим доступа: https://www. technologika.ru/ blog/how-to- create-great-dataset (дата обращения: 01.03.2023).

3. О важности датасета и о том, как сделать его лучше [Электронный ресурс]. Режим доступа: https://www. technologika.ru/ blog/how-to-create- great-dataset (дата обращения: 15.03.2023).
4. 10 библиотек Python для машинного обучения и искусственного интеллекта [Электронный ресурс]. Режим доступа: https://vc.ru/u/1389654-machine-learning/ 588349-10- bibliotek- python-dlya- mashinnogo- obucheniya-i- iskusstvennogo-intellekta?comments (дата обращения: 15.03.2023).
5. Абросимов К.И., Львутина Т.В., Суркова А.С. Метрики качества генеративных моделей/ Абросимов К.И., Львутина Т.В., Суркова А.С. [Электронный ресурс]. Режим доступа: https://www.graphicon.ru/html/2021/papers/01/Volume1_paper_13.pdf (дата обращения: 02.03.2022).

УДК 004.8

ПРОГНОЗИРОВАНИЕ ЦЕН АКЦИЙ НА БИРЖЕ С ИСПОЛЬЗОВАНИЕМ РАЗРОЗНЕННЫХ ИСТОЧНИКОВ

Газизулин А.Ф.[1]
Научный руководитель – к.т.н., доцент Ключев А.О.[1]
1 – Университет ИТМО
e-mail: afgazizulin@itmo.ru

Аннотация
В данной работе рассматривается вопрос прогнозирования цен акций на бирже. Предлагается использовать эффективные методы машинного обучения, такие как LR (Logistic regression), SVM и ANN, которые будут обучаться на данных из разрозненных источников. Данные формируются на основе технических индикаторов и политических, экономических новостей и новостей по отдельным компаниям, для которых проводится анализ.
Ключевые слова
машинное обучение, биржа, цены акций, обработка текста, большие данные, новостные источники.

Инвестиции и торговля на фондовом рынке могут быть сложными, но при правильном анализе и возможности прогнозирования приносят прибыль. На цены акций влияют как несколько технических индикаторов, так и множество других источников информации: новости, аналитика, настроение пользователей и др. В теории финансов есть гипотеза эффективного рынка, она гласит, что цены активов не могут полностью зависеть от устаревшей информации, а рыночные цены реагируют на новую информацию, например статьи финансовых новостей. Чтобы это проверить были рассмотрены работы других исследователей.

Так, в работе Bin Weng и др. использовались данные из открытых онлайн-источников с характеристиками, извлеченными из котировок и технических индикаторов [1]. При этом онлайн-источники должны состоять в базе финансовой экспертной системы. Исследователи использовали следующие алгоритмы: один из алгоритмов ANN (artificial neural network) – многослойный перцептрон MLP (multilayer perceptron), дерево принятия решений DT (decision tree) – метод C5 вместо ID3, C4.5 или C&RT, а также метод опорных векторов SVM (support vector machines) с использованием функции ядра RBF (radial basis function).

Shilpa Gite и др. прогнозировали изменение цены по анализу настроений исходя из новостных сводок и с обработанными данными из Yahoo Finance, а также с рыночными показателями цен OHLC (open high low close) [2]. Для обучения использовались алгоритмы глубокого обучения LSTM (long short term memory), лучший для долгосрочных зависимостей, а также XAI (explainable artificial intelligence) чтобы модель давала не только прогноз, но и объясняла почему он получился именно таким.

В собранном наборе исследователей Abdullah H. Al-Nefaie и Theyazn H. H. Aldhyani включены стандартные входные данные – цены из котировок и новые данные – название компаний, торговый сектор, изменение фондового рынка в последний день и количество торгов на рынке [3]. На их основе сравнивались эффективности алгоритмов MLP и LSTM.

Salvatore M. Carta и др. в своём исследовании подготовили только текстовые данные, состоящие из международных статей и газет из набора Dow Jones DNA, а для обучения модели на подготовленном лексиконе используют дерево принятия решений без уточнения метода [4].

Manushi Munshi и др. два набора данных с платформы Kaggle: «Все Индийские IPO» и «Данные по IPO Индии», содержащие дату, размер эмиссии (общее количество

акций) несколько категорий инвесторов и др. [5]. Используемые методы обучения в данной работе были регрессор KNN (k-nearest neighbors); дерево принятия решений; радиочастотный регрессор; регрессор XGBoost (eXtreme gradient boosting).

Во всех рассмотренных работах использовались стандартные методы предобработки данных, методы отбора предикторов и алгоритмы отбора гиперпараметров. Однако в данных работах не рассматривается корреляция между новостями и целевой переменной, поэтому была выдвинута гипотеза, что новостные статьи в совокупности с данными по ценам акций дают лучшую точность предсказания для модели машинного обучения.

Была сформулирована цель: выявить корреляцию между новостными источниками и ценой акции.

Также были поставлены следующие задачи для её достижения:

- использовать различные типы данных, связанных с предметной областью в качестве обучающего набора;
- для преобразования данных использовать совокупность приёмов и методов, которые перекрывают недостатки методов из сводной модели;
- обучить несколько моделей на выбранных алгоритмах, проанализировать модели с помощью совокупности (или отдельных) методов, чтобы дать достаточно точную оценку.

Данные

В результате проведения обзора работ у зарубежных коллег, была сформулирована следующая проблематика: необходимо предсказывать вариацию цены акции компании, а также интерпретировать результаты предсказания в удобной для пользователя форме.

Для решения данного вопроса предлагается методика, в которой в качестве исходных данных будут выступать:

- ценах OHCL, по интересующей компании;
- новости компаний;
- новости политики;
- новости экономики.

Котировки загружались программой-парсером с сайта Финам при использовании шаблона запроса. Новости собирались также скриптом на сайте агрегатора новостей, где сразу делились на категории, указанные выше, при изменении параметра запроса.

Предобработка данных

Для обработки и анализа данных из котировок использовались как стандартные методы первичной обработки, например такие как отчистка или замена пропущенных значений, а также преобразование данных (извлечение дня недели из даты, подсчёт стандартного отклонения и т.д.), так и методы отбора предикторов по их значимости: Random Search. Далее будет указано, что данный метод отбора отрицательно влияет на конечный результат и этот этап в дальнейшем был отброшен.

Первоначальный набор преобразованных данных с получасовым тиком содержал: цену открытия/закрытия, объём торгов, максимальную цену открытия, минимальную цену открытия, новостные статьи 3-х категорий, а также преобразованные в числовой формат дата, неделя, месяц, час и минуты. Последнее необходимо для объединения котировок и новостей, т.к. новости не так регулярны, как котировки, то после объединения появляются незаполненные места, поэтому они заполняются последней новостью, т.е. как новость вышла то она на протяжении какого-то времени влияет на цену акции.

Для текстовых данных будет использоваться следующие методы обработки: токенизация, удаление/добавление стоп-слов, удаление часто/редко встречающихся слов, стемминг.

Также для текстовых данных был проведён частотный анализ с помощью TF-IDF (term frequency – inverse dense frequency). Текст векторизовался и его размерность была уменьшена с помощью метода главных компонент PCA (principal component analysis) для отсеивания наименее полезных слов в словаре.

После предобработки данных, была построена тепловая корреляционная карта. По ней определялись данные, которые слабо влияют на целевую переменную, и которые в процессе эксперимента можно исключить из выборки. Далее на рисунке 1 представлена сжатая корреляционная карта, где кроме данных по котировкам представлены данные по 4 главным компонентам (x0,x1,x2,x3).

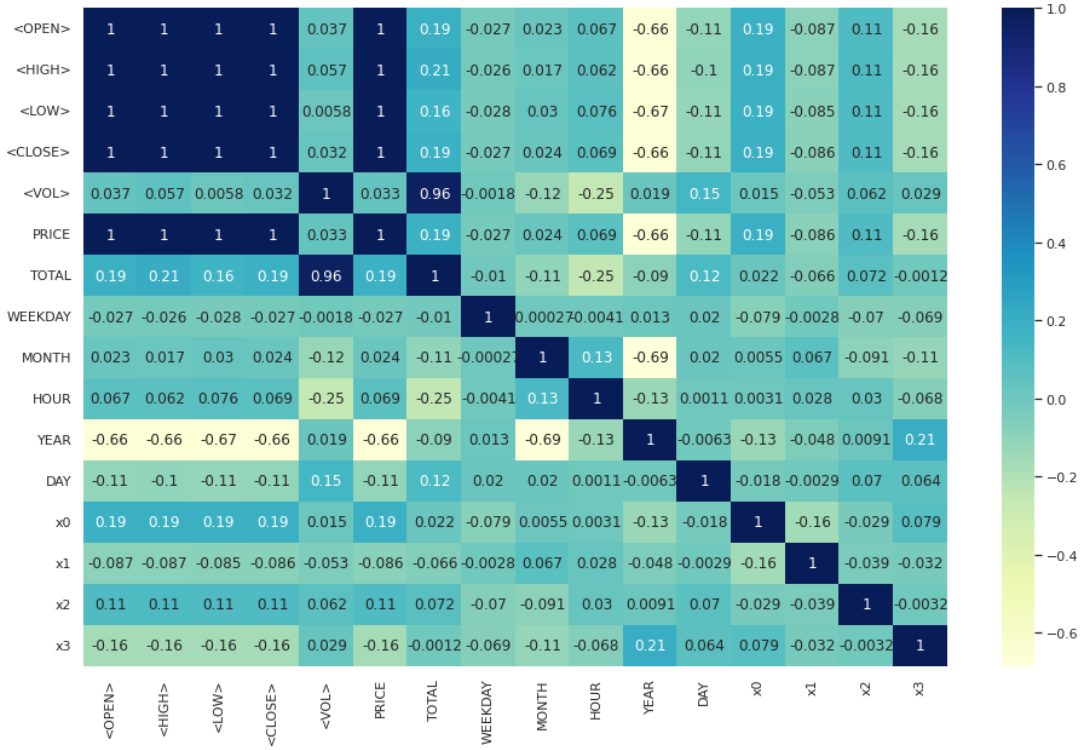


Рис. 1. Тепловая карта для набора данных

На карте видно, что преобразованные текстовые данные имеют с целевой переменной PRICE как положительную, так и отрицательную корреляцию.

Обучение модели

В качестве методов обучения были выбраны несколько алгоритмов с возможностью интерпретирования модели: LR (linear regression), самый простой из методов искусственных нейронных сетей ANN – многослойный перцептрон MLP, а также SVM/ Последний всегда показывает средние результаты и может использоваться в качестве опорного метода. Все обученные модели оценивались стандартными методами, такими как матрица ошибок (accuracy), MSE (mean squared error), RMSE (root mean square error) и MAE (mean absolute error).

Далее для трех выбранных алгоритмов с помощью модуля случайного леса подбирались гиперпараметры, применение которых к модели дало наилучший прирост точности. Такие параметры были выбраны конечными для каждого алгоритма:

- для линейной регрессии были установлены параметры регуляризации $\lambda = 0.67$ и скорость обучения $\alpha = 1.0$;
- для MLP activation='identity', alpha=0.28, learning_rate='adaptive', random_state=42;
- для метода опорных векторов это C = 9, gamma 0.0001, kernel='rbf'.

Результаты

В результате обучения и тестирования моделей, лучшие результаты показали модели с алгоритмами SVM и линейной регрессии с точностью 55% и 53%. С ANN был получен результат в 50 %. Для их обучения и тестирования из набора данных были исключены столбцы минимальной цены закрытия и выбором 6 главных компонент PCA для TF-IDF словаря новостных статей. Далее будут представлены результаты только модели SVM, по большей части его результаты практически совпадают с результатами LR.

Так, на рисунке 2 представлен график ошибок модели SVM со средней ошибкой равной 0.00908 и стандартной ошибкой равной 0.083.

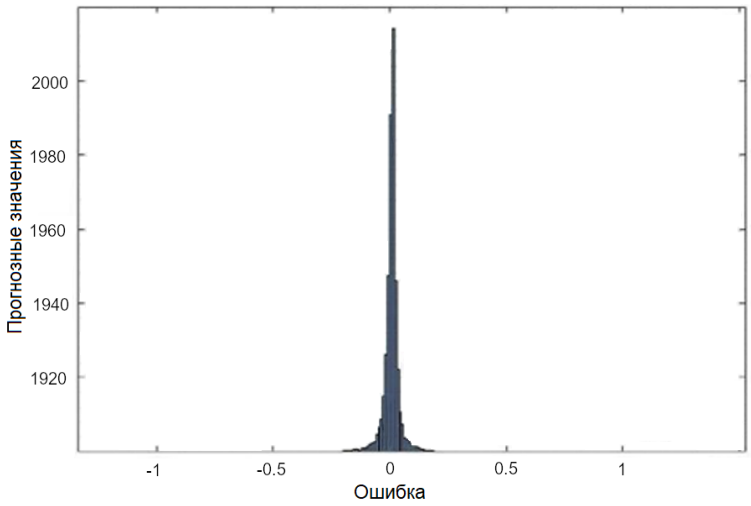


Рис. 2. Гистограмма ошибок SVM

В результате конечная модель SVM обладала следующими показателями точности и accuracy=55% и коэффициентом детерминации R^2=0,8801. Ниже на рисунке 3 представлен график, с двумя кривыми. Голубому цвету соответствует кривая, построенная по данным реальных цен, красному – прогнозируемая цена. Период времени взят с 1 ноября 2022 по 14 февраля 2023. Заметно, что с 10.01.23 график начинает вести себя непредсказуемо, т.к. значения реальных цен с этой отметки времени не использовались при обучении модели, поэтому дальше кривая показывает тенденцию, но не точные значения.

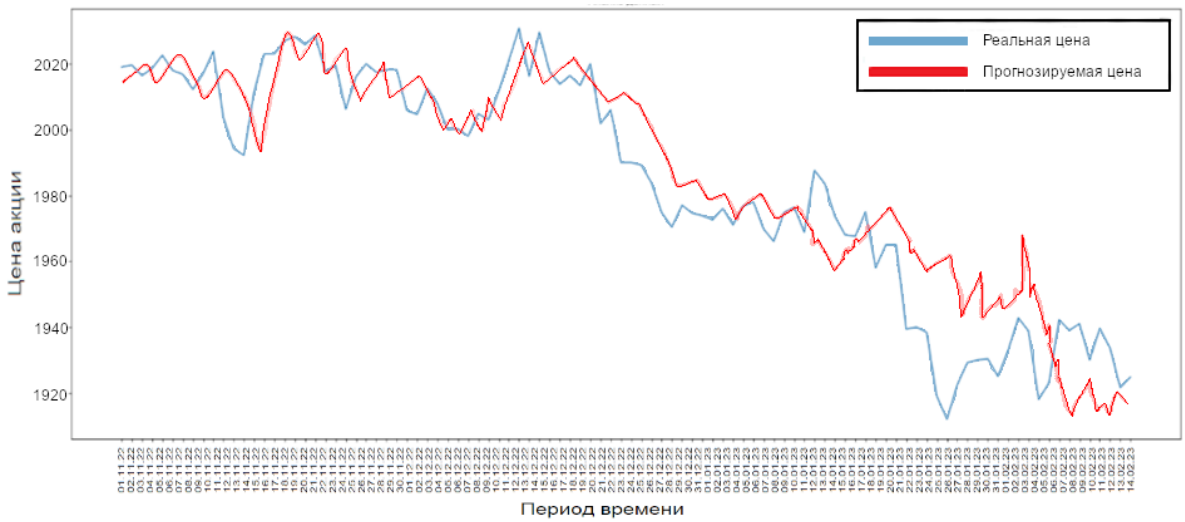


Рис. 3. Прогнозные значения для модели SVM

Заключение

В ходе проведения исследования была выявлена корреляция между текстом новостей и ценой акции. В результате обучения трёх моделей на данных из разрозненных источников, была выявлена одна с лучшими показателями accuracy=55 % и R2=0,880. Таким образом, использование разрозненных источников в качестве входных данных для модели МО позволило получить показатели точности прогнозирования цены акции более 50 %.

Литература

1. Weng B., Ahmed M. A., Megahed F. M. Stock market one-day ahead movement prediction using disparate data sources // Expert Systems with Applications. 2017. Т. 79. P. 153-163.
2. Gite S., Khatavkar H., Kotecha K., Srivastava S., Maheshwari P. & Pandey N. Explainable stock prices prediction from financial news articles using sentiment analysis // PeerJ Computer Science. 2021. Т. 7, P. 1-21.
3. Al-Nefaie A.H., Aldhyani T.H.H. Predicting Close Price in Emerging Saudi Stock Exchange: Time Series Models // Electronics. 2022. Т. 11. №. 21. P. 3443.
4. Carta S.M., Consoli S., Piras L., Podda A.S. & Recupero D.R. 2021. "Explainable Machine Learning Exploiting News and Domain-Specific Lexicon for Stock Market Forecasting", IEEE Access, vol. 9, pp. 30193-30205.
5. Munshi M. et al. Artificial Intelligence and Exploratory-Data-Analysis-Based Initial Public Offering Gain Prediction for Public Investors // Sustainability. 2022. Т. 14. №. 20. P. 13406.

УДК 004.048

МЕТОДЫ ВЫДЕЛЕНИЯ РЕКЛАМЫ И УТОЧНЕНИЯ ЕЁ КАТЕГОРИЙ НА ОСНОВЕ АНАЛИЗА НЕСТРУКТУРИРОВАННЫХ ДАННЫХ СОЦИАЛЬНЫХ СЕТЕЙ

Дощенко А.И.[1] (студент)
Научный руководитель – младший научный сотрудник, аспирант Филатова А.А.[1]
1 – Университет ИТМО
email: annadoshenko@mail.ru

Аннотация
В данной работе были разработаны методы обнаружения рекламных публикаций, размещённых в городской среде, с помощью данных социальных сетей. При помощи различных моделей тематического моделирования, таких как BigARTM, BERTopic и Zero Shot удалось научиться отделять рекламные посты от не рекламных, а затем определять более узкие тематики. Разработанные методы повысили эффективность выделения рекламы по категориям, что в дальнейшем будет использовано для определения степени эффективности данных публикаций и разработки методов рекомендаций для размещения рекламы в социальных сетях.
Ключевые слова
Обнаружение рекламы, тематическое моделирование, анализ социальных сетей, разметка данных, извлечение информации.

Социальные сети – это платформа для публикации большого количества разнообразной информации, они являются одним из основных способов общения и обмена информацией в окружающем нас мире. Многие компании и частные предприниматели используют их для продвижения своих товаров и услуг. Социальные сети с изображениями и возможностью отмечать геолокацию поста являются особенно популярными, так как они позволяют отслеживать распространение публикаций на карте, видеть и анализировать, где их больше и насколько они эффективны. То же самое можно применить и к рекламным публикациям. Выявление и детальный анализ рекламы по открытым данным социальной сети с изображениями и геолокациями, построение тематической карты предложений и анализ эффективности рекламы в разных частях города с последующим созданием методов рекомендации позволит потенциальным рекламодателям выбирать оптимальную геолокацию для размещения рекламы, а также поможет понять, каких тематических предложений не хватает в интересующем районе города.

Актуальность данной работы заключается в том, что на текущий момент область выделения рекламных публикаций по неразмеченным данным социальных сетей достаточно слабо покрыта исследованиями, а существующие решения покрывают выделение рекламы только в какой-то конкретной области. Существуют алгоритмы, уже разработанные для отслеживания рекламы, как, например, в социальной сети, не распространённой в России, описание которой приводится в работе Márcio et al. [1]. Однако в нашем случае такой подход не сработает, так как в данной работе представлена отдельная система, где заранее известно, есть ли в посте реклама или нет. Другая работа, в которой используется выделение рекламы, берёт за основу Википедию: поскольку это открытая интернет-энциклопедия, каждый может ее редактировать, и иногда существует возможность спама и рекламы, что недопустимо. Поэтому Yuanzhen Guo и др. [2] применили метод опорных векторов на векторах документов, чтобы определить, содержит ли отредактированная пользователем статья рекламу или нет. Практически так же, как и в предыдущем случае, Liang, Xiao и др. [3] в своей работе решали проблему определения нативной рекламы, но в статьях китайской социальной сети WeChat. В

данных работах использовались размеченные данные для выделения рекламы, что так же не подходит в нашем случае, так как наши данные для работы – неразмечены.

Однако, Martijn Cool в своей работе [4] решал ту же проблему, что и мы: автор рекомендовал рекламу на основе геолокаций. Когда были собраны неразмеченные данные не распространённой в России социальной сети, описанной в приведённой статье, к ним был применен процесс фильтрации, а в качестве алгоритма, использованного для тематического моделирования, был выбран LDA. Хоть данная работа и полезна для нас, но в качестве данных берутся данные из социальной сети, не распространённой в России, а в нашей работе используются данные из другой социальной сети с геолокациями и изображениями. Исходя из проведённого анализа литературы можно сделать вывод, что тема нашей работы актуальна, так как область выделения рекламы достаточно слабо освещена исследованиями, а в качестве данных в подобных работах используются социальные сети, непопулярные среди русскоговорящего населения.

Основная часть

Данные, которые используются в данной работе, были собраны из социальной сети с геолокациями и изображениями и содержат в себе информацию о координатах геолокации, непосредственно сам текст публикации, ID пользователя, время публикации, количество лайков и комментариев. Данные представлены за 2017–2020 годы по городу Санкт-Петербург и за 2018–2020 год по городу Нью-Йорк. Первой задачей в данной работе было отделение рекламных публикаций от не рекламных. Ниже приведено краткое описание использованных в работе методов и результаты их применения на наших данных. Также, для русских данных было размечено 1000 постов 500 рекламных и 500 не рекламных, они были использованы для анализа результатов моделей.

Разделение данных на 2 тематики – реклама и не реклама

Неотрицательная матричная факторизация (NMF) – это метод, который можно использовать при обучении на неразмеченных данных для извлечения нескольких тематик [5]. NMF – это статистический метод уменьшения размерности входных корпораций. Слова с меньшей связностью получают относительно меньший вес по сравнению со словами, которые более связны. Для данных Санкт-Петербурга, после применения NMF к нашим данным, были получены 2 темы, которые представлены следующими топ-словами:

1 тема – ['день очень человек друг время сегодня любить самый любимый жизнь просто каждый праздник фото первый ребёнок работа знать питер рождество'].

2 тема – ['цена заказ размер доставка петербург директ запись санкт руб маникюр наличие whatsapp профиль спб писать ссылка прямое платье шапка'].

Выделенные темы достаточно репрезентативны, и можно понять, что 1 тема относится к обычным сообщениям, а вторая - к рекламным публикациям. Чтобы оценить, насколько хорошо представлены темы и насколько они интерпретируемы для человека, был использован показатель когерентности. Он измеряет относительное расстояние между словами в темах. В данном случае был получен показатель когерентности, равный 0,485. Результаты на тестовых данных представлены ниже:

Accuracy = 86%, Precision = 86,3%, Recall = 85,8%, F-measure = 86,1%.

Что же касается данных на Нью-Йорке, были получены следующие ключевые слова по темам:

1 тема – [new, year, new year, happy, happy new, new york, york, year new, happy birthday, everybody, resolution, family, start]

2 тема – [day, love, night, time, one, today, get, good, pm, last, come, great, see, back, nyc, thank, best, thirst].

Здесь не так очевидно, какая тема что обозначает, однако, судя по словам 'great', 'pm', 'first', можно предположить, что вторая тема обозначает рекламу. Показатель когерентности в данном случае 0.619.

Следующий метод, который использовался в данной работе – Латентное распределение Дирихле (LDA) [6]. Латентное распределение Дирихле – это генеративная вероятностная модель, которая представляет каждую тему как смесь слов, а каждый текст – как смесь тем. Она основана на двухуровневой байесовской генеративной модели, которая предполагает распределение Дирихле для распределений тем и слов.

Для данных Санкт-Петербурга были получены следующие топ-слова:

Тема: 1 Word: 0.004*"запись" + 0.004*"цена" + 0.004*"заказ" + 0.004*"директ" + 0.004*"маникюр" + 0.003*"размер" + 0.003*"руб" + 0.003*"доставка" + 0.003*"петербург" + 0.003*"спб"

Тема: 2 Word: 0.004*"день" + 0.003*"рождество" + 0.003*"друг" + 0.003*"очень" + 0.003*"человек" + 0.003*"любить" + 0.002*"праздник" + 0.002*"сегодня" + 0.002*"любимый" + 0.002*"питер"

Здесь первый топик характеризует рекламу. Были посчитаны метрики аналогично предыдущему случаю:

Coherence score = 0.609, Accuracy = 92,2%, Precision = 96,1%, Recall = 98,8%, F-measure = 97,4%.

Для данных по Нью-Йорку были получены следующие результаты по топикам:

Topic 1: 0.009*"new" + 0.008*"night" + 0.007*"pm" + 0.007*"de" + 0.006*"nyc" + 0.006*"u" + 0.006*"york" + 0.005*"th" + 0.005*"tonight" + 0.005*"show"

Topic 2: 0.010*"day" + 0.008*"love" + 0.007*"happy" + 0.007*"time" + 0.007*"wa" + 0.006*"one" + 0.006*"good" + 0.005*"like" + 0.005*"great" + 0.005*"get"

Здесь так же, как и по аналогии с NMF, не было возможности достаточно точно определить, какая тема характеризует рекламу, поэтому было сделано предположение, что первая тематика характерна для рекламы. Показатель когерентности здесь равен 0.63.

Переходим к следующей теме – BigARTM. Это проект с открытым исходным кодом, основанный на аддитивной регуляризации на тематических моделях (ARTM) [7]. По сравнению с моделью LDA, здесь добавлен регуляризатор, а в качестве функции регуляризатора используется расхождение KL. Перед обучением модели были созданы словарь и партии данных. После создания модели с 2 темами она была обучена на текущих данных.

Ниже представлены 2 результирующие темы по данным Санкт-Петербурга:

тема 1:

['цена', 'кожа', 'директ', 'запись', 'заказ', 'цвет', 'волос', 'руб', 'работа', 'доставка', 'размер', 'профиль', 'любой', 'мастер', 'писать', 'маникюр', 'процедура', 'whatsapp', 'вопрос', 'стоимость', 'ссылка', 'наличие', 'прямой'].

тема 2:

['день', 'человек', 'очень', 'время', 'жизнь', 'самый', 'друг', 'ребёнок', 'сегодня', 'просто', 'каждый', 'первый', 'знать', 'любить', 'место', 'праздник', 'фото', 'любимый', 'январь', 'делать', 'работа', 'дом', 'сделать', 'петербург', 'город', 'любовь'].

Среди этих тем также очевидно, какая тема включает рекламные публикации, а какая не рекламные - первая предназначена для рекламы.

Ниже представлены результирующие метрики для данной модели:

Perplexity score = 5407,91, Sparsity Score = 0,289, Accuracy = 97,4%, Precision = 89,8%, Recall = 95,2%, F-measure = 92,4%, Coherence score = 0.834.

Переходя к данным по Нью-Йорку, были получены следующие топ-слова:

тема 1

день, один, время, любовь, новый, сегодня, получить, нравится, видеть, год, ночь, большой, спасибо, прийти, хороший, последний, сделать, работа, жизнь, вернуться,

первый, счастливый, идти, друг, показать, лучший, неделя, знать, ха, удивительный, всегда, получил, красивый, путь, много, люди

тема 2

pm, de, new, que, en, free, el, night, Sunday, com, happy, York, tonight, available, nyc, special, drink, party, open, birthday, brunch, day, Saturday, Friday, music, food, come, bottle, delicious, para, ave, chicken, dinner, order

Стоит отметить, что вторая тема больше похожа на рекламу, и в ней содержится много информации о еде. Показатель когерентности в данном случае равен 0.795.

Этот метод превосходит оба предыдущих метода и показывает довольно хорошие результаты на размеченных тестовых данных, что означает, что эта модель может быть использована для извлечения рекламных публикаций среди прочих.

И последний метод, который был рассмотрен в данной работе для выделения рекламы и не рекламы – Zero-shot learning [8]. Данный подход хорошо работает с данными, которых не было в тренировочной выборке. Что касается набора данных по Санкт-Петербургу, были получены следующие топ-слова:

1: ['прямой', 'срочно', 'Санкт-Петербург', 'whatsapp', 'viber', 'приложение', 'снег', 'смс', 'распродажа', 'январь'],

2: ['mua', 'friends', 'january', 'snow', 'saintpetersburg', 'immediately', 'ph', 'petersburg', 'vk', 'com'].

Здесь можно увидеть, что 1 тема представляет рекламные публикации, а вторая - обычные сообщения. Метрики на тестовых данных:

Accuracy = 75,9%, Precision = 94,5%, Recall = 55%, F-Measure = 69%. Coherence score = 0.812.

На данных Нью-Йорка был получен показатель когерентности, равный 0.831.

В заключение необходимо подвести итоги того, что было достигнуто. Ниже приведена общая таблица 1 с результатами по разделениям на тематики рекламы и не рекламы.

Таблица 1

Сравнение метрик для рассмотренных методов на данных Санкт-Петербурга

Метод	Accuracy	Precision	Recall	F-score	Coherence
NMF	0.86	0.863	0.858	0.861	0.485
LDA	0.922	0.961	0.988	0.974	0.609
BigARTM	0.974	0.898	0.952	0.924	0.834
Zero-shot	0.759	0.945	0.55	0.69	0.812

Можно увидеть, что LDA превосходит другие модели на тестовых данных, однако по показателю точности побеждает модель BigARTM. Говоря о представлении тем, BigARTM и Zero-Shot показали хорошие результаты, и поэтому модель BigARTM была выбрана для выделения рекламных и нерекламных публикаций на данных по Санкт-Петербургу (табл. 2).

Таблица 2

Сравнение метрик для рассмотренных методов на данных Нью-Йорка

Метод	LDA	NMF	BigARTM	Zero-Shot
Показатель когерентности	0.609	0.619	0.795	0.831

В данном случае можно увидеть, что Zero-Shot модель превосходит другие, и поэтому она была выбрана в качестве модели для разделения рекламы и не рекламы на неразмеченных данных социальных сетей по городу Нью-Йорк.

Разделение рекламных данных на более узкие тематики

Для разбиения рекламных данных на более узкие тематики использовались модели LDA, NMF, BigARTM и Zero-shot. Для выбора оптимального количества необходимых тем был использован показатель когерентности в качестве триггера для определения лучшего количества тематик для разделения.

Однако, здесь возникла проблема, связанная с наличием большого количества общих слов в текущих рекламных данных. Такие слова, как “заказ”, “директ”, “доставка”, “наличие” и т.д. были очень полезны при разделении данных на рекламу и не рекламу, в то время как при разделении рекламных постов на более узкие тематики модели выделяли общие темы, относящиеся к рекламе, но не относящиеся к чему-то конкретному. Для решения данной проблемы был составлен список из порядка 200-х топ-слов, которые были убраны из данных. После очистки данных удалось повысить качество разделения и определения конкретных тем.

На данных Санкт-Петербурга были получены следующие результаты (табл. 3).

Таблица 3

Сравнение метрик на использованных методах на данных Санкт-Петербурга

Метод	Показатель когерентности	Лучшее количество тематик
NMF	0.384	10
LDA	0.632	25
Zero-shot	0.574	15
BigARTM	0.675	13

В результате анализа было определено, что лучший результат показала модель BigARTM и данные разбиваются на 13 отдельных хорошо разделимых тематик. Регулязаторы для модели были выбраны следующие: sparse_phi_reg = -1.0, sparse_theta_reg = -0.5, decorellator_phi_reg = 2.5e+5.

Темы, которые были выделены, относятся к уходу за собой, тренировкам, еде, уходу за ногтями, подаркам и т.д.

На данных Нью-Йорка были получены следующие результаты (табл. 4).

Таблица 4

Сравнение метрик на использованных методах на данных Санкт-Петербурга

Метод	Показатель когерентности	Лучшее количество тематик
NMF	0.468	5
LDA	0.708	15
Zero-shot	0.741	15
BigARTM	0.748	10

В результате анализа было определено, что лучший результат так же показала модель BigARTM и данные разбиваются на 10 хорошо разделимых тематик. Регулязаторы для модели были выбраны следующие: sparse_phi_reg = -0.8, sparse_theta_reg = -0.55, decorellator_phi_reg = 2.8e+5.

Темы, которые были выделены, относятся к уходу за волосами, еде, путешествиям, тренировкам и т.д. Стоит отметить, что существует пересечение между тематиками, выделенными на данных двух городов: и там и там присутствуют темы по уходу за волосами, тренировкам, еде. Но в то же время, есть и различия, например, в городе Санкт-Петербург одна из самых популярных тем относится к уходу за ногтями, в то время как в Нью-Йорке популярная тематика относится к путешествиям.

Для визуализации выделенных тематик в Санкт-Петербурге, на карту были нанесены 3 самые популярные темы, распределение представлено ниже (рисунок).

Рисунок. Распространение рекламных публикаций по карте Санкт-Петербурга (красный – маникюр, синий – уход за волосами, фиолетовый – одежда)

Проанализировав результаты визуализаций, можно сделать вывод, что реклама маникюра более распространена на окраинах города, так как чаще всего этим занимаются частные лица. В то же время реклама одежды более сосредоточена в центре города, так как одежда чаще всего продается в магазинах и торговых центрах, которые расположены в городе.

В результате данной работы были проанализированы существующие методы по выделению рекламных и не рекламных публикаций, определены их плюсы и минусы, была определена актуальность данной работы. Методы тематического моделирования были применены сначала для выделения рекламных публикаций среди сырого текста, для данных Санкт-Петербурга была выбрана модель BigARTM, для данных Нью-Йорка – Zero-Shot. Далее, после предварительной отчистки текста, были определены модели и оптимальное количество тематик на двух разных наборах данных. В СПБ было выделено 13 тем при помощи модели BigARTM, на данных Нью-Йорка было выделено 10 тем также при помощи модели BigARTM. Все полученные результаты были проанализированы и нанесены на карту города.

Методы, представленные в работе, позволили повысить эффективность выделения рекламных публикаций и рекламных категорий, основываясь на анализе неструктурированных данных социальных сетей города Санкт-Петербург. В дальнейших исследованиях разработанные методы будут использованы для определения эффективности рекламных публикаций (по метахарактеристикам постов), и разработки методов рекомендаций геолокаций для оптимального размещения рекламы в социальных сетях. Сам же пайплайн планируется применять в социальных сетях для более рационального размещения рекламы в городе, а также для анализа распределения тематик на карте с последующим применением результатов данного анализа в проектах, связанных с городской средой.

Литература

1. Silva M. et al. Facebook ads monitor: an independent auditing system for political ads on Facebook //Proceedings of The Web Conference 2020. 2020. С. 224-234.
2. Guo Y., Iwaihara M. Detection of text-based advertising and promotion in wikipedia by deep learning method //Proc. DEIM. 2015. С. 1-7.
3. Liang X., Wang C., Zhao G. Enhancing content marketing article detection with graph analysis //IEEE Access. 2019. Т. 7. С. 94869-94881.
4. Cool M. Topic modelling: Location-based offline advertising using Twitter: дис. Dublin, National College of Ireland. 2018.
5. Topic Modelling using NMF [Электронный ресурс]. Режим доступа: https://medium.com/voice-tech-podcast/topic-modelling-using-nmf-2f510d962b6e (дата обращения: 12.06.2022).
6. Evaluate Topic Models: Latent Dirichlet Allocation (LDA) [Электронный ресурс]. Режим доступа: https://towardsdatascience.com/evaluate-topic-model-in-python-latent-dirichlet-allocation-lda-7d57484bb5d0 (дата обращения: 15.06.2022)
7. Beyond LDA: State-of-the-art Topic Models With BigARTM [Электронный ресурс] – Режим доступа: https://www.databricks.com/blog/2022/02/16/beyond-lda-state-of-the-art-topic-models-with-bigartm.html (дата обращения: 20.06.2022)
8. Bianchi F. et al. Cross-lingual contextualized topic models with zero-shot learning //arXiv preprint arXiv:2004.07737. 2020.

УДК 004.822

ОБЗОР МЕТОДОВ ИЗВЛЕЧЕНИЯ ГРАФОВ ЗНАНИЙ ИЗ ИЗОБРАЖЕНИЙ

Дубовской А.А.[1] (аспирант)
Научный руководитель – к.т.н., доцент Баймуратов И.Р.[1]
1 – Университет ИТМО
e-mail: maverick_142@mail.ru

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

Аннотация
Создание, управление и использование коллекций изображений может быть автоматизировано с помощью извлечения графов знаний из изображений. Однако на данный момент не существует систематического обзора данной исследовательской области. Таким образом, целью данной работы является систематический обзор и анализ существующих решений по извлечению графов знаний из изображений. В результате работы существующие решения сгруппированы по целям и подходам, выявлены наиболее часто используемые методы, лучшие результаты и пробелы в исследованиях.

Ключевые слова
Коллекции изображений, компьютерное зрение, граф знаний, извлечение онтологий, классификация.

Создание, управление и использование коллекций изображений требует большого количества трудозатрат. Эти задачи могут быть частично автоматизированы с помощью извлечения графов знаний из изображений. На данный момент уже предложен ряд подобных решений, однако не существует систематического обзора данной исследовательской области. Таким образом, цель данной работы состоит в поиске и анализе существующих решений по извлечению графов знаний из изображений.

Поиск источников осуществлялся в системе Google Scholar по ключевым словам: “knowledge graph extraction from images”, “ontology extraction from images”, “ontology extraction from videos”, “knowledge graph extraction from videos”, "building knowledge graph based on videos", "building knowledge graph based on images", "building ontology based on images", "building ontology based on videos".

Результаты поиска отбирались по следующим критериям:

1) статья была опубликована не ранее 2016 года;
2) статья содержит информацию по рассматриваемой теме;
3) статья включает описание архитектуры решения, приведенные в статье эксперименты воспроизводимы.

При исследовании литературы, посвященной проблеме извлечения графов знаний из изображений, были найдены существующие обзорные статьи, связанные с данной тематикой. В статье [1] представлен систематический обзор методологий на основе семантических графов для представления и понимания семантики аэросъемки. Данный обзор не включает в себя описание архитектур, применяемых в методах извлечения графов знаний из изображений. В статье [2] обозреваются методы извлечения информации и дальнейшее построение графов знаний на основе этой информации. Данный обзор не включает в себя информацию о методах извлечения графов знаний из изображений. В [3] авторы предоставляют всесторонний обзор графа знаний. В данном обзоре нет информации о генерации графов знаний. В обзоре [4] основное внимание уделяется подходам к обучению с помощью визуального переноса с использованием графов знаний. Данный обзор направлен не на извлечение графов из изображений, а на улучшение результатов классификации.

Также были найдены статьи, не подходящие под один из критериев, а именно в них не включены описания архитектур решений, что делает невозможным воспроизведение результатов. В статье [5] предлагается метод извлечения онтологии из видео. Он основан на технологии глубокого обучения для идентификации объектов и движений в видеосценах. Метод включает в себя использование сети глубокого обучения и системы построения онтологий Protégé. В статье [6] авторы рассматривают задачу автоматического построения визуальной онтологии на основе фолксономии (FBVO) из созданных пользователем аннотированных изображений. Предлагается систематическая структура, состоящая из трех этапов: обнаружение тематики, выделение взаимосвязи объектов тематики и построение иерархии объектов.

В результате систематического анализа статей, подходящих под критерии, были сформированы две группы по целям исследований:

1) извлечение знаний;
2) уточнение классификации.

А также были сформулированы два подхода к извлечению графов знаний:

1) с использованием существующего графа знаний;
2) без использования существующего графа знаний.

Соответственно, источники были сгруппированы по четырем категориям:

1) извлечение графов знаний из изображений и генерация графа знаний с нуля;
2) уточнение результатов классификации объектов на изображении и генерация графа знаний с нуля;
3) извлечение графов знаний из изображений и дополнение существующего графа знаний;
4) уточнение результатов классификации объектов на изображении и дополнение существующего графа знаний.

В таблице 1 представлены статьи, попадающие под первую категорию.

Таблица 1

Извлечение графов знаний из изображений и генерация графа знаний с нуля

Статья	Нейронная сеть для распознавания объектов	Фреймворк генерации графа знаний	Набор данных	Результат
[7]	YOLOv3, MTCNN, ResNet-152	MMKG	GoodNews, NYTimes800k	CIDEr 54,02 и 54,47
[8]	VGG-16	HRE	Visual Genome	recall@k 0,57
[9]	Mask R-CNN	Neo4j	ADE20K	Mean IoU 79,98%
[10]	VGG16	MMKG	FB-DB, FB-Yago	PoE-lrni 82% и 74,6%
[11]	TinyYOLO	VEKG	Pexels, Le2i	F-Score от 0,75 до 0,86
[12]	VGG	WordNet	DBpedia	Точность 80%
[13]	Faster R-CNN, MAP, MTCNN	DARPA AIDA	TAC SMKBP 2019	F1 29,7%

Под вторую категорию попала лишь одна статья [14], использующая нейронную сеть Faster R-CNN и GSNN в качестве фреймворка генерации графа знаний. Оценка системы, представленной в работе, производилась на наборе данных MS COCO и ее точность составила 73,97%.

Под третью категорию попала также одна статья [15], дополняющая граф знаний WordNet при помощи нейронных сетей VGG19 и I3D. Точность работы системы составила 22,16% и 59,19% на наборах данных MSVD* и MSR-VTT* соответственно.

Статьи, попадающие под четвертую категорию представлены в таблице 2.

Таблица 2

Уточнение результатов классификации объектов на изображении и дополнение существующего графа знаний

Статья	Нейронная сеть для распознавания объектов	Граф знаний	Набор данных	Результат
[16]	Inception-ResNetV2	IKG	ImageNet	Точность 82,43%
[17]	YOLO9000	Concept-Net	MS COCO	CIDEr-D 107,2
[18]	CB-CNN	KGRF	Caltech-UCSD Birds	Точность 88,49%
[19]	Модель HMAX	WordNet	ImageNet	Точность 63,8%
[20]	Inception V3	WordNet	ILSVRC2010, SUN397, Caltech256	Точность 68,75%, 53,13% и 79,92%

Общим принципом, применяемым во всех найденных статьях, является использование сверточной нейронной сети. Для извлечения сущностей и их связей на изображениях авторы применяют такие архитектуры, как VGG16, VGG19, YOLO, Faster R-CNN, Mask R-CNN и InterseptionResNetV2. Результаты оцениваются различными метриками, наиболее часто используется точность, наилучший показатель был достигнут в решении [18] – 88,49%.

Результатом работы является систематический обзор литературы на тему извлечения графов знаний из изображений, а также их применения для улучшения результатов классификации объектов. Данный обзор позволяет выявить наиболее оптимальный подход к извлечению графов знаний из изображений и улучшению результатов классификации объектов. Во всех обозреваемых статьях используются только методы обучения с учителем, что существенно повышает количество трудозатрат и стоимость обучения систем. Дальнейшее развитие работы предполагает исследование систем обучения без учителя для генерации графов знаний на основе изображений.

Литература

1. Sun S., Dustdar S., Ranjan R., Morgan G., Dong Y., Wang L. Remote Sensing Image Interpretation With Semantic Graph-Based Methods: A Survey // IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing. 2022. №15. P. 4544–4558.
2. Zamini M., Reza H., Rabiei M. A Review of Knowledge Graph Completion // Information. 2022. №13(8). P. 396.
3. Ji S., Pan S., Cambria E., Marttinen P., Yu P.S.A Survey on Knowledge Graphs: Representation, Acquisition, and Applications // IEEE Transactions on Neural Networks and Learning Systems. 2022. №33(2). P. 494–514.
4. Monka S., Halilaj L., Rettinger A. A survey on visual transfer learning using knowledge graphs // Semantic Web. 2022. №13(3). P. 477–510.
5. Bornia J., Ali F. Combining Deep Learning and Ontology to Reveal Video Sequences Semantics // Revue d'Intelligence Artificielle. 2021. №35(2). P. 131–138.
6. Fang Q., Xu C., Sang J., Hossain M. S., Ghoneim A. Folksonomy-Based Visual Ontology Construction and Its Applications // IEEE Transactions on Multimedia. 2016. №18(4). P. 702–713.
7. Zhao W., Hu Y., Wang H., Wu X., Luo J. Boosting Entity-aware Image Captioning with Multi-modal Knowledge Graph. 2021. arXiv:2107.11970.
8. Mao J., Yao Y., Heinrich S., Hinz T., Weber C., Wermter S., Liu Z., Sun M. Bootstrapping knowledge graphs from images and text // Frontiers in neurorobotics. 2019. №13. P. 93.

9. Fang W., Ma L., Love P.E.D., Luo H., Ding L., Zhou A. Knowledge graph for identifying hazards on construction sites: Integrating computer vision with ontology // Automation in Construction. 2020. №119.
10. Liu Y., Li H., Garcia-Duran A., Niepert M., Onoro-Rubio D., Rosenblum D.S. MMKG: Multi-modal Knowledge Graphs // The Semantic Web. 2019. P. 459–474.
11. Yadav P., Curry E. VEKG: Video Event Knowledge Graph to Represent Video Streams for Complex Event Pattern Matching // First International Conference on Graph Computing (GC). 2019. P. 13–20.
12. Zhang X., Sun X., Xie C., Lun B. From Vision to Content: Construction of Domain-Specific Multi-Modal Knowledge Graph // IEEE Access. 2019. №7. P. 108278–108294.
13. Li M., Zareian A., Lin Y., Pan X., Whitehead S., Chen B., Wu B., Ji H., Chang S., Voss C., Napierski D., Freedman M. GAIA: A Fine-grained Multimedia Knowledge Extraction System // Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics: System Demonstrations. 2020. P. 77–86.
14. Marino K., Salakhutdinov R., Gupta A. The More You Know: Using Knowledge Graphs for Image Classification. 2016.
15. Mahon L., Giunchiglia E., Li B., Lukasiewicz T. Knowledge Graph Extraction from Videos // 19th IEEE International Conference on Machine Learning and Applications (ICMLA). 2020. P. 25–32.
16. Zhang D., Cui M., Yang Y., Yang P., Xie C., Liu D., Yu B., Chen Z. Knowledge Graph-Based Image Classification Refinement // IEEE Access. 2019. №7. P. 57678–57690.
17. Zhou Y., Sun Y., Honavar V. Improving Image Captioning by Leveraging Knowledge Graphs // IEEE Winter Conference on Applications of Computer Vision (WACV). 2019. P. 283–293.
18. He Y., Tian L., Zhang L., Zeng X. Knowledge Graph Representation Fusion Framework for Fine-Grained Object Recognition in Smart Cities // Complexity. 2021. P. 1–9.
19. Filali J., Zghal H., Martinet J. Ontology and HMAX Features-based Image Classification using Merged Classifiers // Proceedings of the 14th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications. 2019. P. 124–134.
20. Zhang Y., Qu Y., Li C., Lei Y., Fan J. Ontology-driven hierarchical sparse coding for large-scale image classification // 2019. №360. P. 209–219.

УДК 004.891

РАЗРАБОТКА ЧАТ-БОТА ДЛЯ ФОРМИРОВАНИЯ РЕКОМЕНДАЦИЙ МЕСТ ОБЩЕСТВЕННОГО ПИТАНИЯ С УЧЕТОМ ПРЕДПОЧТЕНИЙ ПОЛЬЗОВАТЕЛЯ

Егорова Е.В.[1] (студент), **Чеканова Е.Д.**[1] (студент)
Научный руководитель – к.т.н., доцент Воробьева М.С.[1]
1 – ФГАОУ ВО «Тюменский государственный университет»
e-mail: egorovaelizaveta.study@gmail.com

Аннотация
В настоящей работе представлено описание разработки чат-бота для формирования рекомендаций мест общественного питания с учетом пользовательских предпочтений. Рассматривается создание рекомендательного программного продукта на основе методов: косинусного расстояния, сингулярного разложения. Для взаимодействия с пользователем спроектирован и разработан чат-бот в мессенджере Telegram.

Ключевые слова
Анализ данных, формирование рекомендаций, рекомендательные системы, контентный подход, косинусное расстояние, матричная факторизация, сингулярное разложение, SVD, метрики ранжирования.

В современном обществе все чаще возникают проблемы анализа большого объема данных и предоставления рекомендаций на его основе [1]. Используя алгоритмы для формирования рекомендаций, вместо предоставления статической информации можно сформировать релевантный контент, основанный на знаниях о пользователе и объектах [2]. В настоящее время подбор персонального контента используются во многих автоматизированных системах для решения разных задач. В статье [3] описывается работа Л. Л. Волковой и М. М. Токарева по созданию рекомендательной системы, которая учитывает различные параметры прочитанных книг, включая автора, название, страну происхождения, стиль, период написания и жанр. В работе Тео Иеремия [4] предлагается использование коллаборативной фильтрации для создания системы для подбора ресторана с учетом особенностей входных данных. Решение аналогичной задачи рассматривается в публикации [5], в которой автор описывает разработку рекомендательной модели на основе популярности заведений, данный подход рекомендует пользователю самые модные кафе и рестораны, находящиеся в непосредственной близости.

Каждый человек ежедневно сталкивается с проблемой выбора, например: выбор места общественного питания. Большое количество заведений, недостаточный или некорректный набор информации, ограничение по времени, которое можно затратить на подбор места для посещения, компетентность лица, производящего сравнение и выбор – все это причины, которые усложняют принятие решения. С этой проблемой часто сталкиваются студенты, участвующие в мероприятиях в других городах, участники студенческого совета, решившие провести встречу в кафе, сотрудники, которые переехали на новое место, и др. Чтобы упростить процесс выбора мест общественного питания, необходимо сформировать подборку наиболее релевантных заведений для таких пользователей.

Предметом исследования является упрощение процесса выбора места общественного питания. Научная новизна заключается в проведении сравнительного анализа различных методов формирования рекомендаций, а также в разработке чат-бота с использованием рекомендательной модели, предоставляющей пользователю список заведений в соответствии с его предпочтениями.

Источником данных является сервис отзывов для путешественников TripAdvisor, по данным на февраль 2023 года в Тюмени находится 788 мест общественного питания. В качестве объектов для формирования рекомендаций используются данные о ресторанах, кафе и барах Тюмени. После проведения опроса были выделены критерии, на которые опираются пользователи для выбора подходящего заведения. Наиболее значимыми характеристиками являются рейтинги, местоположение, средний чек, уникальные предложения программы лояльности. Наличие сложностей при определении местоположения заведений в сервисе TripAdvisor обусловлено отсутствием единого формата обработки и хранения адресной информации. Для определения координат заведения была проведена предварительная унификация адресов с использованием регулярных выражений, а затем процедура геокодирования.

На основе обзора статьей был проведен анализ существующих методов фильтрации и алгоритмов, используемых для построения рекомендательных систем [6]. Были выбраны такие подходы как: контентная фильтрация и матричная факторизация. Оба способа позволяют учитывать индивидуальные предпочтения пользователей и повышают точность рекомендаций, благодаря чему алгоритм формирования рекомендаций становится более эффективным и удобным для пользователей.

Контентный метод можно реализовать с помощью косинусного расстояния: определения сходства между двумя векторами (вектором оценок пользователя и вектором оценок заведений) путем вычисления косинуса угла между ними (1). Вектор оценок пользователя представлен вектором частот характеристик посещенных заведений. Чем чаще определенная характеристика была выбрана, тем выше ее значимость в векторе оценок пользователя.

$$\text{cosine}(A,\ B) = \frac{\sum_{i=1}^{n} A_i B_i}{\sqrt{\sum_{i=1}^{n} A_i^2}\sqrt{\sum_{i=1}^{n} B_i^2}}, \quad (1)$$

где A – вектор оценок заведений, B – вектор оценок пользователя, n – количество объектов, i – индекс объекта.

Подход матричной факторизации реализуется с помощью метода сингулярного разложения (Singular Value Decomposition), целью метода в контексте формирования рекомендаций является получение прогноза оценок пользователя для заведений. Благодаря SVD можно представить User-Item (*UI*) матрицу в виде выражения (2):

$$UI_{N\times M} = U\sum V^{*}, \quad (2)$$

где U – матрица оценок пользователей (строки матрицы – пользователи), V^* – сопряженно-транспонированная матрица к V (столбцы матрицы – заведения), Σ – матрица размеров $N\times M$, на главной диагонали которой находятся сингулярные числа [7].

В частично заполненной предыдущими оценками пользователей матрице User-Item требуется восполнить пропущенные значения. Для этого необходимо приближенно вычислить сингулярное разложение с помощью минимизации функции (3) градиентным спуском.

$$L(\mu,\ b_i,\ b_j,\ u_i, v_j) = \sum_{(i,j)\in D}\left(r_{i,j} - \mu - b_i - b_j - v_j^T u_i\right)^2 \rightarrow \min, \quad (3)$$

где b_i - среднее значение оценки отдельных пользователей, b_j – среднее значение оценки отдельных ресторанов, μ – общий средний рейтинг по матрице, $v_j^T u_i$ – скалярное произведение соответствующих строк и столбца по ресторану и пользователю.

После получения итоговой матрицы *UI* для пользователей формируется список наиболее подходящих мест общественного питания по спрогнозированным оценкам.

Одним из ключевых элементов в разработке рекомендательного сервиса является точность рекомендаций, которые предоставляет модель. Для оценки качества эффективности использовались метрики ранжирования: Mean Average Precision at K (map@K) и Average Precision at K (ap@K).

На основе результатов составляется список лучших заведений, которые будут рекомендованы пользователям. Чем более подходящее место питания будет отмечено для пользователя, тем большую важность это заведение имеет в ранжировании. Оценки работы алгоритмов на основе косинусного расстояния, сингулярного разложения и пакетной реализации SVD представлены в таблице. Исходя из результатов, можно сделать вывод о том, что использование сингулярного разложения, по сравнению с косинусным расстоянием, обеспечивает более релевантные рекомендации для пользователей. Точность алгоритма, основанного на косинусном расстоянии, измеряется по метрике map@K и составляет 40%. Однако данный алгоритм не учитывает различия в абсолютных значениях, что является важным аспектом при формировании рекомендаций для мест общественного питания. Абсолютные значения, такие как оценки заведений, выражаются в разных единицах, например, в диапазоне от 1 до 5 звезд. Сингулярное разложение же учитывает эти абсолютные значения и преобразует данные в матрицу, где каждый элемент соответствует оценке пользователя для конкретного заведения. Такой подход позволяет с точностью 80% определить место общественного питания, которое может заинтересовать пользователя. Для интеграции в чат-бот был выбран метод сингулярного разложения.

Таблица

Оценки точности работы алгоритмов

Метод / Метрика	Косинусное расстояние	Сингулярное разложение	Пакетная реализация SVD
map@K	40%	79,3%	80%

Архитектура сервиса включает в себя 4 модуля: Получение данных, Хранение (База данных MongoDB), Обработка данных, Обновление информации. Интерфейс чат-бота, для создания которого применяется фреймворк aiogram, определяется мессенджером Telegram (рис. 1).

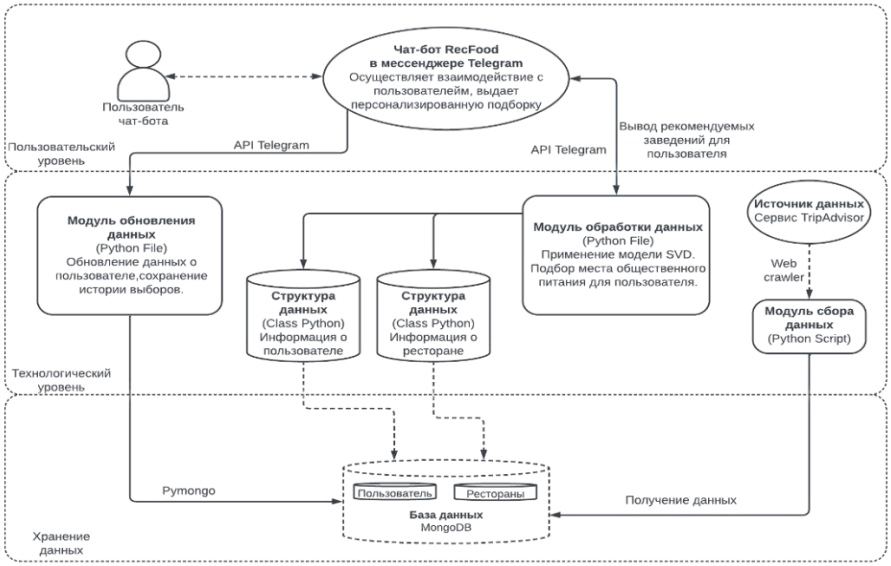


Рис. 1. Архитектура сервиса

Рассмотрим процесс взаимодействия чат-бота с пользователем, который состоит из трех этапов: формирования личных предпочтений пользователя, получения списка рекомендаций в виде сообщения с данными о 2-х наиболее релевантных заведениях и оценивания полученной подборки с заведениями (рис. 2).

Функции чат-бота:

1. Выбор заведения по личным предпочтениям пользователя.
2. Получение рекомендации места общественного питания.
3. Поиск ближайших заведений.

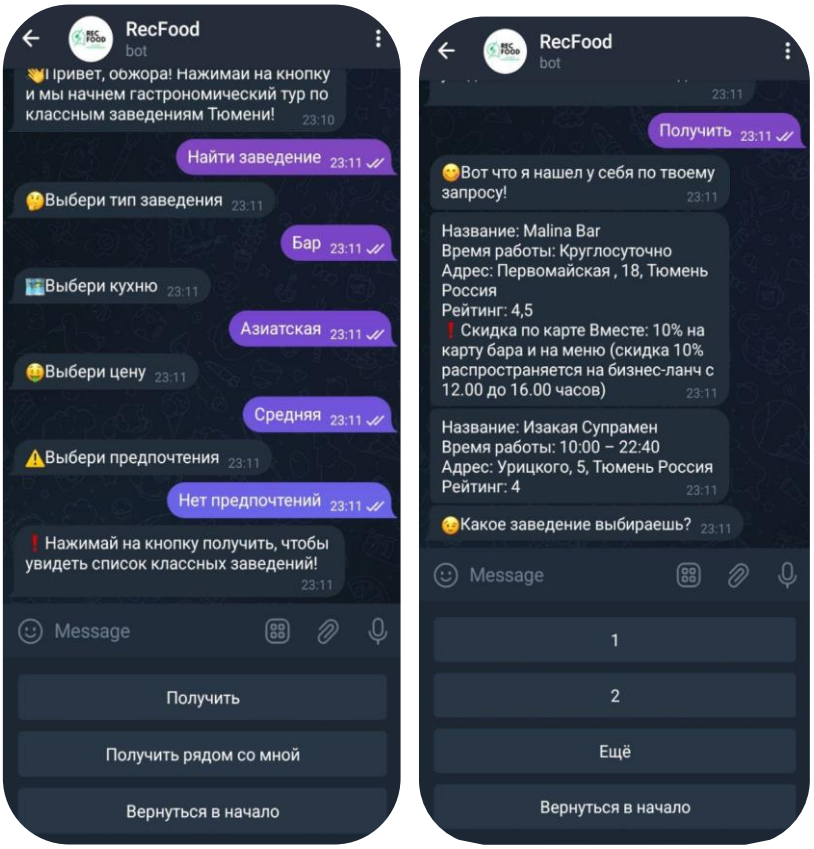


Рис. 2. Пример работы чат-бота

В итоге был спроектирован и разработан чат-бот, который, используя информацию из сервиса TripAdvisor и WebCrawler для получения данных о местах общественного питания, предоставляет рекомендации заведений пользователю, учитывая его предпочтения и скрытые факторы мест общественного питания с помощью сингулярного разложения (SVD). Такой подход приводит к точным и подходящим рекомендациям, а также повышает эффективность и полезность чат-бота для пользователей. На данный момент программный продукт находится в стадии тестирования. В период с января по март в чат-бот поступили запросы от 100 студентов ТюмГУ и местных жителей, получено более 300 подходящих мест общественного питания города Тюмень. Пользовательские отзывы, которые получены в результате сбора обратной связи, помогли выявить неточности, улучшить логику работы чат-бота перед запуском в постоянную эксплуатацию.

В дальнейшем планируется рассмотреть другие подходы для формирования рекомендаций, которые повысят точность подбора релевантных заведений, а также добавить новые функции в чат-бот, например функция расчета времени, которое пользователю необходимо будет затратить, чтобы добраться до места общественного питания с учетом загруженности автомобильных дорог.

Литература

1. Диденко А.А. Большие данные и рекомендательные системы / Диденко А.А. // Россия молодая: Сборник Материалов XII Всероссийской, Научно-Практической Конференции Молодых Ученых С Международным Участием, Кемерово, 21–24 апреля 2020 года. Кемерово: Кузбасский государственный технический университет имени Т. Ф. Горбачева, 2020. С. 211561–211567.
2. Сердинская Ю.А. Рекомендательные системы в сфере общественного питания / Сердинская Ю.А. // Приоритетные направления инновационной деятельности в промышленности: Сборник научных статей по итогам четвертой международной научной конференции., Казань, 29–30 апреля 2020 года. Том Часть 2. Казань: Общество с ограниченной ответственностью «КОНВЕРТ». 2020. С. 119–122.
3. Волкова Л.Л., Токарева М.М., Ланко А.А. О разработке рекомендательной системы, предлагающей книги по предпочтениям пользователя / Волкова Л.Л., Токарев М.М., Ланко А.А. // Новые информационные технологии в автоматизированных системах, 2017. No20.
4. Theo Jeremiah, How to Build a Restaurant Recommendation System Using Latent Factor Collaborative Filtering [Электронный ресурс] // towardsdatascience.com: платформа для обмена идеями и расширения понимания науки о данных. 2019. 29 ноября. URL: https:// towardsdatascience. com/ how-to-build-a- restaurant- recommendation-system-using-latent-factor-collaborative-filtering-ffe08dd57dca (дата обращения: 02.03.2023).
5. Sathyaseelan Krishnaraj, Restaurant Recommendation System Using Machine Learning Algorithms [Электронный ресурс] // International Journal of Science, Engineering and Technology: электронный журнал. 2021. URL: https://www.ijset.in/wp-content/uploads/IJSET_V9_issue3_233.pdf (дата обращения: 02.03.2023).
6. Исламова С.А. Обзор рекомендательных систем / Исламова С.А., Липатникова Н.В. // Modern Science. 2019. № 4–3. С. 174–176.
7. Афанасьев А.А. Вычисление сингулярного разложения матриц / Афанасьев А.А. // Сборник статей Всероссийской молодежной научной конференции «Все грани математики и механики», Томск, 25–28 апреля 2017 года. Томск: Томский государственный технический университет. 2017. С. 162–167.

УДК 004.891.2

КЛАСТЕРИЗАЦИЯ ВАКАНСИЙ ПО КЛЮЧЕВЫМ НАВЫКАМ КАНДИДАТА В СФЕРЕ ИНФОРМАЦИННЫХ ТЕХНОЛОГИЙ

Зайчикова А.И.[1] (магистрант)
Научный руководитель – к.ф.-м.н., доцент Графеева Н.Г.[1]
1 – Университет ИТМО
e-mail: 224525@niuitmo.ru

Аннотация
В работе рассматривается подход, который позволяет извлечь и структурировать информацию о профессиональных навыках кандидатов из вакансий, размещенных на открытых платформах интернет-рекрутмента, в частности, российской платформе Head-Hunter.
Как результат выполнения работы, найдены комбинации профессиональных навыков, ранжированные по степени важности каждого навыка, для различных специальностей в сфере информационных технологий. Полученные данные могут быть эффективным дополнением профессиональных портретов в реестре профессиональных навыков Министерства Труда России.

Ключевые слова
Вакансии в сфере информационных технологий, онлайн-вакансии, спрос на навыки, рынок труда, анализ вакансий, кластеризация вакансий.

Полный реестр профессиональных стандартов востребован во многих организациях, таких как государственные учреждения, учреждения социального обслуживания, высшие учебные заведения, крупные промышленные предприятия, для эффективного и своевременного получения информации о профессиях и профессиональных навыках. Традиционным решением в различных странах мира является стандартизация и унификация профессиональной структуры общества с помощью стандартов классификации профессий [1]. Стандарты и должностные инструкции, такие как содержание профессиональной деятельности, требования к навыкам и квалификационные требования, вместе образуют систему профессионального портрета [2]. Репрезентативными стандартами являются ONET в Соединенных Штатах [3], ESCO в Европейском союзе [4], NOC в Канаде [5] и COCGD в Китае [6].

В России также есть свой реестр профессиональных стандартов [7], который охватывает 1521 профессию. Эти стандарты в основном способны всесторонне и объективно описать профессии в России. Однако, по сравнению с Европой и Соединенными Штатами, российская система профессионального портрета имеет некоторые недостатки. Одним из них является отсутствие профессиональных навыков или квалификационных требований по некоторым профессиям, в частности очень скудно описаны профессии в сфере информационных технологий.

Другим существенным недостатком является то, что актуализация данных серьезно запаздывает, больше половины профессиональных портретов не обновлялись уже более 10 лет, что для многих профессий является не допустимым, поскольку появление новых технологий и различные научные открытия, как следствие, ведут за собой изменения требований, предъявляемых к кандидатам. Разумеется, главной причиной отсутствия актуальных данных является тот факт, что процесс пересмотра стандартов – это ручная работа, которая отнимает значительное количество времени, рабочую силу и много ресурсов. Поэтому возникает необходимость предоставления информации другим путем и в режиме реального времени.

Сегодня онлайн-рекрутмент стал важным каналом как для рекрутеров, распространяющих информацию об имеющихся вакансиях, так и для лиц, находящихся в поиске работы. Содержание вакансий обычно содержит описание информации о

работе, требования к специальности, образованию, опыту, профессиональным навыкам и т.д. Извлекая информацию о требованиях из вакансий и объединяя ее в профессиональную информацию, ее можно использовать в качестве эффективного дополнения к традиционному профессиональному профилированию.

В работе рассматривается подход, который позволяет извлечь профессиональные навыки из вакансий, размещенных на открытой платформе онлайн-рекрутмента Head-Hunter и провести дальнейшую кластеризацию вакансий по извлеченным навыкам с целью получения достоверных и конкретизированных профессиональных портретов в сфере информационных технологий.

Работа состоит из 5 этапов:

1. Сбор вакансий с ресурса Head-Hunter.
2. Извлечение профессиональных навыков из вакансий, присвоение весов извлеченным навыкам.
3. Формирование векторов вакансий, понижение размерности векторов вакансий.
4. Кластеризация вакансий.
5. Составление профессиональных портретов для каждого кластера.

На первом этапе с помощью официального API с ресурса Head-Hunter выгружено более 25 тысяч вакансий. Ключевым моментом является тот факт, что для работы использовались именно вакансии, а не резюме, поскольку для исследования интересны именно требуемые работодателями навыки, а не те навыки, которые могут предложить кандидаты.

Вакансии выгружались по ключевым словам: «Аналитик», «Программист», «Тестировщик», «Разработчик», «QA» и так далее – для регионов «Москва», «Санкт-Петербург» и «Россия». Вакансии выгружались пакетами по 2 тысячи, так как в API есть ограничение на выгрузку. И после удаления дубликатов осталось 16.439 тысяч вакансий.

Вакансия представляет собой json-объект, который содержит различные поля, такие как «name» - название вакансии, «salary» - з/п, «education» - образование и так далее. В первую очередь интересно поле «key_skills», в котором содержатся ключевые навыки, которыми должен обладать кандидат, откликающийся на вакансию. Для дальнейшей работы использовалось 12.809 вакансий, в которых это поле заполнено.

На втором этапе осуществляется извлечение навыков и присвоение им весов. Сначала для собранного набора данных осуществляется токенизация навыков, всего неуникальных оказалось 87.765. Затем производится механическое и ручное удаление дубликатов, остается 7.080 уникальных навыков. После этого удаляются навыки, которые встречаются во всем наборе данных менее 5 раз, для того чтобы исключить опечатки, а также «специфичные» требования свойственные конкретным компаниям, в результате остается 1.335 навыков.

Далее нужно присвоить каждому навыку вес, для этого используется два метода, а в конце рассчитывается их среднее арифметическое взвешенное значение с коэффициентами 3 и 1 для первого и второго метода соответственно.

Первый метод определяет вес навыка по позиции навыка в вакансии. Он основан на том, что работодатели на первом месте указывают самые важные для них навыки кандидатов, а дальше значимость навыков уменьшается. Таким образом, если в вакансии содержится:

– 1 навык – вес первого навыка равен 1;
– 2 навыка – вес первого навыка 0,7, вес второго навыка 0,3;
– 3 навыка – вес первого навыка 0,6, вес второго навыка 0,3, вес третьего навыка 0,1;
– 4 навыка – вес первого навыка 0,5, вес второго навыка 0,25, вес третьего навыка 0,15, вес четвертого навыка 0,1;
– 5 и более навыков – вес первого навыка 0,5, вес второго навыка 0,2, вес третьего навыка 0,1, вес остальных навыков равен и в сумме не превышает 0,2.

Второй метод определяет вес навыка по его встречаемости в наборе данных. Самый редкий навык, который встречается 5 раз, – имеет вес 0, а самый популярный навык, который встречается 2331 раз – имеет вес 1. Самым популярным навыком является SQL.

После присвоения навыкам весов осуществляется векторизация вакансий, каждой вакансии сопоставляется вектор размерностью 1.335, где каждому значению соответствует вес навыка в этой вакансии. Из полученного набора векторов удаляются вектора, сумма значений которых меньше 0,25, для того, что отсеять вакансии, в которых практически не встречаются отобранные ранее навыки. Такие вакансии вносят шум в кластеризацию и с большой долей вероятности попали в выборку случайно, так как они не относятся к информационным технологиям. В результате остается 12.185 векторов вакансий.

На третьем этапе для того, чтобы визуально определить возможна ли вообще прокластеризовать полученный набор данных, использовались стандартные методы понижения размерности: PCA (Principal component analysis) [8], t-SNE (t-distributed stochastic neighbor embedding) [9], MDS (Multidimensional Scaling) [10] и SVD (Singular Value Decomposition) [11].

Результаты работы всех методов представлены на рисунке 1. Как видно, наиболее хорошие результаты показал t-SNE, его было решено использовать для кластеризации методом GMM (Gaussian mixture model) [12].

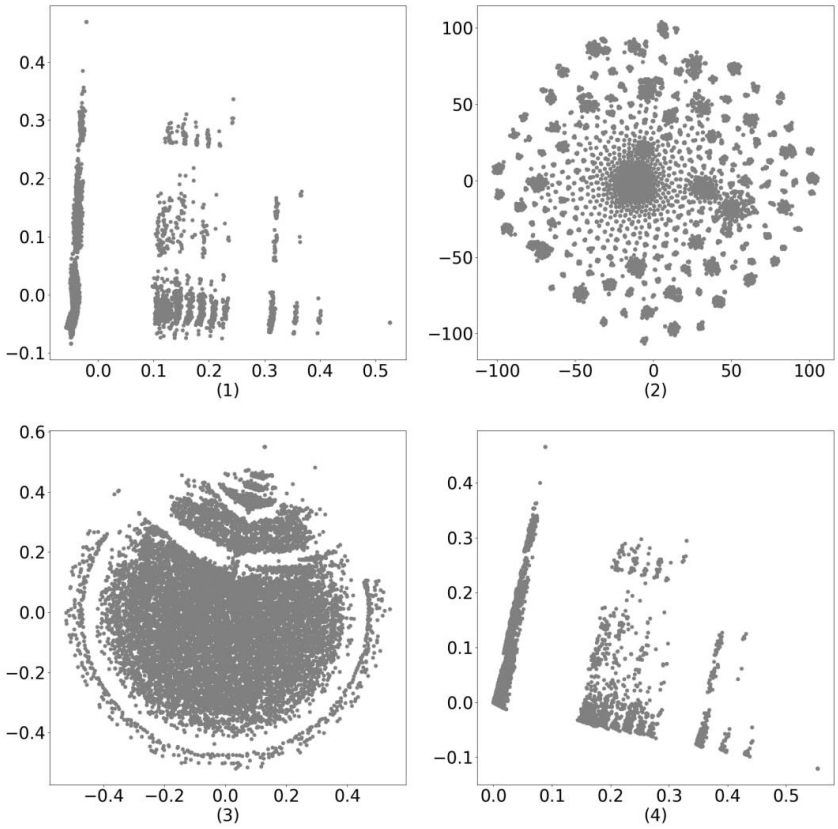


Рис. 1. Понижение размерности методами: (1) PCA, (2) t-SNE, (3) MDS, (4) SVD

На четвертом этапе проводилась кластеризация четырьмя методами: иерархическая кластеризация [13], k-means [14], DBSCAN (Density-based spatial clustering of applications with noise) [15] и GMM [12].

DBSCAN и иерархическая кластеризация показали плохие результаты, поскольку структура данных мало похоже на иерархическую и кластеры достаточно близко расположены к друг другу. В дальнейшем DBSCAN использовался только для удаления шумов.

k-means показал неплохие результаты, за счет того, что в выборке облака точек во много расположены по кругу и не имеют сложно структуры. Но он определил слишком много вакансий в 0 кластер.

Наилучший результат показал GMM после понижения размерности данных методом t-SNE. В результате кластеризации получилось 59 кластеров, которым можно дать какое-либо осмысленное описание. Тут важно отметить, что не все из этих кластеров являются репрезентативными. Результат кластеризации представлен на рисунке 2.

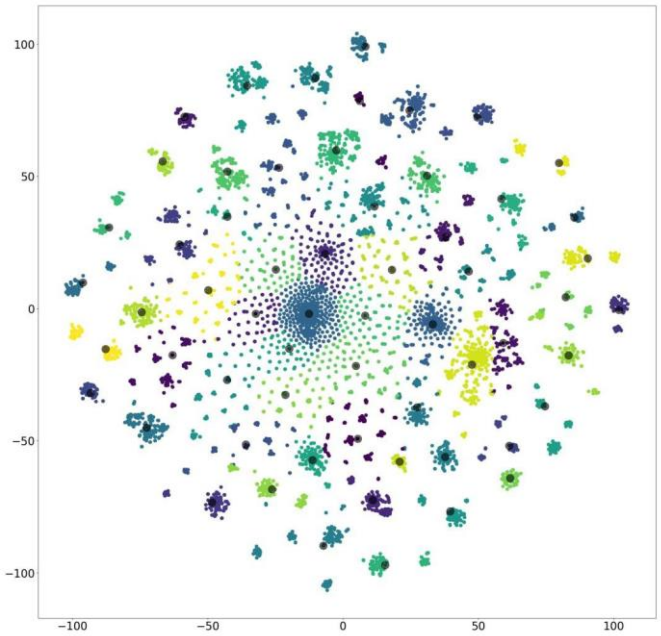


Рис. 2. Кластеризация методом GMM

На пятом этапе для каждого кластера определены 5 центральных вакансий и составлен топ профессиональных навыков с учетом веса каждого навыка в этом конкретном кластере. В качестве примера приведем два полученных кластера, которым условно определены названия «Системный аналитик низкого уровня квалификации» и «Тестировщик ПО низкого уровня квалификации», топы профессиональных навыков для каждого кластера представлены на рисунке 3.

Рис. 3. Профессиональные портреты: (1) Системный аналитик низкого уровня квалификации, (2) Тестировщик ПО низкого уровня квалификации

В первый кластер попало 193 вакансии, центральными вакансиями кластера являются: «Системный аналитик», «Системный аналитик (Junior)», «Младший системный аналитик 1С», «Аналитик» и «Системный инженер». Во второй кластер попала 271 вакансия, центральными вакансиями кластера являются: «Инженер по тестированию / Software Tester», «QA инженер», «Тестировщик QA (web)», «Тестировщик советников Форекс» и «QA Engineer (Кастомизация)». Думаю, осмысленность результатов видна не вооруженным взглядом.

Помимо профессиональных навыков также возможно определить диапазон заработных плат для каждого кластера, правда насколько будут корректны полученные вилки определить сложно, поскольку рекрутеры не всегда указывают в вакансиях фактический уровень заработка, а сравнить полученные результаты с какими-либо реальными данными возможности нет, поскольку зарплатные вилки в каждой компании являются закрытой информацией.

В рамках работы:

1. Собран датасет, состоящий из 12.185 вакансий в сфере информационных технологий, в датасете в явном виде содержится 1.335 профессиональных навыков.
2. Проведена кластеризация вакансий по профессиональным навыкам 4 методами. Лучший результат обеспечил метод GMM с использованием метода понижения размерности t-SNE.
3. В результате кластеризации получено 59 кластеров, для каждого из которых составлен топ профессиональных навыков с учетом весов, определяемых позицией навыка в вакансии и его встречаемостью в датасете.

Также планируется получить экспертную оценку результатов. Для этого 3 кадровых сотрудника составят 5 случайно выбранных профессиональных портретов, и они будут сопоставлены полученным кластерам.

Результаты кластеризации могут быть улучшены за счет дополнительной обработки извлеченных профессиональных навыков, например, определения их семантической близости [16].

Кроме того, в данной работе из вакансий извлекаются явно заданные признаки, которые содержатся в поле «key_skills», и они участвуют в дальнейшей кластеризации. Однако в json-объекте вакансии есть еще поле «snippet», в котором содержаться требования вакансии, то есть неявно заданные признаки. Извлечение этих признаки с помощью решения задачи распознавания именованных сущностей с применением модели W2NER [17] также может улучшить результаты кластеризации. Поскольку в текстах требованиях вакансий более четко соблюдается правило об упоминании наиболее важных навыков в начале, веса навыков в вакансиях будут определены более корректно и это может улучшить результаты кластеризации. К тому же набор значений в поле «key_skills» ограничен самой платформой Head-Hunter, тогда как в тексте требований вакансии есть свобода действий, что, разумеется, несет как плюсы, в виде дополнительных навыков, так и минусы – дублирования, опечатки и прочее.

Литература

1. Emmel A., Cosca T. Occupational classification systems: Analyzing the 2010 standard occupational classification (soc) revision //Bureau of Labour Statistics. 2010.
2. Executive Office of the President, Office of Management and Budget. Standard Occupational Classification Manual. United States. 2018.
3. Occupational Information Network [Электронный ресурс]. Режим доступа: https://www.onetonline.org/ (дата обращения: 16.04.2022).
4. European Skills, Competences, Qualifications and Occupations [Электронный ресурс]. Режим доступа: https://esco.ec.europa.eu/en (дата обращения: 16.04.2022).
5. National Occupational Classification [Электронный ресурс]. Режим доступа: https://noc.esdc.gc.ca (дата обращения: 16.04.2022).

6. COC Groningen & Drenthe [Электронный ресурс]. Режим доступа: https://www.cocgd.nl/ (дата обращения: 16.04.2022).
7. Профессиональные стандарты - Минтруд России [Электронный ресурс]. Режим доступа: https://profstandart.rosmintrud.ru/ (дата обращения: 16.04.2022).
8. Maćkiewicz A., Ratajczak W. Principal components analysis (PCA) //Computers & Geosciences. 1993. Т. 19. №. 3. С. 303-342.
9. Arora S., Hu W., Kothari P. K. An analysis of the t-sne algorithm for data visualization //Conference on learning theory. PMLR. 2018. С. 1455-1462.
10. Morris J. N. et al. MDS cognitive performance scale© //Journal of gerontology. 1994. Т. 49. №. 4. С. M174-M182.
11. Hoecker A., Kartvelishvili V. SVD approach to data unfolding //Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment. 1996. Т. 372. №. 3. С. 469-481.
12. Reynolds D.A. et al. Gaussian mixture models //Encyclopedia of biometrics. 2009. Т. 741. №. 659-663.
13. Жамбю М. Иерархический кластер-анализ и соответствия. 1988.
14. Likas A., Vlassis N., Verbeek J. J. The global k-means clustering algorithm //Pattern recognition. 2003. Т. 36. №. 2. С. 451-461.
15. Khan K. et al. DBSCAN: Past, present and future //The fifth international conference on the applications of digital information and web technologies (ICADIWT 2014). IEEE. 2014. С. 232-238.
16. Ботов Д.С., Кленин Ю.Д. Подход к определению меры семантической близости результатов обучения //Информационные технологии и системы. 2016. С. 76-82.
17. Li J. et al. Unified named entity recognition as word-word relation classification //Proceedings of the AAAI Conference on Artificial Intelligence. 2022. Т. 36. №. 10. С. 10965-10973.

УДК 004.41

ПРОЕКТИРОВАНИЕ КЛИЕНТ-СЕРВЕРНОГО ПРИЛОЖЕНИЯ ДЛЯ РАБОТЫ С ФИНАНСОВЫМИ ЗАДОЛЖЕННОСТЯМИ

Здоровеннов М.Р.[1] (магистрант)
Научный руководитель – старший преподаватель Береснев А.Д.[1]
1 – Университет ИТМО

Аннотация
В работе рассмотрен этап проектирования мобильного приложения для работы с финансовыми задолженностями, целью которого являлась подготовка к разработке приложения, для чего были проанализированы приложения-аналоги, были определены объекты предметной области, сформирован список функциональных требований и соответствующих им сценариев использования приложения.

Ключевые слова
Финансовые операция, долговая история, финтех, финансовые технологии, клиент-серверное приложение, клиент-серверная разработка, мобильно приложение, Android, мобильная разработка, проектирование приложений.

С каждым годом все больше людей по всему миру начинают пользоваться финтех приложениями, по оценкам экспертов количество пользователей финтех индустрии ежегодно увеличивается на 15-20% [1]. Банки по всему миру начинают предоставлять свои услуги в электронном виде с помощью сервисов и приложений. Только в России в 2020-2021 годах количество загрузок банковских приложений составило 114 миллионов, что на 28,6% больше, чем в 2019-2020 годах [2]. Все больше ценится скорость операций и минимизация посредников, а оцифровывание процессов как раз ведет к снижению операционных затрат и ускорению их предоставления [3, 4, 5, 6]. Также немаловажной причиной возрастающего интереса к финансовым технологиям можно считать невероятную прибыльность финтех приложений, например чистая прибыль цифрового банка “Тинькофф банк” за два года (с 2015 по 2017 гг.) выросла почти в 10 раз [1].

Финансовые технологии не обязательно должны придумывать новые концепции, они могут просто оцифровать уже существующий процесс. Так, в Китае снижается потребительский спрос на традиционные банковские кредиты, потому что его вытесняет оцифрованный вариант [7, 8]. А в США был создан финтех сервис Lending Club, представляющий из себя одноранговую кредитную платформу. Этот сервис позволяет пользователям брать кредиты у других пользователей, а не у традиционных банков [9].

Для изучения в данной работе выбран концепт долговых займов. Предполагается что финансовые технологии могут взять на себя ответственность за хранение актуальной и достоверной информации о долговой истории пользователей.

Для оценки существующих приложений которые предоставляют необходимый функционал, были выделены следующие критерии:

1) позволяют вести историю долгов между пользователями;
2) позволяют авторизовываться пользователям чтобы обеспечить доступ к данным с любого устройства;
3) позволяют создавать долг с другим пользователем приложения, который появится в обоих аккаунтах,
4) удобство интерфейса.

Для проведения сравнительного анализа были выбраны наиболее популярные приложения из Google Play по запросу “контроль долгов”:

1. “Мои долги” от компании Anthill station.
2. “Контроль долгов” от компании CleverPumpkin.
3. “Простая Книга Долгов” от компании Breackneck Apps.

4. “Money Pro - личные финансы” от компании iBear LLC.

При помощи сравнительного анализа, результаты которого содержатся в таблице, были выделены их преимущества и недостатки.

Таблица

Сравнение сервисов-аналогов

	Мои долги	Контроль долгов	Простая Книга Долгов	Money Pro - личные финансы
История долгов между пользователями	ДА	ДА	ДА	НЕТ
Аккаунт пользователя	НЕТ	ДА	НЕТ	ДА
Возможность создавать долг с другим пользователем приложения	НЕТ	НЕТ	НЕТ	ДА
Удобство интерфейса	ДА	ДА	НЕТ	НЕТ

В процессе анализа аналогов не было найдено приложение которое бы соответствовало всем необходимым условиям.

Для описания предметной области и дальнейшего рассмотрения приложения обозначим следующие объекты:

1) пользователь – человек, группа лиц, организация, которые используют приложение;

2) долг – объект, который содержит информацию о задолженности одного пользователя другому;

3) список долгов – объект, который состоит из долгов конкретного пользователя;

4) долговая история – объект, который состоит из закрытых долгов конкретного пользователя;

5) список контактов – объект, который состоит из пользователей добавленных в данный список хозяином аккаунта.

Для проектирования будущего приложения необходимо составить и зафиксировать необходимые функции в виде функциональных требований.

Первая группа требований будет связана с аккаунтом пользователя. Это необходимо чтобы была возможна авторизация пользователя и доступ к данным с любого устройства. Выделены следующий требования:

1) у пользователя должна быть возможность создать аккаунт;

2) у пользователя должна быть возможность указывать свой уникальный идентификатор, если он еще не занят другим пользователем;

3) у пользователя должна быть возможность указывать свое имя;

4) у пользователя должна быть возможность указывать пароль от аккаунта при его создании;

5) у пользователя должна быть возможность войти в аккаунт, у пользователя должна быть возможность выйти из аккаунта;

6) у пользователя должна быть возможность восстанавливать доступ к ранее созданному аккаунту;

7) у пользователя должна быть возможность получать информацию о своем аккаунте;

8) у пользователя должна быть возможность редактировать информацию о своем аккаунте.

Вторая группа требований связана с работой пользователя с долгом. Выделены следующие требования:

1) у пользователя должна быть возможность создать новый долг, правильно указав всю необходимую информацию;
2) у пользователя должна быть возможность при создании долга указывать другого пользователя в роли должника;
3) у пользователя должна быть возможность при создании долга указывать другого пользователя в роли одалживающего;
4) у пользователя должна быть возможность отправить долг на подтверждение другому пользователю;
5) у пользователя должна быть возможность отредактировать созданный ранее долг;
6) у пользователя должна быть возможность удалить созданный ранее долг;
7) у пользователя должна быть возможность обновлять информацию о созданном ранее долге;
8) у пользователя должна быть возможность подтверждать получение долга от другого пользователя;
9) у пользователя должна быть возможность отвергать долг от другого пользователя.

Третья группа требований связана со списком долгов. Были выделены следующие требования:

1) у пользователя должна быть возможность получать список долгов связанных с данным аккаунтом;
2) у пользователя должна быть возможность переходить к конкретному долгу из списка для просмотра подробной информации;
3) у пользователя должна быть возможность сортировать список долгов по дате их создания;
4) у пользователя должна быть возможность сортировать список долгов по указанному должнику.

Четвертая группа требований связана со списком пользователей. В приложении предполагается взаимодействие различных пользователей путем заведения долгов в которых один пользователь будет представлен в роли должника, а второй в роли одалживающего. Были выделены следующие требования:

1) у пользователя должна быть возможность получать список других пользователей;
2) у пользователя должна быть возможность искать пользователя по его уникальному идентификатору;
3) у пользователя должна быть возможность искать пользователя по его имени;
4) у пользователя должна быть возможность добавлять другого пользователя в список контактов.

Пятая группа требований связана со списком контактов. Чтобы упростить поиск других пользователей в приложении будет реализован функционал “списка контактов” в котором будут хранится пользователи, добавленные хозяином аккаунта. Для работы со списком контактов были выделены следующие требования:

1) у пользователя должна быть возможность добавлять новый контакт с список контактов;
2) у пользователя должна быть возможность удалять ранее добавленного пользователя из списка контактов;
3) у пользователя должна быть возможность получать список контактов;
4) у пользователя должна быть возможность искать пользователя в списке контактов по его имени;

5) у пользователя должна быть возможность искать пользователя в списке контактов по его уникальному идентификатору.

По ранее сформированным функциональным требованиям написаны сценарии использования разрабатываемого приложения чтобы подробнее обозначить как оно должно будет работать. Сформированы следующие сценарии:

1) создание аккаунта;
2) вход в аккаунт;
3) выход из аккаунта;
4) восстановление пароля;
5) получение информации о своем профиле;
6) Редактирование учетной записи;
7) создание долга;
8) редактирование своего долга;
9) удаление долга;
10) подтверждение получения долга;
11) отвержение запроса долга от другого пользователя;
12) просмотр списка своих долгов;
13) переход к конкретному долгу из списка;
14) добавление нового контакта;
15) удаление пользователя из списка контактов;
16) просмотр списка контактов.

В ходе данной работы был проведен анализ предметной области и обзор аналогов, сформулированы основные функциональные требований к системе, создана диаграмма вариантов использования, составлена логическая модель данных, а также выбраны технологии и архитектура для реализации системы. На этом стадия проектирования может быть законченная и можно приступить к первой фазе разработки мобильного приложения.

Литература

1. Влияние цифровых технологий на сферу государственных и бизнес-услуг в России / Громов И.А. Проблемы современной экономики. 2018. 43 с.
2. Финтех - через трансформацию к новой экономике / Катрич А.С. Глобальные рынки и финансовый инжиниринг. 2017. 65 с.
3. Как создать финтех-приложение в 2023 году: Пошаговое руководство [Электронный ресурс]. Режим доступа: https://www.purrweb.com/ru/blog/kak-sozdat-fintekh-prilozhenie (дата обращения: 19.01.23)
4. Направления трансформации банковского обслуживания корпоративных клиентов в условиях перехода к цифровой экономике / Уразова С.А., Корсунова Н.Н. Вестник Ростовского государственного экономического университета (РИНХ). 2021. 194 с.
5. Модель предметной области на языке описания онтологий / Верхотурова Ю.С. Вестник Бурятского государственного университета. Математика, информатика, 2013.
6. Тенденции развития финтех-экосистемы в российской экономике / Свиридов О.Ю., Некрасова И.В. Вестник Волгоградского государственного университета. Серия 3: Экономика. Экология. 2019. 197 с.
7. Влияние финтех-компаний на банковский сектор Китая / Синь Яньлян – Инновации и инвестиции. 2022. 37 с.
8. Особенности цифровой трансформации экономики Китая / Пряжникова О.Н. Социальные и гуманитарные науки. Отечественная и зарубежная литература. Сер. 9, Востоковедение и африканистика: Реферативный журнал. 2021. 105 с.
9. Развитие рынка финансовых технологий: зарубежный опыт и отечественная практика / Ваганова О.В., Коньшина Л.А. Научный результат. Экономические исследования. 2021. 80 с.

УДК 658.5.012.7

АНАЛИЗ ПРЕДСТАВЛЕННЫХ НА РЫНКЕ CRM-СИСТЕМ ДЛЯ УЧЕБНЫХ ЯЗЫКОВЫХ ЦЕНТРОВ

Зиманова Е.А.[1] (студент)
Научный руководитель – д.э.н., профессор Максимова Т.Г.[1]
1 –Университет ИТМО
e-mail: katya.zimanova.99@mail.ru

Работа выполнена в рамках темы НИР «Сравнительных анализ существующих на рынке CRM-систем для учебных языковых центров».

Аннотация
Данная работа проводилась с целью сравнить представленных на рынке CRM-систем, которые способны обслуживать языковые школы. Для это было отобрано 19 систем. Для проведения сравнительного анализа было отобрано 38 параметров, которые были разбиты на 7 подгрупп. Методом сравнения был выбран квадрант CRM. Для каждой отобранной системе рассчитывались две координаты (X,Y). И в результате, в зависимости от того какие у системы координаты, оно располагалось в той или иной части квадранта CRM. И вследствие этого CRM-системы были поделены на четырех участников рынка: лидеры, претенденты, провидцы, нишевые игроки.
Ключевые слова
CRM-системы, сравнительный анализ, языковые центры, позиции на рынке, квадрант CRM.

Современные предприятия и учебные заведения ставят перед собой задачу автоматизации своих бизнес-процессов. В условиях взаимодействия с цифровыми технологиями и повышенной конкуренции в бизнесе, автоматизация позволяет увеличивать эффективность и оптимизировать затраты на проведение того или иного процесса. Процессы автоматизации в учебных заведениях включают в себя учет и поддержку учеников, контроль знаний и управление учебным процессом в целом. При этом, автоматизация позволяет как повысить качество образовательных услуг, так и увеличить потенциал учебного заведения и его конкурентоспособность. Таким образом, стремление к автоматизации для бизнеса и учебных заведений становится все более важным фактором развития и успешности в условиях современного мира.

В настоящее время муниципальным учреждениям для ведения учета учащихся и контроля качества обучения предлагается единое решение, которое называется Дневник.ру. Это удобная и эффективная платформа, позволяющая автоматизировать многие процессы и значительно упростить работу администрации и учителей. Однако, ситуация с частными языковыми и учебными центрами выглядит иначе. На данный момент на рынке существует более 20 готовых решений, что делает вопрос выбора наиболее подходящей платформы очень актуальным. При выборе необходимо учитывать множество факторов и условий, таких как размер и особенности центра, его потребности и потребности его клиентов, особенности обучения и т.д. От правильного выбора платформы зависит эффективность работы учебного центра и удовлетворенность клиентов, поэтому выбор нужно делать тщательно и обдуманно.

Цель данной работы упростить выбор CRM-системы при помощи составления сравнительного анализа. Для этого были поставлены и выполнены следующие задачи. Во-первых, был составлен список CRM-систем. Во-вторых, был выбран метод анализа, которым оказал квадрант CRM, логика построения которого основывается на логике построения квадранта Гартнера [1]. В-третьих, были выбраны параметры по которым был произведен сравнительный анализ. В-четвертых, был проведен сам сравнительный анализ, который распредели CRM-системы по четырем участникам квадранта CRM.

В ходе анализа было выделено общее количество 19 CRM-систем, которые прошли отбор по критерию их ориентации на языковой учебный центр. Такой список был получен при помощи обращения к сайтам-агрегаторам [2 ,3] и поисковой системе “Яндекс”Однако, дальнейшее изучение показало, что среди 19 систем имелись две, которые являются именно CRM для языковых учебных центров. Остальные 17 систем были отнесены к выборке по критерию, что они содержат в себе модули для работы с языковыми учебными центрами, но не всегда ориентированы строго на эту сферу.

Методом которым сравнивались между собой 19 систем называется квадрант CRM. Квадрант CRM – это метод, который заключается в расположении CRM-систем на плоскости квадранта в соответствии с полученными координатами (X, Y). При этом координаты X и Y, а также размер, определяются в результате сравнительного анализа. Ось X соответствует полноте функционала, то есть чем больше функций предоставляет CRM-система, тем больше ее значение на этой оси. Ось Y отражает востребованность системы у потребителей. Чем более популярна система, тем выше ее значение на этой оси. Размер же определяется ценой системы. В результате такого анализа CRM-системы располагаются в соответствующих квадрантах.

Координата Y представляет собой комбинацию двух значений: общей суммы отзывов на различных сайтах-агрегаторах [2, 3] и отзывов на официальных сайтах CRM-систем. В свою очередь, размер определяется в зависимости от цены CRM за 1 месяц на 400 учеников. Расчет размера осуществляется по формуле:

$$M = I \div N,$$

где M - размер, N - цена за ежемесячное обслуживание центра на 400 учеников. при этом чем больше размер, тем дешевле будет цена.

Для определения координаты X было выбрано 38 параметров, каждый из которых представляет собой присутствие или отсутствие функции или возможности в CRM-системе, обозначенное соответствующим значением 1 или 0. Но для учета неравнозначности этих параметров был использован метод экспертной оценки.

Параметры были разделены на 7 подгрупп, каждая из которых содержала от 4 до 6 параметров. Они включали Денежные средства и кассовый журнал (табл. 1), контроль администратора (табл. 2), работу администратора (табл. 3), педагогов (табл. 4), сбор и анализ информации (табл. 5), клиентов (табл. 6) и маркетинг (табл. 7). Каждому параметру был присвоен вес, при расчете которого учитывалось соотношение весов других параметров в подгруппе. Таким образом, сумма весов в каждой подгруппе равнялась единице.

Таблица 1

Денежные средства и кассовый журнал

Параметр	вес
Бухгалтерский учёт	0,25
Онлайн-платежи	0,3
Расчет зарплаты	0,2
Учет доходов и расходов	0,25

Таблица 2

Контроль администраторов

Параметр	вес
Мониторинг эффективности персонала	0,15
Настройки доступа	0,2
Отслеживание времени	0,13
Отчёты	0,19
Разделение лидов на сегменты	0,13
Управление абонементами	0,2

Таблица 3

Работа администратора

Параметр	вес
Многопользовательский режим	0,2
Добавление сотрудников	0,25
Задачи и напоминания	0,15
Корпоративная программа	0,1
Личный кабинет	0,2
Шаблоны	0,1

Таблица 4

Предагоги

Параметр	вес
Личный кабинет ученика	0,2
Отслеживание посещаемости	0,25
Проведение тестирования	0,1
Прогресс обучения	0,2
Расписание занятий	0,15
Создание курсов	0,1

Таблица 5

Сбор и анализ информации

Параметр	вес
База клиентов	0,2
Воронка продаж	0,15

продолжение таблицы

Параметр	вес
История взаимодействия с клиентом	0,15
Отчётные документы	0,2
Отчёты и аналитика	0,2
Хранение файлов	0,1

Таблица 6

Клиенты

Параметр	вес
Email-рассылка	0,2
Интеграция с социальными сетями	0,15
Мобильная версия	0,1
Предварительная запись клиентов	0,2
Рассылка уведомлений	0,15
Телефония	0,2

Таблица 7

Маркетинг

Параметр	вес
Инструменты маркетинга	0,25
Настройки скидок	0,2
Система дисконтных карт	0,2
Создание коммерческих предложений	0,35

Наконец, координата X была определена как сумма взвешенных 38 параметров. Вес каждого параметра учитывал значимость этого параметра по сравнению со всеми другими параметрами, присутствующими в CRM-системе. Такой подход позволил учесть все важные параметры и получить объективную оценку качества системы.

После того как координаты X и Y определены, и был рассчитан размер, строится сам квадрат (рисунок).

В результате проведенного анализа можно выделить четыре группы игроков, каждая из которых имеет свои особенности и преимущества. Первая группа – лидеры, которые занимают ключевое положение на рынке благодаря высокому качеству своих продуктов и широкой известности. К ним относятся фирмы Мой класс, Alfa CRM и CRM Отмечалка. Вторая группа – провидцы, которые не попали в число лидеров, но продолжают успешно конкурировать на рынке и готовы к изменениям в отрасли. В список провидцев не вошло ни одной CRM из первой группы. Третья группа – нишевые игроки, которые специализируются на определенных рыночных сегментах и пытаются занять свою нишу. Среди них можно назвать AXKED, OKOCRM, Clients24, CRM

PushtoLearn, Контур Класс, Топкурс. Наконец, четвертая группа – претенденты, которые только начинают свой путь на рынке CRM и пытаются пробиться среди более опытных компаний. В эту группу входят CRM HOLLIHOP, Big Ben CRM, ListOk CRM, Параплан CRM, ImpulseCRM, Tallanto, Faience, 2Курс CRM, Roboseller, Olla crm bercana.

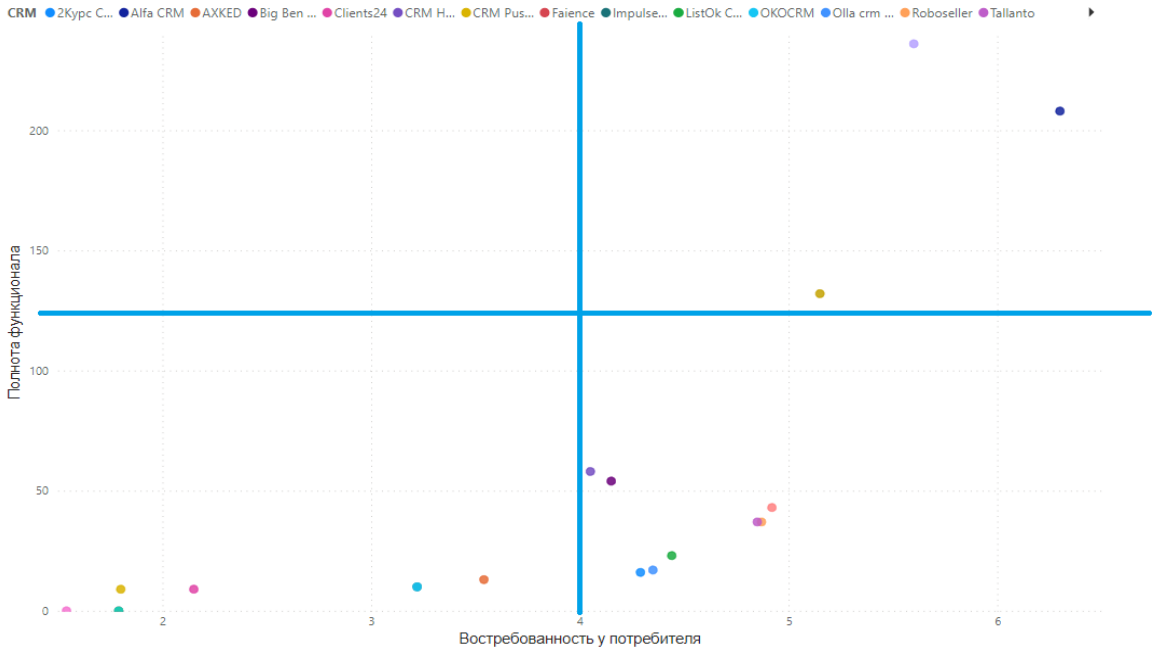


Рисунок. Квадрант CRM

При выборе CRM-системы для языковой школы имеет значение несколько важных параметров, которые необходимо учитывать. В первую очередь, необходимо обратить внимание на полноту функционала системы, так как она должна удовлетворять требованиям и потребностям конкретной школы. Второй параметр - востребованность у потребителя, то есть, система должна быть популярной и широко используемой, чтобы можно было использовать уже проверенные и эффективные методы управления учебными процессами. И, наконец, еще одним важным параметром является ежемесячная стоимость системы, которая должна быть не только доступной и соответствовать финансовым возможностям языковой школы, но и оправдывать свою цену эффективностью и удобством использования. Таким образом, правильный выбор CRM-системы может значительно ускорить и упростить процессы управления учебным процессом в языковой школе, отразится на качестве обучения и удовлетворении потребностей учеников.

Литература

1. Magic Quadrant Research Methodology [Электронный ресурс]. Режим доступа: https:// www. gartner.com/en/ research/methodologies/ magic-quadrants-research (дата обращения: 11.12.2022).
2. ТОП-8 Лучших CRM для учебного центра 2022 [Электронный ресурс]. Режим доступа: https://a2is.ru/ catalog/crm-dlya- uchebnogo-tsentra?page=1 (дата обращения: 06.12.2022).
3. 10 лучших CRM для языковых школ. Рейтинг и сравнение CRM [Электронный ресурс]. Режим доступа: https://crmindex.ru/for/yazukovaya_shkola (дата обращения: 06.12.2022).

УДК 004

ПРОЕКТИРОВАНИЕ ФУНКЦИОНАЛЬНОСТИ И ДИЗАЙНА МОБИЛЬНОГО ПРИЛОЖЕНИЯ ДЛЯ КОЛЛАБОРАЦИИ МУЗЫКАНТОВ

Коровин А.И.[1] (студент)
Научный руководитель – доцент Валитова Ю.О.[2]
1 – Университет ИТМО
e-mail: akorovin815@gmail.com

Аннотация
В работе было проведено исследование существующих аналогов, с помощью которых можно находить музыкантов. Были выявлены преимущества и недостатки этих приложений. Был проведен опрос представителей целевой группы пользователей приложения. Он позволил установить, что мобильное приложение будет востребовано, и определить желаемую функциональность приложения.
Ключевые слова
Приложение, поиск единомышленников, музыка.

Идея моего проекта появилась из хобби: я очень люблю рок-музыку. Однажды я захотел собрать группу или найти место в существующей, но так как род занятий у меня совсем иной, и, следовательно, я не принадлежал к музыкальной публике своего города, это стоило бы слишком больших усилий.

Практически любая группа, даже самая известная, с большим трудом находит правильных людей в свой состав. Не думайте, что это касается только начинающих групп, это общая проблема, но не все знают как с ней справиться.

И тогда я пришел к выводу, что нужно подумать о сотнях таких же людей и сделать обычное приложение, на которой в пару кликов смогут отметиться все музыканты.

Ведь никому из них абсолютно не составит труда это сделать, а выгоды для пользователей очевидны: единомышленники из вашего города гарантированно вас заметят. Да и в пределах своей страны иногда любопытно посмотреть, кто где живет и на чем играет. Возможно, вы вообще хотите переехать в другую страну — тогда на знакомства уйдет колоссальное количество времени. А с таким приложением вы сможете уже в первый день найти себе единомышленников.

Приложения для поиска музыкантов являются важными инструментами для любой группы или музыканта, которым необходимо найти людей. Был проведен анализ аналогов, в которых есть возможность поиска единомышленников, были выявлены преимущества и недостатки этих платформ.

Среди недостатков особо выделяются:

1) неудобный интерфейс;
2) приложения больше не поддерживаются разработчиками;
3) нет возможности настройки фильтра;
4) ограниченность поисковых параметров;
5) ограниченный доступ к контактной информации;
6) платный контент.

Преимущества:

1) широкий выбор музыкантов;
2) возможность прослушивания работ музыкантов;
3) возможность поиска музыкантов не только в своем городе, но и в других городах.

Был проведен опрос представителей целевой группы пользователей приложения. Целью опроса было выявление используемых приложений для поиска музыкантов на данный момент, узнать преимущества, недостатки и востребованность. Опрос показал такие результаты. Основные используемые систем на данный момент:

1) vampr;
2) soundcloud;
3) musicforums.ru.

Преимущества:

1) популярность;
2) доступность;
3) возможность ознакомиться с работами музыканта;
4) возможность переписки с музыкантом.

Основные недостатки:

1) ограниченная функциональность бесплатных версии;
2) старый интерфейс.

На основании проведенного исследования в виде глубинного интервью и анализа аналогов. Так же user story и диаграммы вариантов использования [2] (рисунок). Были сформулированы требования по функционалу проектируемого приложения для коллаборации музыкантов:

1) поиск по геолокации;
2) возможность точной настройки поиска;
3) возможность загрузки видео и фото;
4) возможность загрузки собственных музыкальных записей;
5) чаты между пользователями.

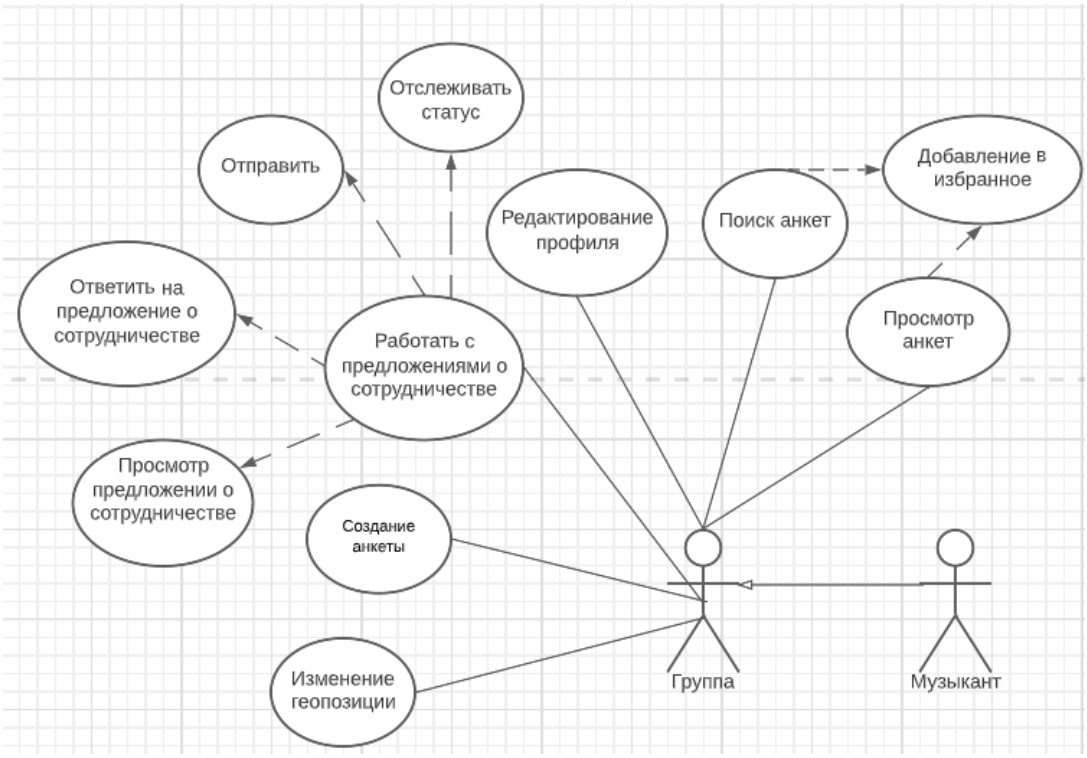


Рисунок. Диаграмма вариантов использования

Зная функциональность мобильного приложения, можно было переходить к разработке интерфейсов, реализующих ее. Представьте, что вы создали классное приложение с широким функционалом и возможностями. Но иконка не выделяется среди тысяч других, приложение неудобное, неправильно подобраны цвета, кнопки размещены так, что большим пальцем на смартфоне их просто не достать. Такое приложение не будет иметь успеха. Поэтому разрабатывается интерфейс мобильного приложения, позволяющий сделать визуальную часть удобной, простой, привлекательной и логичной.

Интерфейс мобильного приложения как раз отвечает за то, чтобы все вышеуказанные условия выполнялись. Это комплекс, позволяющий четко понять, как должно выглядеть приложение, где располагаются все элементы, какая логика программы. Для этого были изучены современные требования к интерфейсу мобильных приложений. Ведь проектируя интерфейс приложения для мобильных устройств, нужно знать об особенностях дизайна, и дело здесь не только в разрешении экрана. То, что хорошо смотрится с ноутбука или стационарного компьютера, может совершенно не подходить для мобильного устройства, и наоборот [1]. На основании изучения теории был разработан дизайн мобильного приложения. На основе разработанного дизайна и определенной функциональности мобильного приложения в онлайн-сервисе Figma были разработаны интерфейсы приложения.

Данный сервис для поиска музыкантов поможет желающим бесплатно найти себе единомышленников. Для разработки такого приложения необходимо провести детальный анализ конкурентов, чтобы выявить их преимущества и недостатки, а также определить уникальный функционал, который будет отличать ваше приложение от других. В итоге, разработка приложения для поиска музыкантов может быть успешной, если разработчики смогут создать уникальный продукт, который будет отличаться от конкурентов, обеспечить удобный и простой интерфейс для пользователей и использовать современные технологии для реализации функционала. В рамках работы были изучены аналоги приложения, изучены правила составления user story. Также была разработана диаграмма вариантов использования.

Литература

1. Создание интерфейса мобильных приложений. [Электронный источник] Режим доступа: https://wezom.com.ua/ blog/dizajn-interfejsov- mobilnyh-prilozhenij (Дата обращения 15.02.2023).
2. Правила написания user story. [Электронный источник] Режим доступа: https://habr.com/ ru/post/577420/ (Дата обращения 15.02.2023).

УДК 004.85

РАЗРАБОТКА МОДУЛЯ БЕЗОПАСНОСТИ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА НА ОСНОВЕ ОГРАНИЧЕНИЙ ПОВЕДЕНИЯ СИСТЕМЫ ПОДДЕРЖКИ И ПРИНЯТИЯ РЕШЕНИЙ

Краснов Т.А.[1], Компаниец Р.И.[1], Дудкин А.С.[1]
Научный руководитель – к.т.н. Дудкин А.С.[1]
1 – Военно-космическая академия им. А.Ф. Можайского
e-mail: vka@mil.ru

Аннотация
В данной статье рассматривается решение проблемы безопасности искусственного интеллекта путем ограничения поведения его системы поддержки принятия решений. Автором обращено внимание на необходимость ввода ограничений в виде модуля безопасности, который будет контролировать поведение системы и предотвращать ее от принятия решений, с высокой долей вероятности являющихся ошибочными.

Ключевые слова
Искусственный интеллект, система поддержки принятия решений, ограничение поведения, многоуровневая система, модуль безопасности.

На текущий момент времени искусственный интеллект (далее – ИИ) развивается быстрыми темпами. Системы ИИ могут превосходить людей в инженерных и научных способностях, а автономные системы способны учиться по ходу работы. Такие проблемы, как прерывание данных систем, влияют на их поведение и на то, как отсутствие руководителя повлияет на ее действия, которые они захотят предпринять. Помимо этого, системы ИИ могут небезопасно исследовать свою среду, тем самым вызывая негативные побочные эффекты. В результате возникает вопрос о разработке системы или методики, которая позволит минимизировать риск принятия системой поддержки принятия решений (далее – СППР) решений, являющихся с высокой долей вероятности ошибочными.

Обратимся к такому понятию, как процессная аналитика (далее – ПА) – аналитическому подходу, направленному на построение исчерпывающего и объективного видения процессов на основе фактических данных. Механизмы ИИ обеспечивают дальнейшее совершенствование процессов принятия решений за счет получения знаний и адаптации рабочих процессов: прогнозирование переменных позволяет добавлять информацию, тем самым оптимизируя решения в среднесрочной и долгосрочной перспективе [1]. Смоделированные процессы, улучшенные с помощью ИИ, называются моделями ПА. Их объединение вместе с ИИ составляет СППР, способную прогнозировать качество процессов и риски для сотрудников, а также управлять человеческими ресурсами. СППР, реализующая многоуровневую логику, может быть хорошим решением для повышения эффективности процессов мониторинга в отношении сценариев работоспособности, интегрирующих алгоритмы прогнозирования ИИ [2]. Тем самым попытаемся реализовать модель ПА для обнаружения, мониторинга и улучшения процессов с помощью многоуровневой СППР, таким образом ограничивая ее поведение и сводя к минимуму принятие неверных решений.

Для решения текущей проблемы может использоваться модуль безопасности ИИ, основанный на многоуровневой СППР, где каждый уровень характеризует условие оповещения. Состояние предупреждения будет определять уровень риска: первый уровень риска (Уровень 1) – состояние незначительной опасности, второй уровень риска (Уровень 2) более опасен, n-й уровень (Уровень n) представляет наибольшую опасность. На рисунке 1 показана блок-схема предлагаемой многоуровневой модели,

интегрирующая СППР. Она структурирована для работы с различными обработками данных, представляющими конкретный процесс принятия решений. Алгоритмы искусственного интеллекта применяются в основном в процессе для контроля условий безопасности работников. Выходные данные модели определяют возможные условия оповещения и процедуры, которые следует принять.

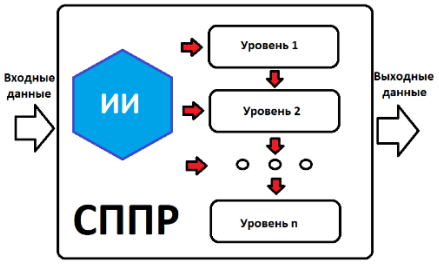


Рис. 1. Блок-схема используемой многоуровневой модели

Первый уровень оповещения - вспомогательный рабочий процесс, образованный первого уровня СППР, основанный на одновременном анализе, проверке порогового значения (обнаружение условий оповещения за короткий период), и прогнозировании с помощью ИИ, поддерживающем оценку риска в среднесрочном или долгосрочном периоде. Обработанный набор данных является исходным набором данных, полученный путем фильтрации наиболее релевантных атрибутов. Смоделировано логическое условие: для положительной проверки контролируемого значения переменных (значения, не превышающие определенные пороговые) гарантируется условие безопасности, и никаких вмешательств не требуется. Кроме того, в отрицательном случае система обеспечивает первое состояние оповещения, позволяя вмешиваться и данные уже обрабатываются во второй СППР, предоставляющий более подробную информацию о рисках для работников (более высокий уровень риска). В случае прохождения положительной проверки для дополнительной безопасности может быть выполнен рабочий процесс второго уровня СППР. Блок-схема работы одного из уровней показана на рисунке 2.

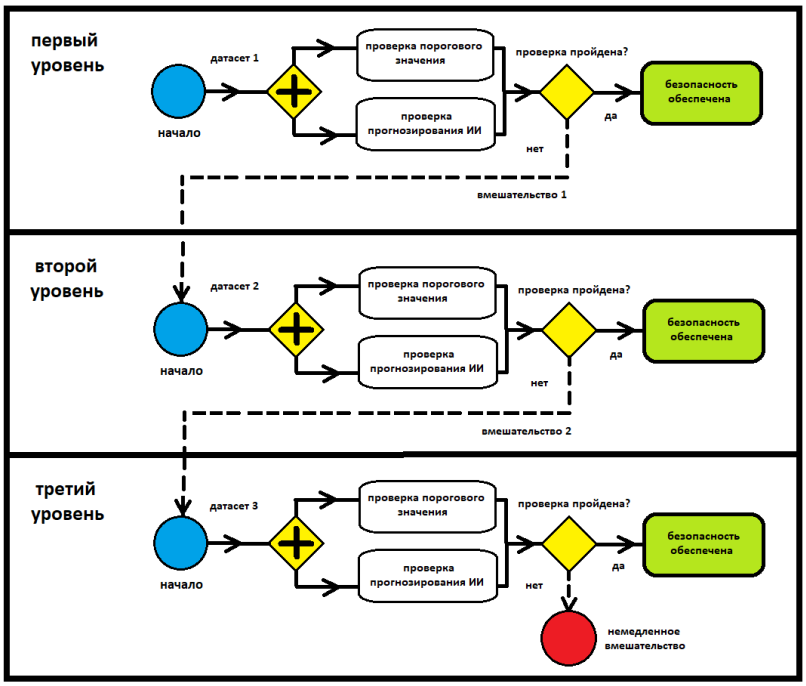


Рис. 2. Блок-схема рабочего процесса модели, имеющей три уровня СППР

Второй уровень оповещения представляет рабочий процесс, включающий ‘И’ - логическое условие между истинными результатами уровня 1 оповещения СППР и других переменных. Для положительной проверки контролируемые переменные также не должны превышать определенных пороговых значений. В этом случае процесс можно считать безопасным. Кроме того, в случае отрицательного результата система предоставляет второе условие оповещения, представляющий более серьезный риск, и данные обрабатываются в третьем уровне анализа СППР.

Третий уровень оповещения представляет собой рабочий процесс, имеющий ту же структуру, что и предыдущий. Если третий уровень также указывает на риск, выполняется немедленное вмешательство (состояние более серьезного риска).

Такая структура может включать в себя n уровней. Крайний уровень СППР в случае отрицательного результата говорит о гарантированном риске для работников и отклоняет принятие решения. Количество используемых уровней зависит исключительно из соображений рациональности и использования ресурсов системы. Данный метод может быть применим к любой СППР системы ИИ и подогнан под ее правила и организационную структуру.

Таким образом, предложенная теоретическая многоуровневая модель, лежащая в основе разработанного модуля безопасности, позволяет свести к минимуму вероятность принятия СППР системы ИИ ошибочных решений, тем самым делаю систему безопасной для практического использования. Также особенностью данного модуля является его применимость к большинству существующих СППР систем ИИ.

Литература

1. Интеллектуальные системы поддержки принятия решений // Хабр URL: https://habr.com/companies/ods/articles/359188. (дата обращения: 12.02.2023).
2. Системы поддержки принятия решений: проектирование, применение, оценка эффективности // Studfile URL: https://studfile.net/preview/5367046. (дата обращения: 12.02.2023).

УДК 004.8

СРАВНЕНИЕ ЭФФЕКТИВНОСТИ ПРОГНОЗИРОВАНИЯ ВРЕМЕННЫХ РЯДОВ С ПОМОЩЬЮ МАШИННОГО ОБУЧЕНИЯ И РАЗЛИЧНЫХ ВИДОВ НЕЙРОННЫХ СЕТЕЙ

Кривоносов Е.Д.[1] (студент), **Нечкасова О.А.**[1] (студент), **Перл И.А.**[1]
Научный руководитель – к.т.н., доцент Перл И.А.[1]
1 – Университет ИТМО
e-mail: trottimus.2001@mail.ru, lesyanechkasova@gmail.com

Аннотация
В докладе обсуждаются вопросы, связанные с анализом временных рядов с использованием машинного обучения. Представлено сравнение эффективности прогнозирования временных рядов на примере различных видов нейронных сетей.
Ключевые слова
Машинное обучение, временные ряды, рекуррентные нейронные сети, сети с долгой краткосрочной памятью, сверточные нейросети, пространственно-временные нейросети.

Повышение точности прогнозирования временных рядов в настоящее время является актуальной и важной задачей во многих областях, в особенности на финансовых платформах. Способы решения данной проблемы обсуждались еще до возникновения машинного обучения. Поэтому на данный момент существует много различных методов для прогнозирования временных рядов, однако нейронные сети стали одним из самых эффективных инструментов для решения этой проблемы. За все время было создано множество моделей для предсказания, каждая из которых подходит под свои конкретные условия.

Эту же задачу по прогнозированию данных нам необходимо было решить для внедрения в платформу sdLitica, предназначенную для анализа временных рядов. Это один из проектов, который разрабатывается в лаборатории COSM (Cloud Oriented Solutions and Modeling Laboratory) на базе университета ИТМО. Прежде чем приступить к реализации этой задачи мы решили выяснить, какое решение лучше всего подойдет для нашего конкретного случая.

Основная цель этой статьи - анализ четырех различных типов нейронных сетей для прогнозирования временных рядов: рекуррентные нейронные сети (RNN), сети с долгой краткосрочной памятью (LSTM), сверточные нейронные сети (CNN) и пространственно-временные нейронные сети (STNN).

Прогнозирование временных рядов — это процесс прогнозирования будущих значений, основанный на анализе предыдущих данных. Эта задача имеет большое значение во многих областях, таких как экономика, финансы, метеорология, медицина и т. д. Прогнозирование временных рядов может быть достаточно сложной задачей, особенно в случае, когда в данных есть множество факторов, которые могут влиять на их поведение. Машинное обучение может помочь в решении этой задачи, используя различные методы, включая нейронные сети.

1. Рекуррентные нейронные сети (RNN) – это класс нейронных сетей, которые имеют способность сохранять информацию из предыдущих шагов и использовать ее для принятия решений на следующем шаге. Это свойство называется «авторегрессия». RNN имеет специальный нейрон, называемый «состояние ячейки», который сохраняет информацию из предыдущих шагов. RNN обычно используется для анализа последовательных данных, таких как ежедневные изменения цен на акции, прогнозирование погоды и т.д. RNN может иметь проблему «затухающего градиента», когда веса обновляются в процессе обучения, и некоторые значения становятся слишком маленькими, что затрудняет обучение [1].

2. Сети с долгой краткосрочной памятью (LSTM) – это модификация рекуррентных нейронных сетей, разработанная для решения проблемы затухающего градиента. LSTM также имеет состояние ячейки, но вместо одного слоя обработки информации оно имеет три: входной, выходной и забывающий. Это позволяет LSTM «запоминать» важную информацию, не теряя ее в процессе обучения. LSTM широко используется в задачах, связанных с обработкой текста, генерированием текста и музыки, а также в обработке звука.

3. Сверточные нейронные сети (CNN) – это класс нейронных сетей, используемых для анализа изображений, но они также могут быть использованы для анализа временных рядов. Они работают путем сканирования данных на различных уровнях детализации. CNN используется для анализа нескольких каналов данных, таких как временные ряды, где каждый канал представляет различные аспекты исходных данных.

4. Пространственно-временные нейронные сети (STNN) – это класс нейронных сетей, используемых для анализа пространственных и временных данных. Они могут работать с данными, в которых каждая точка данных содержит несколько значений, например изображения. Это позволяет им работать с многомерными данными, что делает их эффективными в задачах прогнозирования временных рядов.

Анализ и сравнение позволяют оценить эффективность каждой модели и выбрать наиболее подходящую для конкретной задачи прогнозирования временных рядов. Для анализа были выбраны несколько датасетов: данные фьючерсов МосБиржы за 2018 год, ежечасовые метеоданные в период с 2006 по 2016 год. Эти данные частично представлены в таблицах 1 и 2.

Таблица 1

Пример использованных данных с МосБиржи

<DATE>	<TIME>	<OPEN>	<HIGH>	<LOW>	<CLOSE>	<VOL>
03/01/18	10:05:00	113670.0	115990.0	1136710.0	115990.0	4
03/01/18	10:15:00	113430.0	113430.0	113430.0	113430.0	1
…	…	…	…	…	…	…
20/12/18	18:45:00	109380.0	109390.0	109380.0	109380.0	28

Таблица 2

Пример использованных данных о погоде

Formatted Date	Temperature (C)	Apparent Temperature	Humidity	Wind Speed (km/h)	Visibility (km)	Pressure (millibars)
2006-01-01 00:00:00+01:00	0.577778	-4.050000	0.89	17.1143	9.9820	1016.66
2006-01-01 01:00:00+01:00	1.161111	-3.238889	0.85	16.6152	9.9015	1016.15
…	…	…	…	…	…	…
2016-12-31 23:00:00+01:00	-0.677778	-3.888889	0.92	9.2253	8.8711	1020.72

В первом случае определить зависимость довольно сложно, потому что без учета дополнительных факторов, которые могут влиять на результат изменения переменных,

зачастую эти данные представляют случайное блуждание с возможным наличием трендов и цикличности. Поэтому, как и во втором случае при использовании метеоданных, где есть сезонность, данные предстоит сначала привести к стационарному временному ряду.

Рассмотрим стационарность временного ряда, характеризующего погодные условия в течение месяца. О наличии сезонности и трендов можно судить по критерию Дики-Фуллера, который в нашем случае равен 0.3906, или же, что является более наглядным, по графику автокорреляции. На рисунке 1 (а) вполне можно заметить сезонность при значении лага = 24. После дифференцирования временного ряда по рисунку 2 (б) и значению критерия, меньшем 0.05, можно принять гипотезу о стационарности ряда.

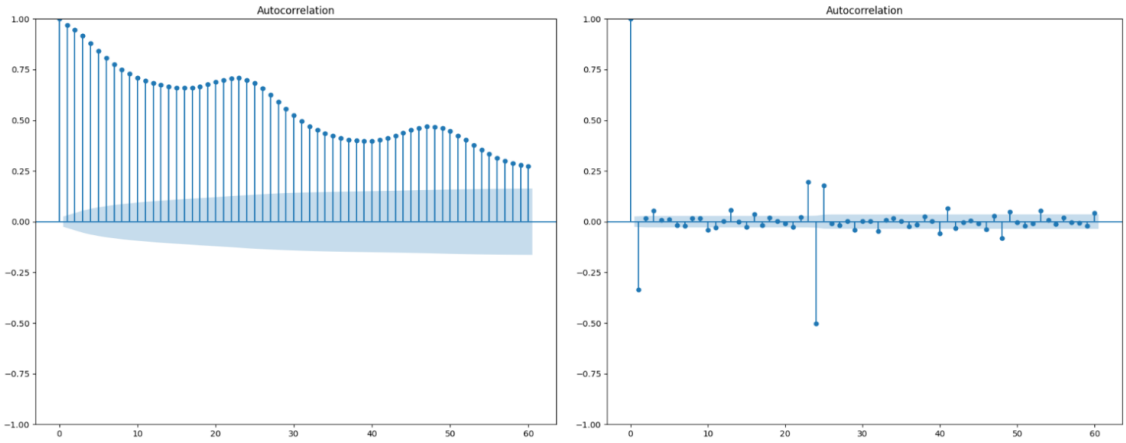


Рис. 1. Автокорреляция метеоданных:
а – исходные данные, б – дифференцированные данные

После нормирования и преобразования на основе имеющихся значений модели четырех нейронных сетей были обучены и протестированы. Для разработки мы использовали библиотеку TensorFlow и специальную надстройку Keras для более удобной конфигурации и управления моделями.

На рисунке 2 представлены результаты прогнозирования временных рядов с помощью сети с долгой краткосрочной памятью (LSTM). Красным цветом представлены данные, на основе которых модель обучалась. Видно, что зеленая линия (предсказанные значения) достаточно точно повторяют синюю линию (исходных значений).

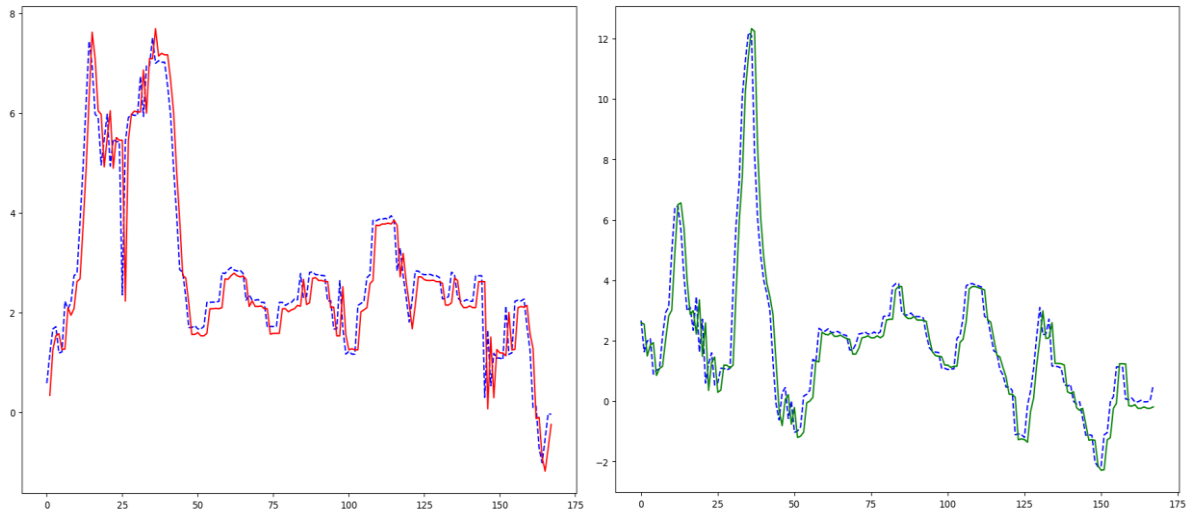


Рис. 2. Прогнозирование значений с помощью LSTM

В экспериментах использовались метрики точности, такие как средняя абсолютная ошибка (MAE), средняя относительная ошибка (MRE) и среднеквадратичное отклонение (RMSE). Они позволяют оценить точность прогнозирования каждой модели.

В рамках исследования мы провели два эксперимента: в первом использовали одномерный временной ряд (набор данных, состоящий из одной серии наблюдений с временным упорядочением), во втором – многомерный временной ряд с указанием четырех дополнительных параметров вместо одного.

В результате было установлено, что для первого эксперимента LSTM является наиболее эффективной моделью, показывая наименьшие ошибки в тестировании. В то же время, CNN и STNN показали сравнительно худшие результаты. Наше предположение о причине такого поведения состояло в том, что эти модели лучше всего подходят для обработки многомерных данных, а не последовательных временных рядов. В этом мы убедились, анализируя результаты второго эксперимента. В этом случае наиболее точные результаты дали модели STNN и CNN. Причем судя по метрикам, точность предсказания оказалась выше, чем в первом эксперименте с наиболее эффективной моделью LSTM.

Таким образом, каждая модель имеет свои особенности и подходит для определенных задач. RNN и LSTM хорошо работают с последовательными данными, в то время как CNN и STNN лучше всего подходят для обработки многомерных данных. Поэтому выбор модели зависит от конкретной задачи прогнозирования временных рядов и характеристик исходных данных.

Анализ и сравнение моделей позволяют определить наиболее эффективную модель для прогнозирования временных рядов с использованием машинного обучения и нейронных сетей. Это помогает решить актуальную задачу прогнозирования, которая имеет широкое применение в финансовых, промышленных и других сферах.

Кроме того, использование нейронных сетей для прогнозирования временных рядов имеет ряд преимуществ. Во-первых, это позволяет автоматизировать процесс прогнозирования и снизить вероятность человеческой ошибки. Во-вторых, нейронные сети могут улавливать сложные нелинейные зависимости между данными, что позволяет повысить точность прогнозирования [2].

Однако у использования нейронных сетей также есть свои ограничения. Во-первых, для обучения и прогнозирования временных рядов необходимо иметь большое количество данных, что может быть проблематично в некоторых случаях. Во-вторых, обучение и настройка нейронных сетей требует значительных вычислительных ресурсов, что может быть затратно по времени и стоимости. В-третьих, помимо построения модели перед работой с данными нужна их предварительная обработка: разделение на обучающую и тестовую выборки, корректное нормирование по строкам или столбцам и преобразование формата (преобразование столбцов в массив) [3].

Тем не менее применение нейронных сетей для прогнозирования временных рядов продолжает развиваться и улучшаться. В настоящее время исследователи работают над созданием новых моделей, которые будут более эффективными и универсальными для различных типов временных рядов.

Вывод

В заключение, прогнозирование временных рядов с использованием машинного обучения и нейронных сетей является актуальной и важной задачей во многих областях. Анализ и сравнение различных моделей позволяют выбрать наиболее подходящую модель для конкретной задачи. При этом необходимо учитывать особенности и ограничения каждой модели, а также возможности и доступность данных и вычислительных ресурсов.

Литература

1. Хайндман Р. Прогнозирование: принципы и практика / Хайндман Р., Атанасопулос Дж. Москва: ДМК-Пресс. 2023.
2. Нильсен Э. Практический анализ временных рядов: прогнозирование со статистикой и машинное обучение / Нильсен Э. Москва: Вильямс. 2021.
3. Шакла Н. Машинное обучение и TensorFlow / Шакла Н. Санкт-Петербург: Питер. 2019.

УДК 004.8

КЛАССИФИКАЦИЯ ЖАНРА МУЗЫКАЛЬНОГО АЛЬБОМА ПО ОБЛОЖКЕ С ПОМОЩЬЮ СВЕРТОЧНЫХ НЕЙРОННЫЙХ СЕТЕЙ

Лакиза А.Н.[1] (студент)
Научный руководитель – к.п.н., доцент факультета инфокоммуникационных технологий Валитова Ю.О.[1]
1 – Университет ИТМО
e-mail: alelakiza@yandex.ru

Аннотация
В ходе работы был собран набор данных, состоящий из более 5 тысяч изображений обложек 4-х музыкальных жанров: джаз, электронная музыка, хип-хоп и поп-музыка. Была проведена предобработка изображений и обучена сверточная нейронная сеть, которая на тестовой выборке смогла показать точность 52%. Проведена интерпретация предсказаний нейронной сети, благодаря которой был сделан вывод о том, что в ходе обучения нейронная сеть находит общие закономерности в обложках альбомов одного музыкального жанра. На основе проделанной работы можно сделать вывод, что сверточные нейронные сети могут решать задачу классификации жанра музыкального альбома по обложке.

Ключевые слова
Классификация изображений, сверточные нейронные сети, сеть AlexNet, интерпретация, SHAP.

С середины прошлого века все музыкальные произведения издаются вместе с обложкой. Виниловые пластинки и CD-диски продавались в упаковках с изображением обложки. Сегодня в музыкальных стриминговых сервисах изображение обложки композиции занимает центральное место на экране гаджета. Такая важность обложки обусловлена тем, что она уже давно стала неотъемлемой частью музыки, она задает настроение композиции, заставляет слушателя испытывать определенные чувства, чтобы глубже погрузить в атмосферу песни.

В обложках альбомов и песен одного жанра иногда прослеживаются общие черты, на которые слушатель обычно не обращает внимание, но по которым можно сделать предположение о том, к какому жанру относится данная композиция. Именно поэтому люди, которые часто слушают музыку, умеют угадывать жанр композиции исключительно по обложке.

В связи с активным развитием искусственного интеллекта возник вопрос – сможет ли нейронная сеть улавливать те закономерности, которые замечает человек в обложках. Исследованию этого вопроса и посвящена данная работа.

Для данной работы было решено собрать собственный набор данных, так как те, которые есть в открытом доступе, не имеют изображений обложек российских исполнителей. Чтобы собрать набор данных, необходимо к одному из музыкальных стриминговых сервисов. В данной работе изображения были собраны с помощью API сервиса Deezer [1]. В отличие от API других музыкальных сервисов, в Deezer API у объектов «альбом» есть атрибут «жанры».

Взаимодействие с Deezer API происходило с помощью библиотеки для языка программирования Python deezer-python. В результате работы с API удалось собрать 5751 изображение по 4-м жанрам: джаз, электронная музыка, хип-хоп и поп-музыка. Все изображения имели размер 250 на 250 пикселей.

Для того, чтобы использовать изображения в процессе обучения нейронной сети, необходимо переопределить класс torch.data.utils.Dataset из фреймворка для работы с нейронными сетями PyTorch.

Все изображения были поделены на обучающую и тестовую выборку в соотношении 3 к 1. В итоге получилось 4600 изображений в обучающей выборке и 1151 изображение в тестовой выборке соответственно.

Первые попытки обучить нейронную сеть на наборе данных, состоящем из 4600 изображений, не увенчались успехом, тогда было решено увеличить объем обучающей выборки посредством аугментации изображений. Аугментация изображений производилась с помощью библиотеки для языка Python под названием Albumentations. После аугментации объем обучающей выборки для нейронной сети вырос с 4600 до 41400 изображений.

В исходной задаче есть изображение, которое необходимо отнести к одному из 4 жанров. Это задача классификации изображений. В подавляющем большинстве случаев такие задачи решаются с помощью сверточных нейронных сетей.

В данной работе было принято решение воспользоваться сверточной нейронной сетью с архитектурой AlexNet [2].

В ходе серии экспериментов с разными гиперпараметрами сети лучше всего себя показала нейронная сеть, обученная с 50 эпохами, размером батча 256 изображений и скоростью обучения 0,001. На тестовой выборке, включающей в себя 1151 изображение, лучшей нейронной сети удалось добиться точности 52%. Таким образом, нейронная сеть успешно предсказывает жанр каждой второй обложки.

Однако без ответа остался вопрос – удается ли нейронной сети уловить общие закономерности в обложках альбомов одного жанра, вдруг такая точность является лишь случайностью. Для ответа на этот вопрос было решено провести интерпретацию работы нейронной сети. Одной из самых популярных библиотек для интерпретации предсказаний нейронной сети на языке Python является библиотека SHAP [3]. Для расчета важности признаков библиотека использует значения вектора Шепли. В задачах классификации изображений интерпретатор показывает на исходном изображении, какие пиксели привнесли наибольший вклад в сторону предсказания одного из жанров.

Наблюдение за интерпретацией работы нейронной сети позволило сделать вывод нейронная сеть в обложках альбомов жанра «джаз» обучилась замечать воротник белой рубашки. Об этом говорят красные пиксели в данных областях на изображениях во втором столбце рисунка 1.

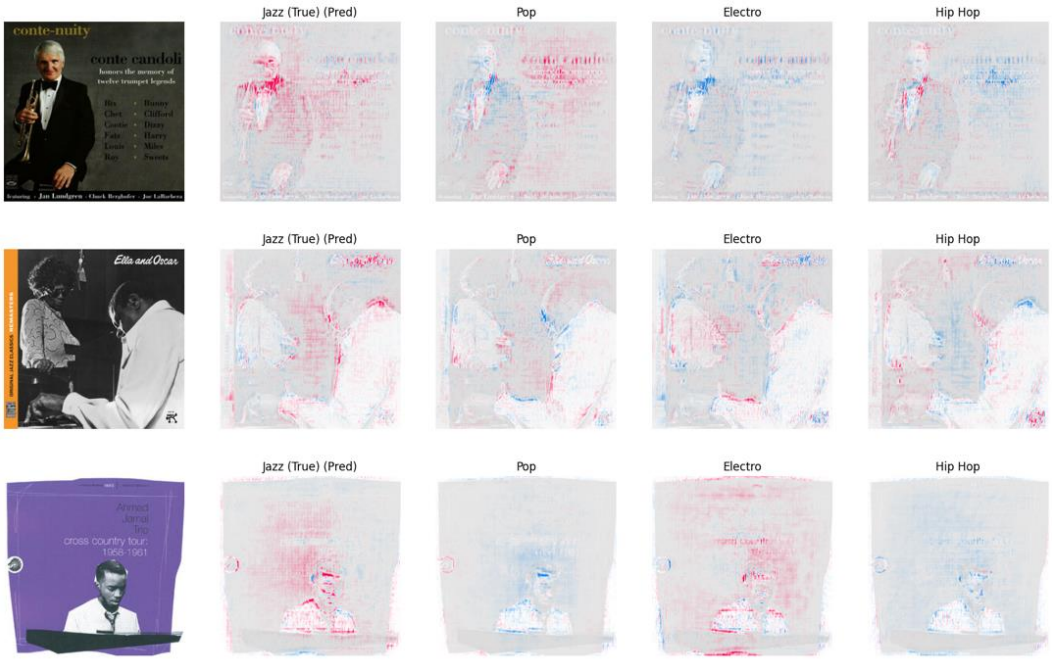


Рис. 1. Интерпретация классификации изображений с белой рубашкой

Нейронной сети также удалось заметить, что в обложках альбомов электронной музыки часто встречаются круглые силуэты в центре изображения. Это можно заметить на рисунке 2.

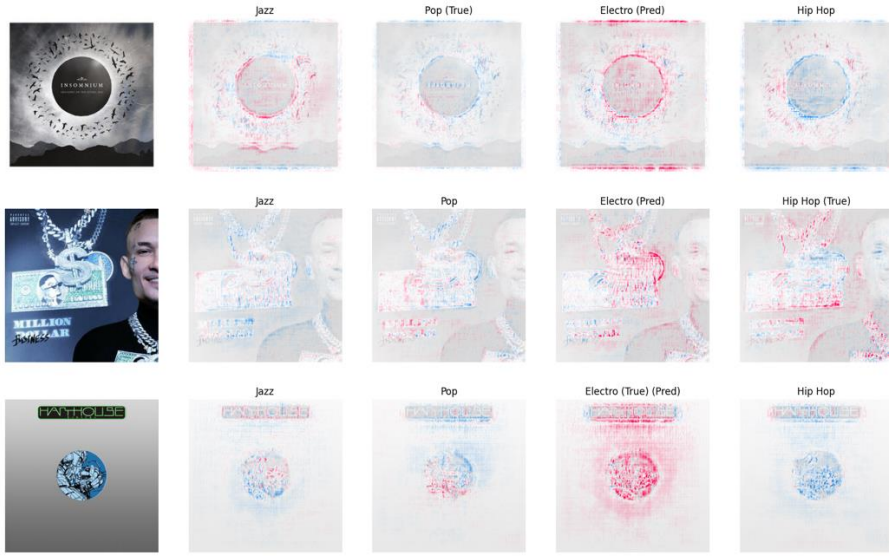


Рис. 2. Интерпретация классификации изображений с круглыми силуэтами

Нейронной сети также удалось уловить следующую закономерность – на обложках хип-хоп альбомов артисты часто помещают наклейку о наличии нецензурного контекста в текстах песен, это смогла заметить нейронная сеть, и поэтому ставила изображению метку «хип-хоп», если видела такую наклейку на изображении. На рисунке 3 можно посмотреть на примеры подобных предсказаний.

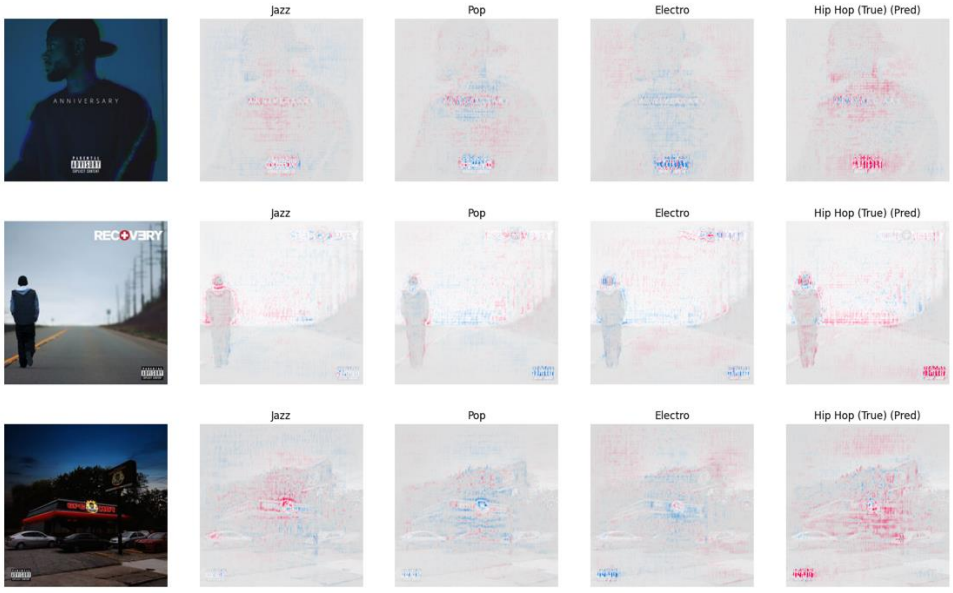


Рис. 3. Интерпретация изображений с наклейкой о нецензурном контенте

Данные наблюдения позволяют сделать вывод о том, что нейронной сети в процессе обучения удается заметить те закономерности в обложках альбомов одного жанра, которые замечает обычный человек.

В дальнейшем планируется увеличить объем исходного набора данных, обучить нейронные сети с разными архитектурами, протестировать интерпретацию работы

нейронной сети с другими алгоритмами, а также провести опрос среди студентов и сравнить, кто предсказывает жанр музыкального альбома по обложке лучше – нейронная сеть или человек.

Литература

1. Deezer for developers [Электронный ресурс]. Режим доступа: https://developers.deezer.com/api (дата обращения: 13.02.2023).
2. Krizhevsky A., Sutskever I., Hinton G. E. Imagenet classification with deep convolutional neural networks //Communications of the ACM. 2017. Т. 60. №. 6. С. 84-90.
3. SHAP [Электронный ресурс]. Режим доступа: https://shap. readthedocs.io/en/ latest/ index.html (дата обращения: 13.03.2023).

УДК 004.67

ПРИМЕНЕНИЕ ТЕХНОЛОГИЙ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА В ЗАДАЧАХ МОНИТОРИНГА И АГРЕГИРОВАНИЯ СОБЫТИЙ ИЗ ОТКРЫТЫХ ИСТОЧНИКОВ С ПРИВЯЗКОЙ К ГЕОГРАФИЧЕСКИМ КООРДИНАТАМ

Бодунов Г.А.[1] (студент), **Лобань И.В.**[1] (студент), **Шишкарев З.А.**[1] (студент), **Дудкин А.С.**[1]
Научный руководитель – к.т.н. Дудкин А.С.[1]
1 – Военно-космическая академия им. А.Ф. Можайского
e-mail: vka@mil.ru

Аннотация
Разработанный программный комплекс включает программные модули по сбору и обработке данных из средств массовой информации в режиме реального времени, с помощью алгоритмов систематизации информации, размещенной на различных новостных ресурсах, посредством технологии асинхронного парсинга веб-страниц и мессенджеров.
Ключевые слова
Искусственный интеллект, обработка естественного языка, парсинг веб-сайтов, открытые источники, веб-приложение, анализ данных.

Открытые источники данных могут предоставить доступ к информации достаточно широкого спектра, однако из-за большого объема данных процесс анализа может занять большое количество времени.

Проблема: Высокое время сбора, обработки и анализа данных из открытых источников.

Цель проекта: Повышение эффективности процессов сбора, обработки и анализа данных с открытых источников.

Задачи проекта для достижения цели:

1) разработать эффективные подходы к сбору информации с открытых источников;

2) реализовать передовые методы обработки и агрегации данных;

2. Разработка алгоритмов сбора данных

2.1. Парсинг и его компоненты.

Для автоматизированного сбора информации из открытых источников применяется технология парсинга.

Парсеры зависят от структуры веб-сайта, и из этого вытекает проблема универсальности применения данной технологии. Для решения проблемы универсальности было разработано 2 подхода к технологии парсинга:

1) целевой парсер;
2) универсальный парсер (метапаук).

2.1.1. Целевой парсер.

Целевой парсер предназначен для сбора информации с определенной структурой, и как следствие, применение данного подхода требует разработки отдельного парсера для каждого веб-сайта, с которого будет производиться сбор информации.

Для разработки целевого парсера была изучена структура веб-сайта, расположение необходимой информативной нагрузки – новостей в тегах веб-страницы.

2.1.2. Универсальный парсер (метапаук).

Метапарсинг предназначен для автоматического сбора и анализа информации о структуре веб-страниц с использованием алгоритмов машинного обучения.

Метапарсинг используется для получения и анализа метаданных, таких как заголовки страниц, описания, ключевые слова, авторы и т.д., а также для получения данных из тегов и разметки HTML, CSS и других языков разметки. В метапарсинге

используются технологии NLP (естественной обработки языка) для обработки содержания веб-страницы.

Реализованный в работе метапаук имеет следующий алгоритм работы.

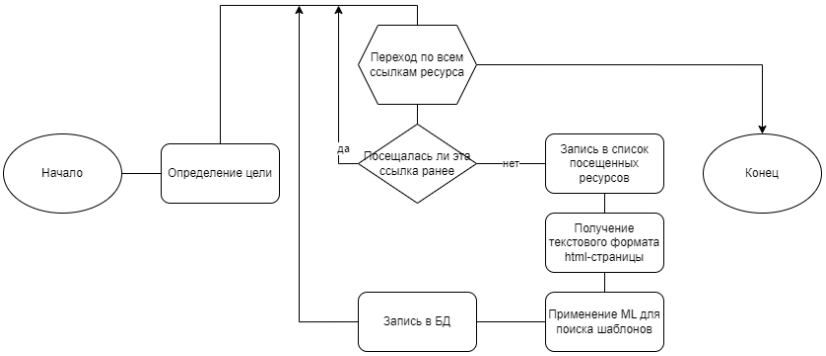


Рис. 1. Алгоритм работы метапаука

3. Разработка алгоритмов обработки данных.

После сбора данных с открытых источников следует этап обработки, полученных данных, в результате которых будут выполнены следующие преобразования:

1) векторизация и кластеризация новостей по категориям;

2) извлечение именованных сущностей;

3) определение географических координат.

3.1. Векторизация и кластеризация новостей по категориям.

В настоящее время объемы текстовой информации, генерируемой и доступной для анализа, растут с каждым днем. Для того чтобы эффективно обрабатывать и анализировать такие объемы данных, необходимы соответствующие методы. Одним из них является комбинация векторизации и кластеризации текста [5]. Использованные ниже методы могут значительно упростить анализ больших объемов текстовой информации и помочь выделить важные тренды и закономерности.

Перейдем к первому этапу.

Перед тем как векторизировать текст необходимо провести его очистку. Существуют различные подходы к предобработке текстовых данных, которые могут повлиять на качество кластеризации. Например, часто используется удаление стоп-слов (наиболее часто встречающихся слов, которые не несут смысловой нагрузки), лемматизация (приведение слов к их базовой форме) и токенизация (разбиение текста на отдельные слова или токены). После проведения преобразований *можно переходить к векторизации текста*.

Векторизация текстовых данных – это процесс преобразования текстовых документов в числовые векторы, которые могут быть использованы для анализа и классификации текста [1].

Существует несколько подходов к векторизации текстовых данных:

1) метод Continuous Bag of Words;

2) методы, основанные на модели Word2Vec (Skip-gram).

Из проведенных экспериментов было установлено, что модель Skip-gram реализации Word2vec обладает наилучшими метриками качества по сравнению с другими.

Перейдем ко второму этапу.

Кластеризация текстовых данных – это метод группировки большого количества текстовых документов по сходству их содержания. Она используется для упрощения и анализа больших объемов текстовой информации, которые могут быть сложны для понимания и интерпретации вручную. Кластеризация текстовых данных является одной из ключевых задач обработки естественного языка и широко применяется в различных

областях, таких как анализ социальных медиа, информационный поиск, маркетинговые исследования и многое другое [1, 2-4].

Существуют различные методы кластеризации текстовых данных, каждый из которых имеет свои преимущества и ограничения. Некоторые из наиболее распространенных методов включают в себя [5]:

1. K-means и Minibatch K-means.
2. DBSCAN.
3. Иерархическая кластеризация.
4. Минимальное покрывающее дерево.

Сравнительный анализ связки двух алгоритмов векторизации с алгоритмами кластеризации представлен ниже на диаграммах (рис. 2, 3).

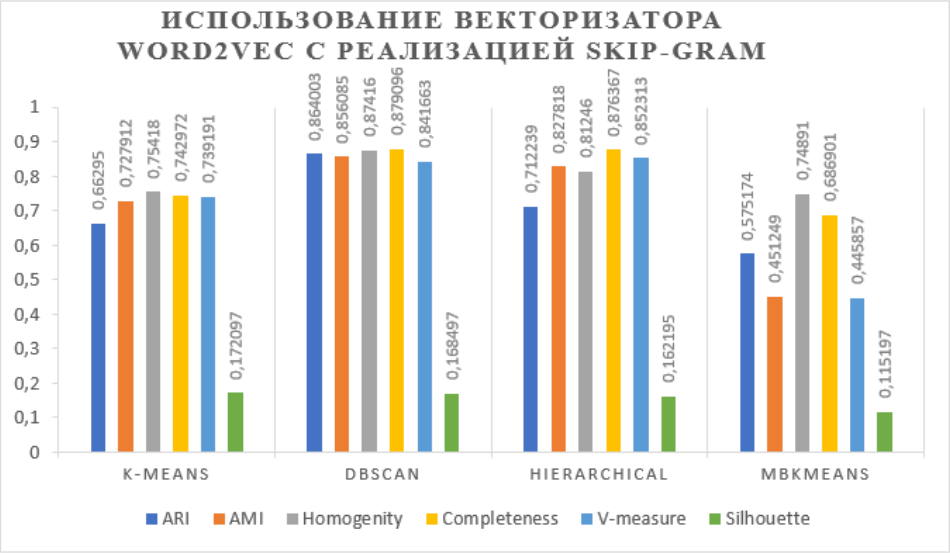


Рис. 2. Диаграмма 1

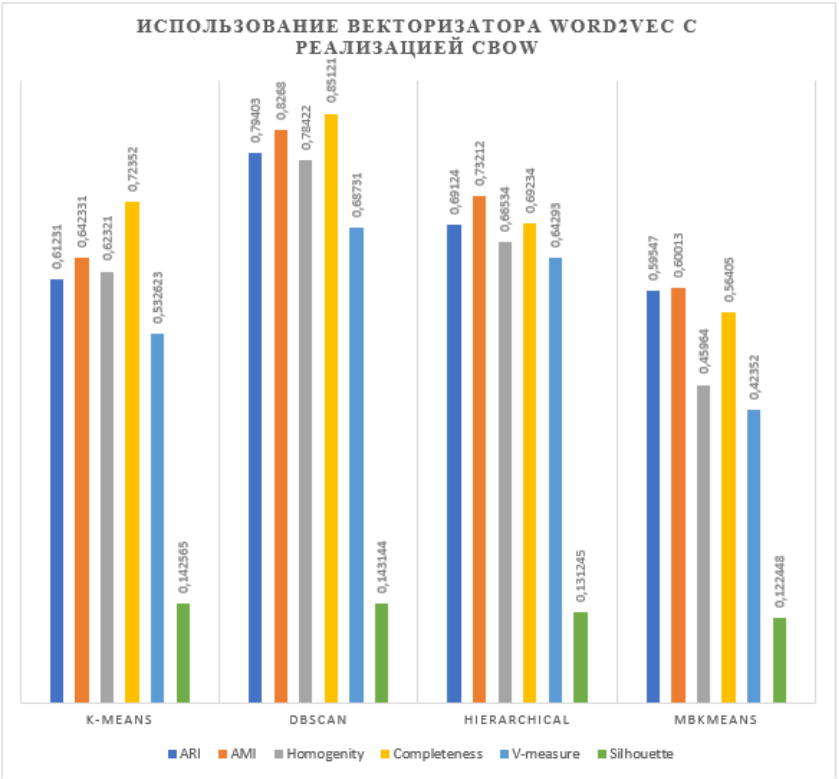


Рис. 3. Диаграмма 2

3.2. Извлечение именованных сущностей.

Именованные сущности могут использоваться для структурирования данных для записи в БД, для нахождения определенных абзацев текста с сущностями определенного вида. Полученные данные сильно облегчают работу семантических алгоритмов NLP, так в предложении: «Вооруженный протестующий приехал на своем автомобиле, после чего сотрудники полиции задержали его в установленном порядке» «протестующий» является сущностью личности, и алгоритм семантического анализа поймет, что сотрудники полиции задержали протестующего, а не автомобиль [6-9].

В работе проблема NER была решена с помощью задачи классификации на уровне токенов, это означает, что каждый токен может относиться к одному из классов. Стандартом решения NER с помощью методов машинного обучения является BIOES-схема. Суть схемы состоит в том, что к метке сущности необходимо добавить некоторый префикс, который обозначает позицию класса в тексте.

Опишем данную схему:

B – beginning. Первый токен в тексте, который состоит из больше, чем 1 слова.

I – inner. Токен, который находится внутри текста, состоящего из больше, чем 1 слова.

O – out. Токен не относится к какой-либо сущности.

E – ending. Последний токен в тексте, который состоит из больше, чем 1 слова.

S – single. Сущность состоит из одного слова (рис. 4).

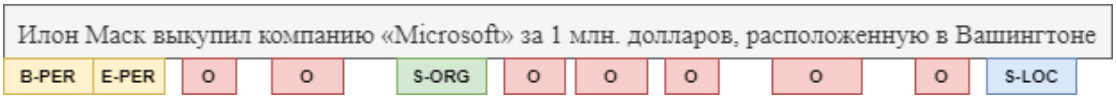


Рис. 4. Текст в представлении схемы BIOES

Таким образом, извлечение сущностей является ключевой задачей в NLP, реализация которого помогает оптимизировать и ускорить работу алгоритмов семантического анализа текста, построения связей «организация-личность», профилирования организаций, личностей и другой информации в тексте, а также определения географических координат сущности локации для дальнейшего отображения на электронной карте.

3.3. Определение географических координат

В третью группу алгоритмов обработки данных, является алгоритм определения географических координат. Алгоритм определения географических координат реализован с применением библиотеки *GeoPy,* которая с помощью встроенных геокодеров автоматически выполняет поиск местоположения после чего на выходе возвращает географические координаты локации (рис. 5).

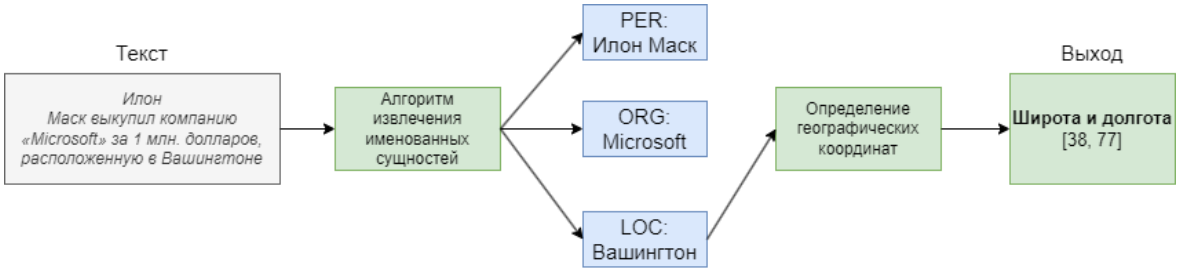


Рис. 5. Принцип работы алгоритма определения географических координат

Поиск географических координат осуществляется в автоматическом режиме после работы алгоритма извлечения сущностей, после чего на вход алгоритму подается сущность с именем *LOC.*

4. Программная реализация работы проекта.

В результате работы было разработано веб-приложение, серверная часть которого написана на языке *Python* с использованием фреймворка *Vue.js* для создания клиентской части приложения (рис. 6).

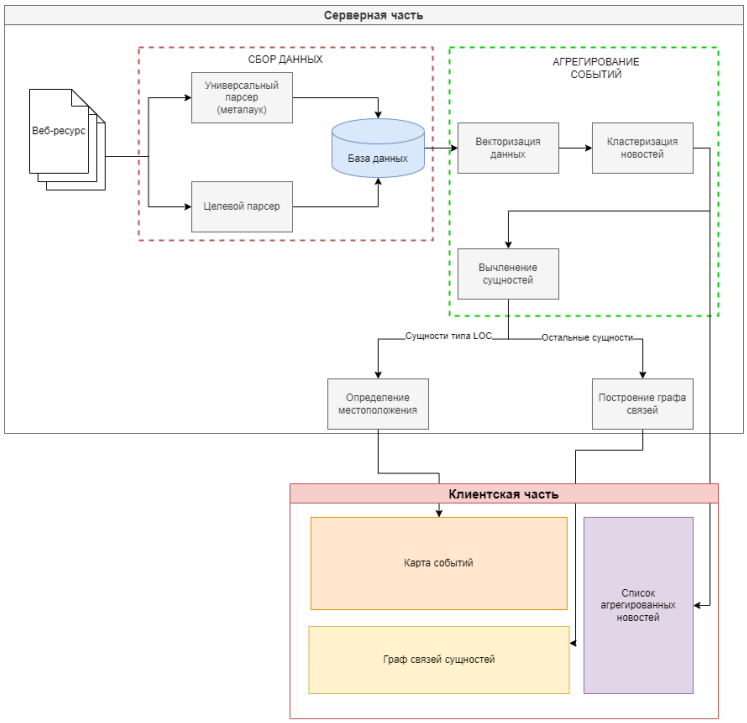


Рис. 6. Архитектура веб-приложения

5. Заключение.

Результатом проделанной работы по созданию сервиса стало решение ранее озвученных задач. Автоматический сбор и обработка данных поможет улучшить планирование и управление ресурсами в кризисных ситуациях. Одной из главных перспектив развития данного решения является улучшение качества аналитических данных за счет использования более сложных алгоритмов машинного обучения и нейросетей. Также стоит отметить, что данный сервис находится на стадии тестирования и существует необходимость в его доработке. В целом, применение искусственного интеллекта в подобных задачах является перспективным направлением в информационных технологиях.

Литература

1. Бирюков Д.Н., Бодунов Г.А., Дудкин А.С., Тельбух В.В. Подход к определению тематической близости текстов. // Труды ВКА имени А.Ф. Можайского. 2022. № 684. С. 84 – 93.
2. Федеральный закон «Об информации, информационных технологиях и о защите информации» от 27.07.2006 N 149-ФЗ [Электронный ресурс]. Режим доступа: https://www. consultant.ru/ document/ cons_doc_LAW_61798/ (дата обращения: 03.03.2023).
3. Пархоменко П.А., Григорьев А.А., Астраханцев Н.А. Обзор и экспериментальное сравнение методов кластеризации текстов // Труды ИСП РАН. 2017. №2. [Электронный ресурс]. Режим доступа: https://cyberleninka.ru/article/n/obzor-i-eksperimentalnoe- sravnenie- metodov- klasterizatsii- tekstov (дата обращения: 03.03.2023).

4. Кравченко Ю.А., Мансур Али М., Мохаммад Ж.Х. Векторизация текста с использованием методов интеллектуального анализа данных // Известия ЮФУ. Технические науки. 2021. №2 (219) [Электронный ресурс]. Режим доступа: https:// cyberleninka.ru/article/n/vektorizatsiya-teksta-s-ispolzovaniem-metodovintellektualnogo-analiza-dannyh (дата обращения: 05.03.2023).
5. Ботов Д.С., Кленин Ю.Д., Николаев И.Е. Извлечение информации с использованием нейросетевых моделей языка на примере анализа вакансий в системах онлайн-рекрутмента // Вестник ЮГУ. 2018. №3 (50) [Электронный ресурс]. Режим доступа: https: // cyberleninka.ru/ article/n/izvlechenie-informatsii-s- ispolzovaniem-neyrosetevyh-modeley- yazyka-na- primere-analiza- vakansiy-v- sistemah-onlayn-rekrutmenta (дата обращения: 03.03.2023).
6. Чанышев О.Г. Критерий близости документов и кластеризация // МСиМ. 2001. №2 (8) [Электронный ресурс]. Режим доступа: https://cyberleninka.ru/article/n/kriteriy-blizosti-dokumentov-i-klasterizatsiya (дата обращения: 08.03.2023).
7. Меньшиков Я.С. Преимущества автоматического сбора данных в сети интернет над ручным сбором данных // Universum: технические науки. 2022. №10-1 (103). [Электронный ресурс]. Режим доступа: https:// cyberleninka.ru/ article/n/ preimuschestva- avtomaticheskogo- sbora-dannyh-v- seti-internet- nad-ruchnym-sborom-dannyh (дата обращения: 10.03.2023).
8. Цхошвили Д.З., Иванова Н.А. Извлечение данных с веб-ресурсов, содержащих разнородную информацию // Ученые записки Брянского государственного университета. 2017. №4 (8). [Электронный ресурс]. Режим доступа: https:// cyberleninka.ru/ article/n/ izvlechenie-dannyh-s-veb- resursov- soderzhaschih-raznorodnuyu-informatsiyu (дата обращения: 12.03.2023).
9. Менщиков А.А., Комарова А.В., Гатчин Ю.А. Изучение поведения средств автоматизированного сбора информации с веб-ресурсов // Вопросы кибербезопасности. 2017. №3 (21). [Электронный ресурс]. Режим доступа: https:// cyberleninka.ru/ article/n/izuchenie- povedeniya- sredstv- avtomatizirovannogo- sbora-informatsii-s-veb-resursov (дата обращения: 12.03.2023).

УДК 004.41

ПРОЕКТИРОВАНИЕ МОБИЛЬНОГО ПРИЛОЖЕНИЯ ДЛЯ КОНТРОЛЯ И ПЛАНИРОВАНИЯ ТРЕНИРОВОЧНОГО ПРОЦЕССА

Макаров К.Д.[1] (магистрант)
Научный руководитель – старший преподаватель Береснев А.Д.[1]
1 – Университет ИТМО

Аннотация
В работе представлен этап проектирования мобильного приложения для контроля и планирования тренировочного процесса при занятиях спортом, целью которого являлась подготовка к разработке приложения. Для этого были проанализированы приложения-аналоги, определены основные функциональные требования, составлена диаграмма вариантов использования, даталогическая модель данных, для реализации интеллектуальных функций выбран ML, а также был проведен поиск и выбор средств, необходимых для его реализации.
Ключевые слова
Мобильная разработка, проектирование приложения, ОС Android, тренировочный процесс, план питания.

За последние сто лет количество людей с избыточным весом значительно возросло, это в первую очередь связано с тем, что образ жизни современного человека сильно изменился – люди в наше время двигаются намного меньше, чем их предки, так как появляется все больше удаленных работ, кто-то переходит на доставку продуктов или готовой еды. Вызовы, с которыми сталкивается современное общество тоже имеют свой вклад, так согласно исследованию платформы Ipsos Global Advisor — около ⅓ населения мира набрало вес за время пандемии [1]. А малоподвижный образ жизни и избыточный вес тела немаловажные факторы при развитии большинства хронических заболеваний, к тому же COVID у таких людей по статистике протекает в более тяжелой форме. Вместе с этим, как это не странно, интерес к занятиям спортом тоже растет. Так, согласно исследованию ЕМИСС, доля граждан Российской Федерации, систематически занимающихся спортом, выросла с 45.4 % в 2020-м году до 51.5 % в 2023-м [2]. Среди проблем, возникающих у человека, принимающего решение начать заниматься спортом, существует проблема поиска и структурирования информации, а также проблема того, как оперативно получать доступ к этой информации прямо во время тренировок. И если для решений первой проблемы каждый, основываясь на своих финансовых возможностях, а также на жизненном опыте, выбирает прибегнуть к услугам тренера или воспользоваться открытыми источниками информации, то для решения второй на помощь приходят разнообразные мобильные приложения для тренировок. Но, к сожалению, на сегодняшний момент технологии, которые призваны облегчить тренировочный процесс еще далеки от идеала и заточены под разные нужды: одни составляют тренировочный план и помогают ему следовать, но не позволяют его изменять, другие же достаточно вариативны, но не рассчитаны для новичков, есть приложения и для тренеров, но они не позволяют клиентам заниматься самостоятельно или, например, сменить тренера без потери доступа к аккаунту. Это приводит к тому, что одному человеку приходится параллельно использовать несколько приложений и дублировать в них большую часть информации, что чаще не облегчает, а только усложняет его жизнь. Поэтому было принято решение спроектировать мобильное приложение, с помощью которого пользователи смогут самостоятельно выстраивать свой тренировочный план, план питания, а также делиться своим прогрессом с тренером. А тренеры смогут формировать план тренировок и питания своих клиентов, а также отслеживать их прогресс.

В ходе анализа рынка были рассмотрены сервисы-аналоги были выбраны 3 популярные Android приложения для тренировок разной направленности, имеющие наибольший функционал среди своих прямых конкурентов:

1. GymUp – позволяет самостоятельно выстраивать тренировочный процесс.
2. Тренировки для дома – содержат базу уже готовых тренировочных программ.
3. Trainerize – приложение для фитнес-тренеров, в котором основной функционал у пользователя-тренера.

При помощи сравнительного анализа, результаты которого содержатся в таблице, были выделены их преимущества и недостатки.

Таблица

Сравнение сервисов-аналогов

	GymUp	Тренировки для дома	Trainerize
Регистрация	ДА, по желанию	ДА, по желанию	ДА, обязательно (приглашение от тренера)
Изменение программы	ДА	ДА	Только тренер
Стоимость	Расширенная версия(199р/месяц)	Убрать рекламу(219р)	Для тренера 2-350
Реклама	НЕТ	ДА	НЕТ
Техника выполнения	ДА	ДА	ДА
Самостоятельное составление программы	ДА	НЕТ	Только тренер
Функционал тренера	НЕТ	НЕТ	ДА
Возможность поделиться статистикой	Общая статистика (всего подходов, общий вес и т.д.)	НЕТ	Только тренер может увидеть статистику клиента
Возможность поделиться программой	НЕТ	НЕТ	
Функционал создания/отслеживания плана питания	НЕТ	НЕТ	Создает тренер

На основе проведенного анализа сервисов-аналогов, были выделены основные функциональные требования к системе:

- авторизация и регистрация;
- создание тренировочных программ;
- отображение и отслеживание тренировочных программ;
- создание плана питания;
- отображение и отслеживание плана питания;
- отображение программ в общем доступе;
- передача пользователю информации о программе тренировок и питании;
- отображение панели управления личным кабинетом;
- отображение тренерской панели;
- отображение панели модератора;
- отображение панели администратора.

Было выделено 4 основных актора, а прецеденты разделены на 6 семантический групп и после составлена диаграмма вариантов использования, показанная на рисунке 1.

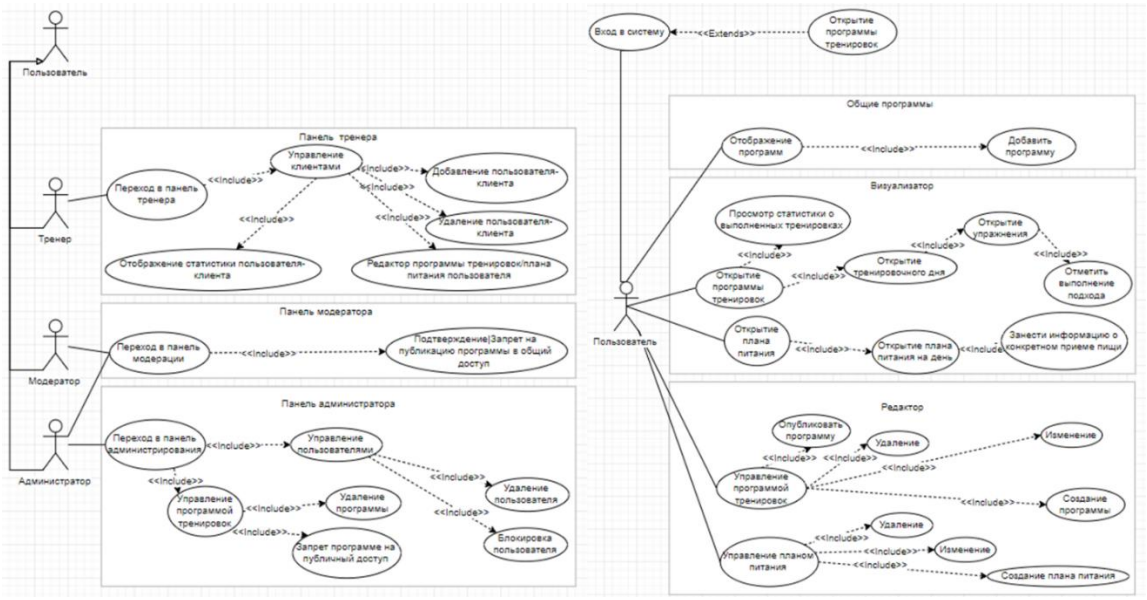


Рис. 1. Диаграмма вариантов использования

При помощи языка UML была создана логическая модель данных, которая отображена на рисунке 2.

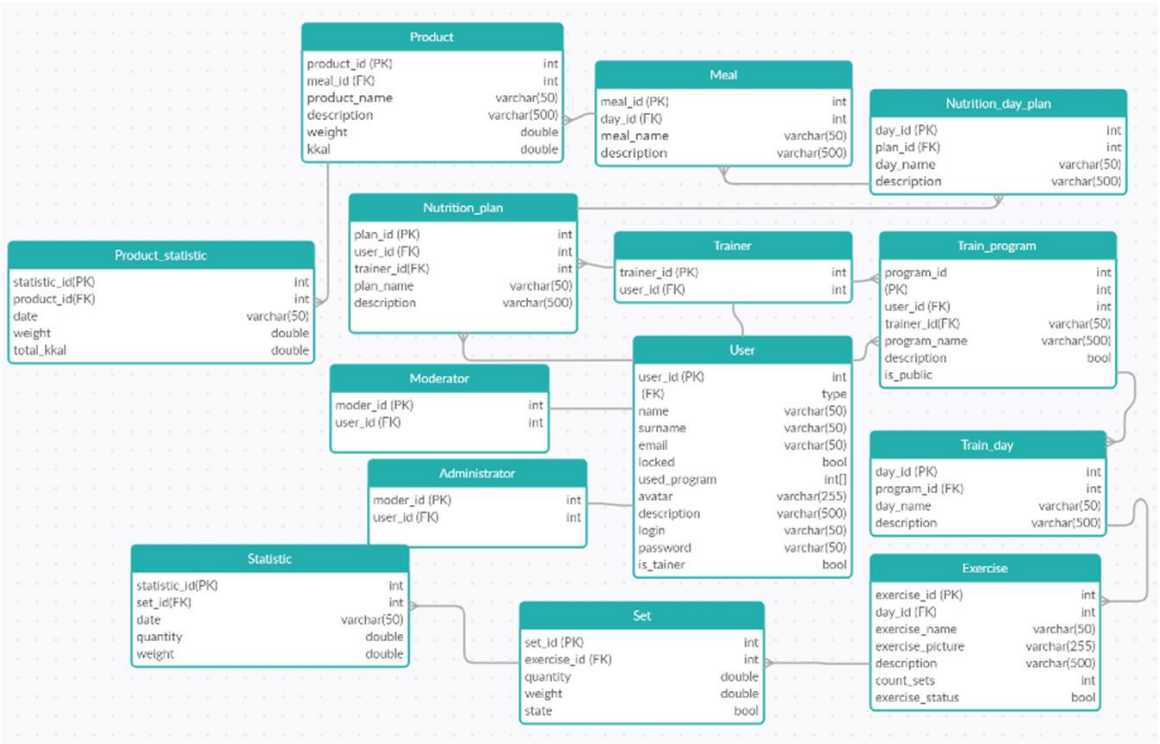


Рис. 2. Даталогическая модель базы данных

Для реализации интеллектуальных функций будущего приложения выбран ML. Для обучения моделей необходим специфичный датасет, который будет формироваться из таблиц Statistic и Product_Statistic, в которые заносятся данные о выполненных упражнениях и приготовленных приемах пищи, которые в будущем могут быть использованы для создания рекомендаций по изменению режимов питания и тренировок.

В качестве основных средств разработки были выбраны:

1. Android Studio – интегрированная среда разработки под платформу Android.
2. Kotlin – язык программирования рекомендуемый Google для разработки Android приложений.
3. SQLite – система управления базами данных (СУБД), используемая для организации локального хранения данных.
4. Room – библиотека, часть архитектурных компонентов Android, предназначена для работы с SQLite.
5. PostgreSQL – СУБД для организации удаленного хранения данных на сервере, которая является одной из лучших на рынке, при этом ее использование бесплатно [3].
6. Retrofit 2 – библиотека сетевого взаимодействия, необходимая для получения и отправки удаленных данных.
7. Jetpack Compose – инструментарий для создания пользовательского интерфейса Android приложения с использованием декларативного подхода, согласно сравнению такого подхода со стандартным, использующим XML, выяснилось, что после рендера, экраны, созданные при помощи Jetpack Compose открываются быстрее [4].

При выборе архитектуры будущего приложения рассматривались Model-View-Presenter(MVP) и Model-View-ViewModel(MVVM) как доминирующие отраслевые подходы к архитетктуре. Была выбрана архитектура MVVM, так как при сравнении этих архитектур, произведенном для доклада в рамках ISSAT 2020, выяснилось, что ее использование повышает быстродействие приложения и в меньшей степени расходует ресурсы процессора [5].

В ходе данной работы был поведен анализ предметной области и обзор аналогов, сформулированы основные функциональные требований к системе, создана диаграмма вариантов использования, составлена даталогическая модель данных, создан принцип формирования датасетов, на основе которых в будущем при помощи ML будут созданы рекомендации по питанию и тренировкам, а также выбраны технологии и архитектура для реализации системы. На этом стадия проектирования может быть законченная и можно приступить к первой фазе разработки мобильного приложения.

Литература

1. Ipsos Global Advisor The implications of COVID—19 on our diet & health – 20.01.2021 [Электронный ресурс]. Режим доступа: https://www.ipsos.com/en/covid—diet—and—health (дата обращения 15.02.2023).
2. Доля граждан, систематически занимающихся физической культурой и спортом. (Р5) [Электронный ресурс]. Режим доступа: https://www.fedstat.ru/indicator/61635 (дата обращения15.02.2023).
3. Kleweka R, Truskowski W., Skublewska-Paszkowska M. Comparison of MySQL,MSSQL, PostgreSQL, Oracle databases performance, including virtualization // Journal ofComputer Sciences Institute. 2020. P. 279-284.
4. Клоков С.А. Сравнительный анализ времени отображения экранов с использованием стандартного подхода в Android и библиотеки Jetpack Compose // Молодой ученый. 2022. № 4 (399). С. 12-16.
5. Wisnuadhi B., Munawar G., Wahyu U. Performance Comparison of Native Android Application on MVP and MVVM // Advances in Engineering Research, volume 198 International Seminar of Science and Applied Technology (ISSAT 2020). 2022.

УДК 004.896

СТАТИЧЕСКИЙ АНАЛИЗ ФРЕЙМВОРКОВ МАШИННОГО ОБУЧЕНИЯ

Менисов А.Б.[1] (докторант), **Захаров О.О.**[1] (студент)
Научный руководитель – к.т.н. Менисов А.Б.[1]
1 – Военно-космическая академия им. А.Ф. Можайского
e-mail: vka@mil.ru

Работа выполнена при финансовой поддержке гранта Президента Российской Федерации для государственной поддержки молодых российских ученых - кандидатов наук, МК-2485.2022.4.

Аннотация
Фреймворки машинного обучения являются основой для разработки, а их ошибки и сбои могут привести к катастрофическому снижению качества систем искусственного интеллекта. Такой негативный эффект демонстрирует необходимость и важность обеспечения качества фреймворков машинного обучения. Понимание характеристик этих ошибок является базовым шагом для обеспечения защищенности систем искусственного интеллекта. В работе представлены результаты статического анализа популярных фреймворков (TensorFlow, PyTorch, MXNet, CNTK, Dlib, Keras, и Sklearn).
Ключевые слова
Искусственный интеллект, фреймворки машинного обучения, информационная безопасность, анализ дефектов и уязвимостей, статический анализ.

В последние годы системы искусственного интеллекта (СИИ) стали одним из самых популярных типов программного обеспечения и широко используются во многих видах деятельности. Однако, как и все программное обеспечение, СИИ также содержат дефекты, которые могут привести к огромным экономическим потерям или даже угрожать жизни людей.

Архитектура СИИ включает три уровня: прикладной уровень (модели машинного обучения), программный уровень (программы, используемые для обучения моделей машинного обучения) и уровень фреймворков. Ошибки и неисправности на любом уровне могут повлиять на качество целой СИИ. Следовательно, необходимо обеспечить качество на всех этих уровнях. Большинство исследований сосредоточены на тестировании моделей машинного обучения прикладного уровня [1-3], а также разработки методов обнаружения и диагностики ошибок компонентов СИИ [4-6]. Однако, дефектам и уязвимостям фреймворков уделяется меньше внимания. На самом деле, фреймворки являются основой для построения всех СИИ и моделей машинного обучения, поэтому их ошибки могут иметь гораздо более широкий эффект, чем ошибки в конкретной программе или модели машинного обучения.

В последнее время качество СИИ стало актуализировано, и были предложены некоторые методы тестирования [7, 8]. Несмотря на то, что некоторые методы доказали свою эффективность в обнаружении некоторых новых ошибок, они, как правило, рассматривают тестируемую СИИ как черный ящик и не имеют полного понимания характеристик ошибок. Такой недостаток существенно ограничивает их производительность и мешает разработке более эффективных методов обнаружения ошибок. Таким образом, всестороннее понимание характеристик ошибок фреймворков машинного обучения является одной из основных задач в области обеспечения качества СИИ.

Анализ научно-технического задела показал, что в мире было проведено несколько исследований по изучению характеристик ошибок СИИ [9, 10], но они нацелены на ошибки прикладной программы, а не на ошибки фреймворка машинного обучения. Из-за существенных различий между прикладными программами и

фреймворками их ошибки различны. В частности, СИИ должна вызывать API, предоставляемые инфраструктурой машинного обучения, для эксплуатации желаемой архитектуры нейронной сети.

Фреймворки машинного обучения (МО) являются основой для реализации СИИ и разработки моделей МО. Помимо реализации различных функций МО, фреймворки также обладают двумя типичными характеристиками: обеспечивают связь между моделями МО и аппаратной инфраструктурой, и реализуют некоторые технологии для поддержки функций машинного обучения.

Машинное обучение является быстрорастущей технологической областью, и поэтому фреймворки МО часто обновляются, чтобы учесть быстрое развитие алгоритмов МО. В частности, они имеют большое разнообразие, например, включают в себя применение как статических, так и динамических вычислительных графов на различных языках программирования. Фреймворки МО очень сложны, и их можно декомпозировать на 4 уровня: прикладной, вычислительного графа, функциональный, уровень утилит и среды выполнения.

Прикладной уровень включает следующие компоненты:

1) обработки данных, предназначенный для конвертации входных данных под соответствующие требования модели МО, например, изменение размера изображения и токенизация текста;

2) разработки модели МО, направленный на построение структуры и поиск группы оптимальных гиперпараметров (например, функции потерь и тип оптимизатора);

3) интеграции модели МО в существующую информационную инфраструктуру, включающий вспомогательные функций для облегчения процесса машинного обучения, например, визуализацию функционирования модели и создание контрольных точек.

Уровень вычислительного графа, в котором каждый узел представляет операцию (например, операцию свертки), также можно декомпозировать на компоненты:

1) создания вычислительного графа;

2) преобразования графа (например, объединение операций или преобразование в формат ONNX);

3) выполнения графа, включая локальное и распределенное выполнение.

Функциональный уровень содержит все подробные реализации операций вычислительного графа. Операция принимает ноль или более тензоров на вход и создает ноль или более тензоров на выходе. В фреймворках МО реализовано большое количество операций, таких как операции свертки, объединения, пакетной нормализации, манипулирования массивами и др.

Уровень утилит и среды выполнения является самым низким и предназначен для поддержки функциональных возможностей фреймворков в различных средах. Типичным примером является распределение памяти на разных устройствах, а также связи с аппаратной (например, GPU), так и программной средой (например, операционные системы).

СИИ могут быть атакованы путем эксплуатации уязвимостей и дефектов программного обеспечения, а также новых уязвимостей различных компонентов машинного обучения, процессом машинного обучения, когда злоумышленники получают доступ и манипулируют незащищенным программным кодом и данными. Например, фреймворки машинного обучения содержат дефекты и уязвимости, которые могут привести к отказу в обслуживании, раскрытию информации или удаленному выполнению кода. В таблице представлены результаты проведения статического анализа с помощью средства Coverity.

Таблица

Дефекты и уязвимости фреймворков машинного обучения

Фреймворки / Дефекты	CNTK 2.7	Dlib 19.24	Keras 2.10.0	MXNet 1.9.1	PyTorch 1.12.1	Sklearn 1.1.2	Tensorflow 2.10
Всего дефектов и уязвимостей	57	185	54	1580	2859	867	191
Дефекты высокого уровня опасности	12	52	0	465	1468	93	43
Дефекты среднего уровня опасности	41	112	16	918	1150	323	140
Дефекты низкого уровня опасности	4	21	38	197	241	451	8
Активность обнаружения (за последние 5 лет)	0	0	3	2	0	2	323
Топ-25 (2022 года)	1	5	3	5	5	2	3
Наиболее уязвимые компоненты	OpenCV MKLML	libpng, zlib		oneDNN, GoogleTest, CUDA SDK	oneDNN, GoogleTest, Tensorpipe, NCCL, Google Protobuf, onnx, ATen, SLEEF, ideep, Google Benchmark, FBGEMM		NVIDIA NCCL

Статистический анализ показал дефекты, связанны с:

1. Неверной обработкой ошибок (error handling issues).
2. Проблемами хранения целых чисел (integer handling issues).
3. Неинициализированными членами классов (uninitialized members).
4. Нарушением потока управления (control flow issues).
5. Разыменовыванием нулевого указателя (null pointer dereference).
6. Некорректными выражениями (incorrect expression).
7. Проблемами использования API (API usage issues).
8. Различными другими ошибками (medium impact quality).
9. Нарушением доступа к данным из нескольких потоков (concurrent data access violation).
10. Небезопасной обработкой данных (insecure data handling).
11. Проблемами с производительностью (program hangs).

Перечисленные дефекты могут повлечь за собой отказ в обслуживании (Denial of Service), например, разыменовывание нулевого указателя может привести к аварийному завершению работы программы. Проблемы с хранением целых чисел могут повлечь за собой исполнение произвольного кода, привести к неопределенному поведению программы и, как следствие, аварийному завершению работы, либо нестабильной работе фреймворка. Нарушение потока управления влияет на логику работы программы, то есть в ряде случаев не будет выполняться ожидаемый код, поскольку условия его выполнения заведомо невозможные (dead code). Нарушение доступа к данным из разных потоков может привести к неопределенному поведению программы, что неприемлимо в системах, спроектированных с расчетом на высокую отказоустойчивость.

Таким образом, в этой работе представлены результаты статического анализа основных фреймворков машинного обучения. Дефекты и уязвимости фреймворков машинного обучения могут повлечь за собой нарушение всех составляющих информационной безопасности – конфиденциальности, целостности и доступности: повреждения памяти могут повлечь за собой исполнение произвольного кода, несанкционированный доступ к конфиденциальной информации, а также атаки типа DoS (Denial of Service); неинициализированные переменные зачастую содержат некорректные данные – это может повлиять на стабильность работы систем искусственного интеллекта.

Литература

1. Braiek H.B., Khomh F. On testing machine learning programs //Journal of Systems and Software. 2020. Т. 164. С. 110542.
2. Zhang J.M. et al. Machine learning testing: Survey, landscapes and horizons //IEEE Transactions on Software Engineering. 2020. Т. 48. №. 1. С. 1-36.
3. Cheng K.S. et al. Tool Support for Improving Software Quality in Machine Learning Programs //Information. 2023. Т. 14. №. 1. С. 53.
4. Amershi S. et al. Software engineering for machine learning: A case study //2019 IEEE/ACM 41st International Conference on Software Engineering: Software Engineering in Practice (ICSE-SEIP). IEEE. 2019. С. 291-300.
5. Lwakatare L.E. et al. A taxonomy of software engineering challenges for machine learning systems: An empirical investigation //Agile Processes in Software Engineering and Extreme Programming: 20th International Conference, XP 2019, Montréal, QC, Canada, May 21–25, 2019, Proceedings 20. – Springer International Publishing, 2019. С. 227-243.
6. Yang Y. et al. A survey on deep learning for software engineering //ACM Computing Surveys (CSUR). 2022. Т. 54. №. 10s. С. 1-73.
7. Martínez-Fernández S. et al. Software engineering for AI-based systems: a survey //ACM Transactions on Software Engineering and Methodology (TOSEM). 2022. Т. 31. №. 2. С. 1-59.
8. Ahmad K. et al. Requirements engineering for artificial intelligence systems: A systematic mapping study //Information and Software Technology. 2023. С. 107176.
9. Andrada G. On human centered artificial intelligence. 2023.
10. Tapeh A.T.G., Naser M.Z. Artificial intelligence, machine learning, and deep learning in structural engineering: a scientometrics review of trends and best practices //Archives of Computational Methods in Engineering. 2023. Т. 30. №. 1. С. 115-159.

УДК 004.85

АНАЛИЗ ДООБУЧЕННЫХ ЯЗЫКОВЫХ МОДЕЛЕЙ ДЛЯ ЗАДАЧИ ГЕНЕРАЦИИ ДИСТРАКТОРОВ

Меньщиков М.А.[1] (студент)
Научный руководитель – к.ф.-м.н., доцент Корзун Д.Ж.[1]
1 – Петрозаводский государственный университет
e-mail: menshikov.mikhail.2001@gmail.com

Работа выполнена в рамках темы проекта НИОКР №79049 «Разработка программного ассистента для аттестации профессиональных знаний на основе методов искусственного интеллекта», который реализуется ООО «Универтус» при финансовой поддержке Фонда содействия инновациям.

Аннотация
В работе рассмотрено применение предобученных языковых моделей на основе Transformer-архитектуры для одной из подзадач по автоматизации процесса создания тестовых заданий для проверки знаний аттестуемых: автоматической генерации грамматически верных и подходящих под контекст вопроса неправильных вариантов ответа (дистракторов). Выдвигается гипотеза, что привлечение дополнительных тренировочных датасетов за счёт использования мультиязычных моделей позволит получить лучшее качество генерации по сравнению со стандартным дообучением для одного естественного языка.

Ключевые слова
Генерация дистракторов, предобученные языковые модели, Transformer-архитектура, мультиязычные модели, T5, GPT3.

Качество тестовых заданий с множественным выбором (Multiple Choice Questions) сильно зависит от созданных к ним неправильных ответов: дистракторов. Если эти отвлекающие варианты не могут запутать тестируемых участников, то правильный ответ может быть найден из логических соображений, никак не основанных на знании предметной области, понимание которой проверяется в данном вопросе. В результате качество всего задания сильно ухудшится, а сам тест теряет возможность объективной оценки знаний аттестуемых. Тем не менее, подбор подходящих дистракторов является сложной задачей. Вместо того, чтобы быть очевидным неправильным вариантом, дистрактор должен обладать правдоподобностью, которая сбивает с толку аттестуемых, не усвоивших проверяемые предмет. Хороший дистрактор должен быть грамматически правильным с точки зрения вопроса и семантически соответствовать его контексту. Составители заданий должны добиться правдоподобия дистрактора, при этом не задев присущую ему некорректность. Таким образом, ручная подготовка дистракторов требует много времени и ресурсов. На сегодняшний день актуальной задачей является разработка методов автоматической генерация "хороших" отвлекающих вариантов ответа, которые смогут облегчить работу составителей вопросов и уменьшить количество необходимых знаний/опыта для их создания по конкретной области [1].

За последние годы языковые модели на основе Transformer-архитектуры прошли путь от экспериментальных исследований до внедрения в большинство задач по обработки естественного языка (NLP). Их популярность началось с ULM-FiT- и BERT-моделей, которые показали потенциал предварительного обучения (Pre-training) большой нейронной сети с использованием подхода обучения без учителя (Unsupervised Learning). После предварительного обучения эти нейронные сети можно дообучить (Fine-Tune) и применить для решения разных конкретных задач. Это привело к существенным улучшениям в решении задач классификации эмоционального состояния по тексту, генерации ответов на вопросы, извлечения признаков из

неструктурированных данных и многих других. С появлением GPT2-архитектуры, был сделан большой скачок по качеству генерации текстовых данных. Было показано, что производительность языковой модели продолжает улучшаться с увеличением её размера: начиная от 117 миллионов параметров для самой маленькой модели GPT-2 до 175 миллиардов параметров для самой большой из моделей GPT-3.

В рамках данного исследования внимание будет сконцентрировано на оценке применимости предобученных языковых моделей на основе Transformer-архитектуры для генерации грамматически верных и подходящих под контекст вопроса неправильных вариантов ответа (дистракторов) [2].

Для получения результатов и сокращения затрат на формирование сравнительной статистики вся работа была разбита на следующие этапы:

1. Выбор предобученных моделей.
2. Формирование тренировочных корпусов.
3. Выбор оптимального формата входной последовательности (Prompts) для исследуемых моделей.
4. Выбор гиперпараметров моделей и их дообучение.
5. Выбор оптимальной стратегии генерации текста (Text Generation Strategies) для исследуемых моделей.
6. Ручная оценка генерируемых дистракторов.

Для проведения исследований был использован следующий инструментарий и источники данных:

1. Предобученные модели были получены из открытых репозиториев на сайте HuggingFace [3] в разделе "Models".
2. Дополнительные тренировочные корпуса текстов были получены из открытых репозиториев на сайте HuggingFace в разделе "Datasets".
3. Реализация архитектуры используемых нейронных сетей и их процесса дообучения проводилось с использованием программных библиотек Transformers и PyTorch на языке Python.
4. Реализация алгоритмов перебора параметров проводилась с использованием программной библиотеки Optuna на языке Python.
5. Для проведения высоконагруженных вычислений была арендована виртуальная машина с доступом к графическому процессору общим объёмом 80 ГБ на платформе ImmersCloud.

В рамках данной работы выдвигалась гипотеза о приросте значений метрик качества генерации при использовании мультиязычных моделей по сравнению с моделями, которые предназначены для генерации текста на одном языке [4]. В результате поиска в качестве предобученных моделей были выбраны следующие экземпляры:

1) mT5 (by google):
   - репозиторий модели: "google/mt5-base";
   - размер: 2.33 GB;
   - модель была предобучена на корпусе текстов mC4, покрывающем 101 язык, но без задач обучения с учителем (Supervised Learning).
2) ruT5 (by cointegrated):
   - репозиторий модели: "cointegrated/rut5-base-multitask";
   - размер: 977 MB;
   - модель представляет собой урезанную версия mT5-модели (by google) только для русского и английского языков;
   - модель была дообучена на следующих задачах: машинный перевод, заполнение пропусков, перефразирование, генерация заголовков, генерация вопросов.
3) mGPT (by sberbank-ai):

- репозиторий модели: "sberbank-ai/mGPT";
- размер: 3.45 GB;
- модель обучена на 60 языках с использование следующих корпусов: Wikipedia, Colossal Clean Crawled Corpus;
- модель была дообучена на задачах классификации, автоматической разметке и генерации текста.

4) ruGPT (by sberbank-ai):
   - репозиторий модели: "sberbank-ai/rugpt3medium_based_on_gpt2";
   - размер: 1.73 GB;
   - модель была дообучена на задаче генерации текста.

Для проведения экспериментов по дообучению выбранных моделей были использованы следующие текстовые корпуса:

1. ОТиПБ (ru):
   - датасет подготовлен компанией ООО "Универтус";
   - для обучения будет использоваться набор из 3498 записей, каждая из которых включает: отрывок текста, вопрос по данному текстовому отрывку, правильный ответ и три неверных варианта ответа;
   - для валидации моделей в процессе обучения используется набор из 80 записей;
   - датасет создавался на основании документов ГОСТ из области охраны труда и техники безопасности.
2. RACE (eng):
   - открытый репозиторий датасета: "datasets/race";
   - набор данных собран из экзаменов по английскому языку в Китае, которые предназначены для учащихся средних и старших классов;
   - для обучения будет использоваться корпус "RACE-middle Train", состоящий из 25421 записей, каждая из которых содержит: отрывок текста, вопрос, 4 варианта ответа, номер правильного ответа.
3. SCIQ (eng):
   - открытый репозиторий датасета: "datasets/sciq";
   - набор данных содержит 13679 экзаменационных вопросов по физике, химии и биологии, каждый из которых содержит: отрывок текста, текст вопроса, правильный ответ, 3 неверных варианта ответа.

В качестве стратегий генерации текста были выбраны три наиболее популярных метода: BeamSearch, Sampling и ContrastiveSearch. Для каждого из данных методов будет производиться перебор гиперпараметров для выбора оптимальной конфигурации, дающей максимальное качество метрик. В качестве таких метрик, определяющих: насколько эталонный вариант дистрактора совпадает с сгенерированным были выбраны BLEU1, BLEU2 и BertScore. Первые две метрики оценивают корреляцию между двумя текстами на основании частотностей употребляемых в них слов, когда последняя метрика выдаёт оценку на основании косинусной близости их векторных представлений. Для каждого из выбранных методов генерации перебирались значения следующих параметров:

1. Beam Search: max_length, num_beams, early_stopping, no_repeat_ngram_size, diversity_penalty, temperature, repetition_penalty, num_beam_groups.
2. Sampling: max_length, early_stopping, no_repeat_ngram_size, top_k, top_p, repetition_penalty, temperature, do_sample.
3. Contrastive Search: penalty_alpha, top_k, max_length, early_stopping, repetition_penalty, temperature, no_repeat_ngram_size.

Ручная оценка качества генерируемых дистракторов с помощью выбранных моделей производится на основе следующих метрик:

- количество принятых дистракторов;

- количество дистракторов, похожих на ответ для данного вопроса;
- количество дистракторов, не подходящих под контекст вопроса;
- количество дистракторов с другими ошибками.

В результате перебора специальных токенов (Prompts), которые будут использоваться при формировании входной последовательности для дообучения моделей, оказалось, что наименьшее значение функции ошибки на валидационной выборке было получено для русскоязычной T5 (by cointegrated). Лучшие форматы входных последовательностей для каждой из исследуемых моделей представлены в таблице 1.

Таблица 1

Лучшие форматы входных последовательностей для исследуемых моделей

Харкте-ристики	mT5	ruT5	mGPT	ruGPT
Eval_loss	51.355	0.027	27.528	17.893
Количество дистракторов за одну итерацию	1	3	1	1
Начало контекста	"lang: ru"	"passage:"	"<s> lang: ru <start_passage>"	нет
Начало вопроса	"</s>"	"question:"	"<start_quest-ion>"	"<s>"
Начало ответа	"</s>"	"answer:"	"<start_answer>"	"<s>"
Начало дистрактора	"</s>"	"distractor:"	"<start_distrac-tor>"	"<s>"
Завершение последовательности	"</s>"	"</s>"	"</s>"	"</s>"

Из-за ограниченных временных ресурсов выбор гиперпарметров обучения для каждой из моделей производился на основании опыта и эмпирических наблюдений при дообучении других языковых моделей. В результате дообучения наименьшее значение функции ошибки на валидационной выборке было получено для русскоязычной T5 (by cointegrated). Результаты дообучения для каждой из исследуемых моделей представлены в таблице 2.

Таблица 2

Характеристики и результаты дообучения исследуемых моделей

Характеристики	mT5	ruT5	mGPT	ruGPT
Тренировочные датасеты	SCIQ, RACE, ОТиПБ	ОТиПБ	SCIQ, RACE, ОТиПБ	ОТиПБ
Eval_loss	0.029	0.012	0.070	0.118
Batch_size	8	8	2	8

продолжение таблицы

Характеристики	mT5	ruT5	mGPT	ruGPT
Epochs	10	20	4	20
Learning_rate	5e-05	2e-05	5e-05	2e-06
Train_seed	0	0	0	0
Weight_decay	0.1	0.1	0.1	0.1

После дообучения моделей и перебора стратегий генерации текста оказалось, что во всех случаях BeamSearch-метод давал лучшие результаты на основании выбранных метрик качества. Было замечено, что для ruT5 (by google) и ruGPT (by sberbank-ai) перебираемые стратегии давали довольно близки результаты, когда в случае с мультиязычными моделями ContrastiveSearch- и Sampling-методы генерировали очень длинные последовательности: данным стратегиям не удавалось найти завершающий токен для прекращения генерации. Конфигурации лучших стратегий для каждой из моделей представлены ниже:

1) mT5 (by google):
   - лучшая стратегия: Beam Search;
   - метрики: BLEU1 – 0.603, BLEU2 – 0.476, BertScore – 0.785;
   - параметры: max_length – 169, num_beams – 20, no_repeat_ngram_size – 2, diversity_penalty – 0.4, temperature – 7.4, repetition_penalty – 3.4, num_beam_groups – 4.
2) ruT5 (by cointegrated):
   - лучшая стратегия: Beam Search;
   - метрики: BLEU1 – 0.601, BLEU2 – 0.466, BertScore – 0.778;
   - параметры: max_length – 157, num_beams – 16, no_repeat_ngram_size – 2, diversity_penalty – 0.4, temperature – 9.0, repetition_penalty – 1.4, num_beam_groups – 2.
3) ruGPT (by sberbank-ai):
   - лучшая стратегия: Beam Search;
   - метрики: BLEU1 – 0.531, BLEU2 – 0.378, BertScore – 0.741;
   - параметры: max_length – 198, num_beams – 28, no_repeat_ngram_size – 0, diversity_penalty – 2.4, temperature – 7.0, repetition_penalty – 1.4, num_beam_groups – 4.
4) mGPT (by sberbank-ai):
   - для данной модели были выбраны параметры генерации дистракторов, полученные для модели ruGPT (by sberbank-ai).

Для анализа качества генерируемых дистракторов было выбрано тридцать записей, каждая из которых относится к одному из трёх классов вопросов: "Определения", "Слова-индикаторы" и "При условии ...". Для каждой записи, состоящей из "контекста", "вопроса" и "ответа" было необходимо получить три дистрактора. В результате генерации и проведения ручной оценки была составлена таблица 3 со статистикой.

Из таблицы 3 можно увидеть, что с генерацией дистракторов лучше справляются модели, предназначенные для одного естественного языка. Мультиязычным моделям не удалось воспользоваться своими знаниями по заданной предметной области из других языков, на которых они были обучены, для того, чтобы генерировать лучшие по качеству текстовые последовательности.

Таблица 3

Результаты ручного анализа генерируемых дистракторов с помощью исследуемых моделей

Характеристики	mT5	ruT5	mGPT	ruGPT
Всего	90	90	90	90
Удалось сгенерировать	80 (88.8%)	80 (88.8%)	56 (62.2%)	83 (92.2%)
Принято	54 (67.5%)	59 (73.7%)	35 (62.5%)	58 (69.8%)
Походят на ответ	20 (25%)	12 (15%)	11 (19.6%)	20 (24.1%)
Не соответствуют контексту	6 (7.5%)	4 (5%)	3 (5.4%)	5 (6.1%)
Содержат ошибки	0 (0.0%)	5 (6.3%)	7 (12.5%)	0 (0.0%)

Можно выделить следующие замечания по характеру генерации дистракторов:

1. Генерируемые дистракторы для одного вопроса очень близки по словарю используемых слов. Отсутствует разнообразие.
2. Можно выделить основной шаблон генерации: используется эталонный ответ на вопрос и в нём заменяются некоторые слова.
3. Модели "плохо" справляются с генерацией "длинных" дистракторов. Из-за общего шаблона построения дистракторов ломается общая семантика текста.

В результате исследования лучшее качество генерации дистракторов показала русскоязычная модель T5 (by cointegrated). Это может обуславливаться тем, что модель помимо обучения без учителя (Unsupervised Learning) на неразмеченных данных была дообучена на большом количестве других Supervised-подзадач обработки естественного языка, когда другие исследуемые модели обучались на 1–2 задачах или только на неразмеченных данных. Можно выделить следующие направления дальнейших исследований по улучшению качества генерируемых дистракторов:

1. Использовать новый алгоритм оптимизации Lion при дообучении моделей.
2. Модифицировать функцию ошибки, чтобы дополнительно наказывать модель за генерацию копии или похожей формы правильного ответа (Answer Negative Regularization).
3. Использовать новые значения гиперпараметров для дообучения моделей.
4. Обучить модель фильтрации генерируемых тестовых заданий (QA filtering model) [5].

Литература

1. Смирнов Н., Семёнов Н., Абашкина Е., Храмцов К. Автоматическая генерация вопросов по текстовому документу // XVI Всероссийская научно-практическая конференция, Петрозаводск. 2022. С. 88–91.
2. Qiu Z., Wu X., Fan W. Automatic Distractor Generation for Multiple Choice Questions in Standard Tests // In Proceedings of the 28th International Conference on Computational Linguistics, Barcelona, Spain (Online). 2020. С. 2096–2106.
3. HuggingFace [Электронный ресурс]. Режим доступа: https://huggingface.co/ (дата обращения 26.04.2023).
4. Armengol-Estapé J., Bonet G., Melero M. On the Multilingual Capabilities of Very Large-Scale English Language Models // In Proceedings of the Thirteenth Language Resources and Evaluation Conference, Marseille, France. 2022. С. 3056–3068.

5. Offerijns J., Verberne S., Verhoef T. Better Distractions: Transformer-based Distractor Generation and Multiple Choice Question Filtering [Электронный ресурс]. Режим доступа: https://arxiv.org/abs/2010.09598 (дата обращения: 26.04.2023).

УДК 51.76

ПРОГНОЗИРОВАНИЕ ГЛИКЕМИИ В УСЛОВИЯХ МАЛЫХ ВЫБОРОК

Мурзина А.А.[1] (аспирант)
Научный руководитель – к.т.н., доцент Гусарова Н.Ф.[1]
1 – Университет ИТМО
e-mail: hloshkin@yandex.ru

Аннотация
В работе рассмотрены различные подходы к анализу моделей, решающих задачу предсказания уровня глюкозы в крови. Описываются статистические подходы сравнения различных моделей, как машинного обучения, так и классических алгоритмов предсказания временных рядов в условиях нестационарности процесса, а также частотных ограничений данных. Рассмотрены преимущества моделей регрессии на небольшом количестве персонализированных данных.
Ключевые слова
Гликемия, сравнение распределений, инсулиновая реакция, близкий горизонт предсказания, анализ интерпретируемости.

Для контроля качества измерителей гликемии (а позднее и моделей) по сравнению с эталонным клиническим значением часто используют анализ диаграммы ошибок Кларка. Несомненно, при разработке предсказательных моделей важно не позволять модели или измерителю совершать оценки гликемии, соответствующие критическим областям диаграммы. Как правило, сравнения значений модели и данных гликемии крови происходит только в моменты, когда были произведены физические измерения. Общепринято использовать среднюю квадратичную ошибку (RMSE) и средняя абсолютная ошибка (MAE) в качестве метрики для сравнения моделей. В связи с широким распространением в различных областях моделей машинного обучения задача сравнения временных рядов приобрела особенную актуальность.

Процесс поступления глюкозы в кровь и расщепления её инсулином нестационарный, зависит от антропометрических параметров человека. Важной проблемой является то, что измерения являются низкочастотными относительно внутренних процессов человека, влияющих на значение гликемии.

Частоты пульсации секреции различных гормонов принадлежат диапазону от двух до пятнадцати минут [1], что не позволяет использовать данные с частотой измерения более пяти минут для построения интерпретируемой и физиологической модели. Так, например, сравнительно недавно появилась возможность собирать данные гликемии человека с минутной частотой, что позволит сформировать гипотезы о секреции гормонов, участвующих в пищеварении, и нарушениях в этих процессах.

Поскольку наибольшая частота одного из источников гормонов, участвующих в усвоении пищи, может превышать половину частоты измерения уровня глюкозы, мы не можем восстановить исходный процесс секреции без искажений согласно теореме Котельникова.

Вследствие того, что процесс линейной интерполяции может сильно не соответствовать истинной функции гликемии от времени. Это, в свою очередь, происходит из-за отсутствия обоснования монотонности уровня глюкозы между измерениями. Таким образом, предсказание значения гликемии на таких доступных данных может иметь вероятностный характер.

Модели с большим количеством параметров склонны к запоминанию, а не обобщению, данных, тогда как могут показывать хорошую точность отклонений от истинных значений в моменты измерения.

В данной работе показаны проблемы применения алгоритмов LSTM и RNN глубокого машинного обучения к задаче прогнозирования гликемии.

Для исследования способности модели к обобщению, а также к возможности интерпретации скрытых представлений рассмотрим распределение ошибок, а также метрики расстояния между распределениями истинных и предсказанных значений глюкозы для RNN на наборе данных OhioT1DM [2].

Ещё одной значимой проблемой моделей машинного обучения на задачах регрессии является неустойчивость, именно поэтому временные ряды не могут предсказывать на длительные промежутки времени вперёд.

Для многих областей задача маломерной регрессии решается классическими алгоритмами, такими как классические и стохастические дифференциальные уравнения и различные вариации моделей авторегрессии. Здесь под небольшим количеством размерностей подразумевается такое количество, которое возможно обобщить для представления в аналитическом виде при тщательном изучении данных экспертом в исследуемой области или же при возможности независимо друг от друга изучить компоненты модели.

Сразу можно отметить, что распределения среднеквадратичных ошибок по пациентам не являются нормальными с p-value=0.001. Рассмотрим результаты работы глубоких сетей с точки зрения смещения распределения ошибок.

Сравнение метрики оптимизации и дивергенции распределений экспериментальных данных и смоделированных представлены на таблице 1 для получасового горизонта предсказания и на таблице 2 для часового.

Таблица 1

Сравнение метрики оптимизации и дивергенции распределений экспериментальных и смоделированных данных для получасового горизонта предсказания

Пациент	RMSE	MAE	Дивергенция Кульбака-Лейблера	Дивергенция Йенсена-Шеннона	Гипотеза о равенстве мат.ожиданий (α=0.99)
540	18,63	13,63	0,008	0,045	Отвергнута
544	16,57	11,46	0,005	0,037	Отвергнута
552	14,92	11,00	0,006	0,038	Не отвергнута
567	19,35	13,33	0.012	0,054	Отвергнута
584	21,55	15,18	0,011	0,052	Отвергнута
596	15,69	10,90	0,005	0,038	Не отвергнута

Таблица 2

Сравнение метрики оптимизации и дивергенции распределений экспериментальных и смоделированных данных для часового для часового горизонта предсказания

Пациент	RMSE	MAE	Дивергенция Кульбака-Лейблера	Дивергенция Йенсена-Шеннона	Гипотеза о равенстве мат.ожиданий (α=0.99)
540	33.01	24,84	0,023	0,076	Отвергнута
544	27,82	20,26	0,015	0,061	Отвергнута
552	26,67	19,96	0,018	0,067	Не отвергнута
567	34,72	24,74	0.032	0,088	Отвергнута
584	34,93	25,52	0,023	0,077	Отвергнута
596	26,72	19,12	0,016	0,064	Не отвергнута

Таблицы получены путём обучения модели RNN с архитектурой, предложенной в работе [3]. На таблицах можно заметить, что не все метрики одинаково отражают качество регрессии на различных пациентах.

Для пациентов 552 и 596 получены лучшие значения метрик RMSE и MAE соответственно. Модель обучилась достаточно для предсказания гликемии выше указанных двух пациентов и нельзя опровергнуть отклонение математического ожидания ошибки предсказания для них.

Для остальных пациентов RNN не смогла так же, как и для 552 и 596 хорошо предсказать значения глюкозы, однако для обозначенных пациентов расхождения по распределениям несколько сильнее, чем для пациента 544, распределение гликемии которого удалось повторить лучше.

На рисунке представлена расходимость предсказания со временем, что отражает проблему неустойчивости модели.

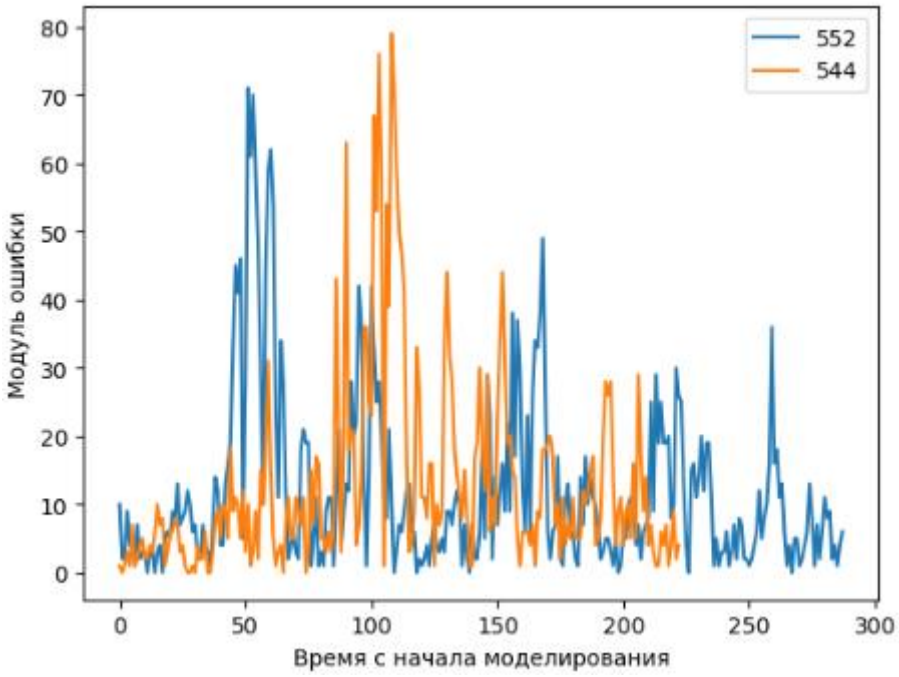


Рисунок. Зависимость абсолютной ошибки гликемии от времени в минутах с момента начала моделирования

Начиная с некоторого момента после получаса от начала моделирования прогноз сильно отклоняется от истинного экспериментального значения и далее лишь иногда бывает близок к данным, однако, основываясь на собственном ошибочно предсказанном временном интервале не имеет возможности вернуться к тому состоянию со скрытым представлением, которое бы не ухудшало значение метрик в дальнейшем.

На рисунке можно увидеть, что временной момент расхождения модели с наилучшим качеством метрик имеет место быть раньше, чем момент расхождения модели, для которой метрики обучения хуже, но само распределение значений уровня глюкозы ближе к истинному экспериментальному.

Поскольку локальные экстремумы могут быть пропущены в ходе измерений, то близость распределений может являться не менее полноценной метрикой качества модели по сравнению со средней квадратической ошибкой.

В работе были проведены статистические тесты и измерение дивергенций между распределениями для оценки применимости методов глубокого обучения для задачи прогнозирования уровня глюкозы. Описаны основные предпосылки для важности сравнения экспериментальных и предсказанных распределений гликемии.

Удовлетворение такими моделями всех обозначенных условий в том числе таких, как временная устойчивость и схожесть распределений, сложная задача, требующая вероятностного подхода к решению. Сочетание интерпретируемых моделей,

соответствующих устойчивым динамическим системам, с методами кластеризации влияний на изменение уровня глюкозы в крови и алгоритмами обучения с учителем будет решением, которое сможет избежать недостатков описанных моделей.

Данные, которые в скором времени будут собраны благодаря прогрессу в области разработки носимых устройства контроля гликемии, позволят использовать знания о пульсирующих процессах секреции инсулина и глюкагона, чтобы построить интерпретируемые устойчивые модели, обладающие сравнимой точностью с результатами методов глубокого обучения.

Литература

1. Pørksen N. The in vivo regulation of pulsatile insulin secretion //Diabetologia. 2002. Т. 45. С. 3-20.
2. Marling C., Bunescu R. The OhioT1DM dataset for blood glucose level prediction: Update 2020 //CEUR workshop proceedings. – NIH Public Access. 2020. Т. 2675. С. 71.
3. Mirshekarian S. et al. LSTMs and neural attention models for blood glucose prediction: Comparative experiments on real and synthetic data //2019 41st annual international conference of the IEEE engineering in medicine and biology society (EMBC). – IEEE. 2019. С. 706-712.

УДК 004.89

КЛАССИФИКАТОР ПРЕДЛОЖЕНИЙ АНГЛИЙСКОГО ЯЗЫКА ДЛЯ ПРОЦЕССНО-СМЫСЛОВОГО АНАЛИЗА ТЕКСТОВ

Никифорова А.Д.[1] (студент), **Роговой В.**[1] (аспирант), **Суров И.А.**[1]
Научный руководитель – к.ф.-м.н. Суров И.А.
1 – Университет ИТМО
e-mail: 334743@niuitmo.ru

Аннотация
Выявление сценарно-логической структуры является важной частью семантического анализа текстов. Мы решаем эту задачу на основе процессной теории смысла. При этом в качестве универсальной структуры процесса используется жизненный цикл, состоящий из этапов наблюдение-восприятие, анализ-объяснение, целеполагание-планирование, исполнение-действие, согласование-доводка и внедрение-апробация. Фрагменты текста классифицируются по этим шести этапам с помощью нейронной сети на основе языковой модели RoBERTa. Модель успешно апробирована на массиве коротких логически завершённых текстов.

Ключевые слова
Сценарный анализ, процессная структура, семантическое моделирование, обработка естественного языка.

Увеличение потоков текстовой информации ставит задачу автоматизации наиболее простых аналитических задач: аннотирования, суммаризации, поиска в текстах нужных смысловых фрагментов. В связи с этим востребованы методы логически-смысловой разметки текста, облегчающие его дальнейшее использование человеком. При этом удобство таких методов прямо зависит от того, насколько результат машинной обработки близок к формату естественного мышления. В этом отношении имеющиеся методы работы с текстом отстают от современных результатов когнитивной семантики, что ограничивает их практическую эффективность.

Одним из интерпретируемых, природоподобных подходов к разметке текстов является сценарный анализ, ставящий задачу выявления в тексте типичных сюжетных линий [1]. Знание таких сюжетов помогает эксперту ориентироваться в незнакомом тексте, увеличивает скорость и повышает качество аналитического результата. Существующие методы сценарного анализа, к сожалению, в основном ограничиваются распознаванием отдельных сюжетных ситуаций, дающих мало информации об общей логически-смысловой структуре текста.

Мы решаем эту задачу на основе универсальной процессно-логической структуры. В этом подходе информационные блоки упорядочиваются путём их размещения в процессном цикле, известном как жизненный цикл [2]. Эта структура, характерная для развития живых организмов и компаний, проектов, продуктов и технологий, социальных и экологических систем, рассматривается в качестве смысловой матрицы естественного мышления (там же). На основе этой структуры и простейшей векторной модели естественного языка возможно нахождение элементарных сценарных циклов [3]. В этой работе ставилась задача развить этот подход с помощью современных методов машинного обучения.

Шестиэтапная модель жизненного цикла [2], показанная на рисунке 1, представляет собой следующую последовательность фаз:

1. Наблюдение и восприятие доступных информационных потоков, заканчивающееся обнаружением требующей внимания новизны.
2. Анализ и объяснение новизны, заканчивающееся построением её модели в когнитивной системе субъекта.

3. Постановка субъектом целей по отношению к понятой новизне, в также построение планов по её достижению, заканчивающиеся созданием проекта.

4. Действия и мероприятия по исполнения плана, заканчивающиеся получением предварительного результата или макета.

5. Согласование и коррекция действий в ходе взаимодействия с объектом, заканчивающееся доводкой результата до окончательного вида.

6. Внедрение и апробация полученного результата в реальных условиях, подведение и рефлексия, заканчивающиеся окончательными выводами.

Рис. 1. Шестиэтапная структура процесса. Надписями в радиальных направлениях показаны точки перехода между этапами

Главным этапам шестиэтапной модели (2, 4, 6, чёрные сектора на рисунке 1) соответствует классическая трёхактная структура литературных и киносценариев завязка – конфликт (борьба) – развязка [4], показанная на рисунке 2. На процессной плоскости последовательности такого типа соответствуют движению по часовой стрелке. Для представления повествования в этой графической форме фрагменты текста должны быть отнесены к соответствующим этапам процесса.

Рис. 2. Трёхэтапная схема литературных и киносценариев [3]

Поставленная задача классификации решена на основе предобученной модели естественного языка семейства трансформеров «RoBERTa» с оптимизатором «Adam» и косинусным планировщиком. Модели дообучалась на вручную размеченных текстах следующих литературных произведений на английском языке:

– русские народные сказки, а также сказки А.С. Пушкина, переведённые на английский язык: «Сказка о золотом петушке» (The Golden Cockerel), «Иванушка-

дурачок» (Ivanoushka the Simpleton), «Сказка о царе Салтане» (The Tale of Tsar Saltan) – 819 предложений;

– роман Курта Воннегута «Завтрак для чемпионов» (Breakfast of Champions) – 649 предложений;

– повесть Джека Лондона «Зов предков» (The Call of the Wild) – 838 предложений;

– исторический роман Джона Уильяма Кея «История войны в Афганистане» (History of the War in Afghanistan) – 630 предложений.

Разметка заключается в отнесении каждого предложения текста двумя либо тремя экспертами к одному из шести этапов, показанных на рисунке 1. Полный набор данных содержал 2936 предложений. Большинство оценок совпадают или расходятся на один класс. Для обучения модели использовались 1287 предложений, для которых все оценки экспертов совпали. Этот набор был разделен на валидационную, тренировочную и тестовую части в соотношении 10:72:18 соответственно.

Существенной проблемой оказалось различие числа предложений в классах, из-за которого предсказания модели перекашивались в сторону наиболее крупного из них. Для решения этой проблемы было опробовано два подхода:

– балансировка классов по нижней границе, при которой содержание каждого класса было сокращено до количества предложений в наименьшем классе;

– использование специальной функции потерь «DiceLoss», предназначенной для классификации при дисбалансе классов.

Все модели обучались на тренировочной части данных и тестировались на тестовой части в течение 30 эпох, причем для моделей с балансировкой классов обрезались обе части данных. Затем выбор лучшей модели осуществлялся на основании одинаковой для всех случаев валидационной части. В качестве метрики использовалась усредненная по всем шести классам f-мера, показанная в таблице. Наилучшей оказалась модель, обученная на полном наборе данных с использованием функции потерь «DiceLoss» (таблица).

Таблица

Точность модели (f1) при различных функциях потерь и балансировки классов в обучающих данных

Функция потерь / Подход	Модель с обычной функцией потерь (кросс-энтропия)	Модель с функцией потерь «DiceLoss»
Балансировка	0.607	0.580
Без балансировки	0.617	0.637

Полученный классификатор был применен к массиву данных ROCStories [5], содержащему 98161 логически целостных текстов, каждый из которых состоит из пяти предложений. Ожидалось, что классификатору удастся выделить в них структуру, соответствующую процессно-смысловому циклу (рис. 1). При этом первое предложение будет классифицировано как наблюдение-восприятие, второе – как анализ объяснение и т.д.

Результат классификации предложений каждого порядкового номера показан в столбцах таблицы на рисунке 3. Так, первое и второе предложения относятся классификатором к первым четырём классам, тогда как пятое (последнее) - к последним трём классам. Усредненная по шести классам f-мера составила 63%. Таким образом предположение частично подтвердилось. Точность классификации при этом может быть значительно улучшена устранением общего перекоса к классу исполнение-действие (четвёртая строка).

	Предложение 1	Предложение 2	Предложение 3	Предложение 4	Предложение 5
Наблюдение	30.33%	17.84%	10.52%	8.48%	9.45%
Анализ	18.59%	14.60%	15.60%	14.31%	13.28%
Целеполагание	16.54%	19.33%	14.48%	10.10%	10.82%
Исполнение	28.41%	35.34%	40.17%	39.78%	25.87%
Согласование	4.07%	8.65%	12.27%	14.83%	15.84%
Внедрение	2.07%	4.24%	6.96%	12.50%	24.74%

Рис. 3. Результат работы классификатора на 98161 текстах датасета ROCStories

Таким образом, разработан алгоритм, классифицирующий предложения естественного языка по этапам универсальной процессно-смысловой структуры. Достигнутая точность определения процессной функции (f1 мера 0.63) может быть повышена путём балансировки выходных классов, расширению массива обучающих данных, а также переходу от задачи классификации к задаче регрессии по фазовой координате. Применение модели к небольшим логически завершённым текстам показало перспективность её использования в задачах сценарного анализа, разметки и аннотирования текстов.

Литература

1. Суворова М.И., Кобозева М.В., Соколова Е.Г., Толдова С.Ю. Извлечение сценарной информации из текстов. Часть 1. Постановка задачи и обзор методов // Искусственный интеллект и принятие решений. 2021. Т. 48. № 6. С. 517-523.
2. Суров И.А. Жизненный цикл: смысловая матрица процессного моделирования // Онтология проектирования. 2022. Т.12. №4(46). С.430–453. DOI:10.18287/2223-9537-2022-12-4-430-453.
3. Суров И.А. Process-Semantic Analysis of Words and Texts // Artificial Intelligence in Models, Methods and Applications / ред. Долинина О. и др. Cham: Springer. 2023. с. 247-260.
4. Field S. Screenplay: The Foundations of Screenwriting. New York: Delta. 2005.
5. Nasrin Mostafazadeh et al. A Corpus and Evaluation Framework for Deeper Understanding of Commonsense Stories // arXiv:1604.01696. 2016

УДК 004.891

ЭВРИСТИКА АЛГОРИТМОВ КЛАССИФИКАЦИИ В ЗАДАЧЕ ОПРЕДЕЛЕНИЯ УБЫТОЧНЫХ СЕГМЕНТОВ ПОРТФЕЛЯ ОСАГО

Раздьяконов А.А.[1] (студент)
Научный руководитель – к.ф.-м.н., доцент Графеева Н.Г.[1]
1 – Университет ИТМО
e-mail: rtm510@mail.ru

Работа выполнена в рамках темы НИР «Эвристика алгоритмов классификации в задаче определения убыточных сегментов портфеля ОСАГО».

Аннотация
В работе представлена оценка использования стандартных алгоритмов классификации данных для фильтрации полисов ОСАГО страховой компании АО «ОСК» посредством функционала методов машинного обучения. Выбор технологий обоснован новыми требованиями МСФО 17, который обязует страховщиков работать с ретроспективой договоров страхования и иначе трактовать заработанную прибыль для правильного расчёта страховых резервов.
Ключевые слова
МСФО 17, ОСАГО, машинное обучение, классификация, алгоритмы сэмплирования.

На 2022 год в России подсчитано более 45 миллионов легковых автомобилей. По статистике на тот же год чуть меньше 105 миллионов человек в РФ страше 18 лет. Таким образом ежегодно каждый 2 совершеннолетний житель нашей страны сталкивается с вопросом оформления договора ОСАГО.

С сентября 2022 года Центробанком был обновлён тарифный коридор для ценообразования ОСАГО: актуальными нижним и верхним значениями являются 1646 и 7535 рублей соответственно [1]. Каждая аккредитованная страховая компания обязана назначать базовый тариф в данных пределах. Но для каждого автолюбителя стоимость ОСАГО меняется из года в год. Причина такого ценообразования кроется в том, что окончательная сумма зависит не только от базовой ставки, а также от набора 6 основных страховых коэффициентов: коэффициент бонус-малус, территориальный коэффициент, возраст и стаж водителя, мощность двигателя, коэффициент сезонности, количество водителей автомобиля [2].

Если взять среднее значение из обозначенного коридора для ОСАГО, то базовый тариф в 4591 рубля под влиянием этих коэффициентов может вырасти для автолюбителя до 20000 рублей. Например, можно смело утверждать: чем меньше у водителя опыта или чем чаще водитель попадает в ДТП, тем выше его коэффициенты и больше стоимость полиса ОСАГО. Механизм оценки клиента по набору определённых показателей прямым образом является прецедентом к уменьшению убыточности компании, а также способствует повышению качества расчётов объёмов страховых резервов.

Основные парадигмы МСФО 17. Так как каждая авария – это убыток для страховой компании, то, как минимум, по набору данных, включающих упомянутые ранее коэффициенты, можно ранжировать клиентов на “потенциально убыточных” и наоборот, “безопасных”. А если добавить такие данные, как: марка автомобиля, год выпуска, иномарка или нет – можно построить куда более точный портрет автомобилиста.

Май 2017 года можно по праву считать новой страницей в страховом деле, так как именно тогда был впервые опубликован новый стандарт финансовой отчетности, именуемый МСФО 17. С тех пор он дополнялся новыми поправками, требованиями и шаблонами, которым должна следовать любая аккредитованная страховая организация. На территории России переход на этот стандарт официально начался только в 2023 году,

что требует от компаний подготовить и – если потребуется – перестроить свою структуру расчета ожидаемой прибыли на новую модель до 2025 года. Однако Центральный Банк РФ ещё с 2021 года частично начал вводить стандарты из МСФО 17, которым должны следовать все страховщики.

Если изучить документ МСФО 17, то, на первый взгляд, глобальная картина ведения бизнеса страховыми компаниями не сильно меняется. Прежний стандарт МСФО 4 требовал строить специальную модель, с помощью которой организация могла рассчитать ожидаемые доходы. МСФО 17 всё так же требует от страховых компаний наличия модели оценки прибыли, но на этот раз, вместо использования фиксированных параметров модели, необходимо принимать во внимание динамические актуарные допущения [3].

В итоге применение подобной модели предполагает позитивный эффект на процессе регулярной оценки деятельности страховой организации, а именно:

1) выделение прибыльных страховые продукты;
2) расчет убытков и соответствующих резервов.

МСФО 17 требует тесную работу с ретроспективой данных страховых договоров компании для построения качественных прогнозов. Впредь компания не может трактовать заработанную премию за отдельно взятый договор как прибыль, так как необходимо принимать во внимание потенциальные убытки и соответственно резервы на месяцы или даже годы вперед на основе исторических данных.

В итоге новый стандарт:

1) повышает прозрачность финансовой информации, предоставляемой страховыми компаниями, что придаёт инвесторам и аналитикам больше уверенности в понимании страховой отрасли;

2) повышает гарантии, что компания сможет выполнить свои страховые обязательства.

Выполнение поставленной задачи прежними средствами классического математического моделирования – чрезвычайно комплексная задача даже для квалифицированных актуариев с большим практическим опытом. Поэтому для построения упомянутой модели большинство компаний прибегают к использованию функционала информационных систем, которые на сегодняшний день предлагают богатый набор способов построения прогнозных моделей, отталкивающихся от исторических данных.

Разработка. За последние 6 лет проводились исследования в применении алгоритмов машинного и глубокого обучения в задачах расчёта убыточности, резервов, прибыли и прочих метрик бизнеса страховщика. В текущей работе приводится оценка применимости алгоритмов машинного обучение к задаче классификации полисов ОСАГО.

Работа состояла из 8 этапов. Данные были предоставлены страховой компанией ОСК по договорам ОСАГО за 5 лет, с 17 по 21 года включительно. После объединения данных по договорам и соответствующим выплатам проводился этап очистки данных, в которых встречались пропущенные записи, ошибки типов и дублирования. Следом был создан столбец целевой переменной, который формировался на основе значений 3 столбцов оригинального фрейма. После оценки корреляции между независимыми и целевой переменными наступил этап работы с моделями машинного обучения. Но сразу же была обнаружена проблема дисбаланса классов. Для решения этой проблемы были проведены эксперименты с алгоритмами сэмплирования, что в итоге привело к позитивному результату, который будет раскрыт подробнее в заключительной части этой работы.

На рисунке 1 представлена информация о датафреймах до подготовки: можно увидеть, что много столбцов содержат пропущенные значения, не все типы соответствуют действительности. Также встречаются столбцы, которые на самом деле

не несут значимой информации, что необоснованно повышает сложность будущих моделей. До подготовки наборы данных содержали в среднем по 250000 записей за каждый год соответственно.

#	Column	Non-Null Count	Dtype
0	Unnamed: 0.1	264457 non-null	int64
1	Unnamed: 0	264457 non-null	int64
2	Договор	264457 non-null	object
3	Unnamed: 1	0 non-null	float64
4	Unnamed: 2	0 non-null	float64
5	Тип значения	264457 non-null	object
6	Unnamed: 4	0 non-null	float64
7	Агент	264456 non-null	object
8	Пролонгация	237559 non-null	object
9	Договорус договор дата	264457 non-null	object
10	А договор дата.Год	264457 non-null	int64
11	А договор дата.Название месяца	264457 non-null	object
12	А договор дата.Квартал	264457 non-null	int64
13	Дата начала	264451 non-null	object
14	Дата окончания	264451 non-null	object
15	Цели использования ТС	263956 non-null	object
16	Стаж	239238 non-null	float64
17	Договорус договор территория по ОСАГО.Субъект РФ	262772 non-null	object
18	Договорус договор территория по ОСАГО.Наименование полное	263630 non-null	object
19	Категория ТС	264447 non-null	object
20	Марка	264348 non-null	object
21	Модель	264252 non-null	object
22	Модель иномарка	264455 non-null	object
23	ТБ	264378 non-null	object
24	КБМ	264485 non-null	object
25	КВС	264378 non-null	object
26	КМ	264378 non-null	object
27	КН	264378 non-null	float64
28	КО	264378 non-null	float64
29	КП	264378 non-null	object
30	КС	264378 non-null	object
31	КТ	264378 non-null	object
32	К прицеп	264375 non-null	object
33	Количество объектов страхования	237565 non-null	float64
34	Экспозиция	237557 non-null	float64
35	Премия (руб)	264254 non-null	float64
36	Заработанная премия	264248 non-null	float64
37	Количество урегулир	13655 non-null	float64
38	Выпл	13533 non-null	float64
39	Колич ЗНУ	11 non-null	float64
40	ЗНУ	11 non-null	float64
41	Оск тип ТС	263966 non-null	object
42	Сумма	217719 non-null	float64
43	Регистратор подразделение	264457 non-null	object
44	Вид договора	264457 non-null	object
45	Прямая выплата	11924 non-null	float64
46	Фикс	11927 non-null	float64
47	Год выпуска	264430 non-null	float64
48	Вид операции	26892 non-null	object
49	Дата начисления	264457 non-null	object
50	Дата расторжения 1	23148 non-null	object
51	Модель строкой	264452 non-null	object
52	Документ основание	2572 non-null	object
53	Тип	5926 non-null	object
54	Договор номер	5926 non-null	object
55	Дата начала ответственности	5921 non-null	object
56	Объект страхования	5926 non-null	object
57	Валюта	5925 non-null	object
58	Страховая сумма	5906 non-null	float64
59	Страховой риск	5926 non-null	object
60	Дата случая	5926 non-null	object
61	Дата Заявления	5926 non-null	object
62	Дата выплаты	5926 non-null	object
63	Вид страхования	5926 non-null	object
64	Учетная группа	5926 non-null	object
65	Дата ПВУ	4243 non-null	object
66	Сумма выплаты	5926 non-null	float64

Рис. 1. Данные до подготовки

На рисунке 2 изображены данные после этапа подготовки. Признаковое пространство уменьшилось до 15, включая one-hot кодировку столбца “Год выпуска” и label-encoding столбца “Марка”. Столбцы “type_flag” и “purp_flag” были сделаны на основе разведывательного анализа данных и анализа корреляций между категориальными фичами и таргетом посредством столбчатых диаграмм. “type_flag” характеризует, принадлежит ли тип транспортного средства страхователя к категориям B или BE. В свою очередь “purp_flag” равен 1, если цель использования ТС - Личная и 0 в прочих случаях.

	Марка	Стаж	Модель иномарка	КБМ	КВС	КМ	КП	КС	КТ	type_flag	purp_flag	Год выпуска
0	Hyundai	14	1	0.50	0.96	1.4	1.0	1.0	1.80	1	1	2011-2022
1	Kia	10	1	0.75	1.01	1.4	1.0	1.0	1.50	1	1	2011-2022
2	ВАЗ/Lada	13	0	0.50	0.96	1.1	1.0	1.0	1.00	1	1	2011-2022
3	Ford	7	1	0.90	1.04	1.2	1.0	0.5	1.80	1	1	2011-2022
4	MAN	6	1	0.70	0.96	1.0	1.0	1.0	1.80	0	1	2006-2010
...	...	...	...	...	...	...	...	...	...	...	...	...
102520	Ford	33	1	0.50	0.93	1.4	1.0	1.0	1.72	1	1	2011-2022
102521	Ford	8	1	0.80	0.94	1.4	1.0	1.0	1.54	1	1	2011-2022
102522	Ford	25	1	0.85	0.93	1.4	1.0	1.0	0.91	1	1	2011-2022
102523	Ford	19	1	0.50	0.94	1.4	1.0	1.0	1.54	1	1	2011-2022
102524	Ford	15	1	0.90	0.94	1.4	1.0	1.0	1.72	1	1	2011-2022

102525 rows × 12 columns

Рис. 2. Данные после подготовки

Как было упомянуто ранее, в процессе подготовки данных был обнаружен дисбаланс: классу 1 целевой переменной принадлежит всего лишь 4 сотых от общего числа записей. В случае обучения моделей на оригинальной выборке получаются очень низкие результаты. Разобраться с этой проблемой было решено с помощью методов оверсэмплинга и андерсэмплинга. В итоге 1-й показал себя гораздо лучше с точки зрения конечного качества, так что выбор остановлен именно на оверсэмплинге [4]. Если конкретно, то использовался алгоритм SMOTE, который синтезирует данные на основе

подхода ближайших соседей. Так как в фичах присутствуют и непрерывные, и номинальные типы данных, использовалась конфигурация SMOTENC [5]. После синтеза классу 1 целевой переменной стала принадлежать половина от общего числа записей. Для обучения моделей было выбрано 5 моделей, а именно:

1. DecisionTreeClassifier.
2. ExtraTreesClassifier.
3. XGBClassifier.
4. GradientBoostingClassifier.
5. AdaBoostClassifier.

На рисунке 3 изображено сравнение графиков ROC-AUC, где слева представлен случай до применения алгоритмов сэмплирования, а справа – после.

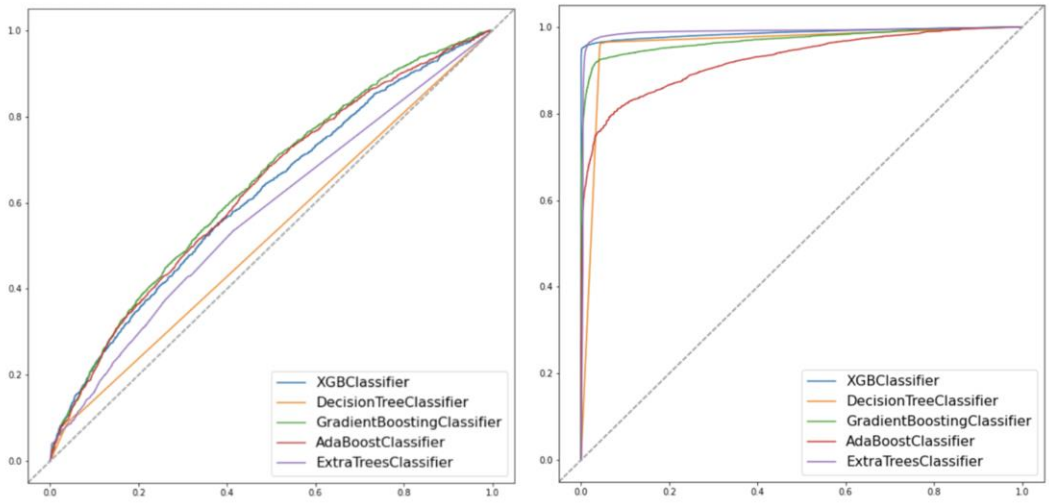


Рис. 3. Графики ROC-AUC классификаторов

Также было разработано приложение с помощью библиотеки Python Eel и фреймворка Vue, которое представлено на рисунке 4. В нём используется модель XGBoost.

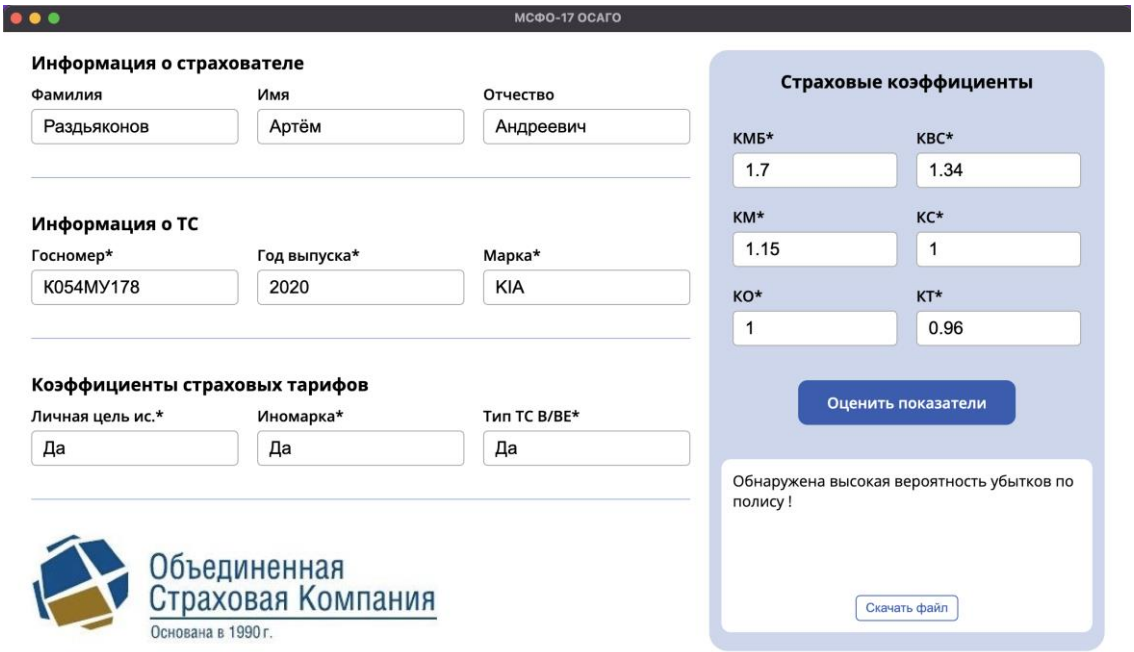


Рис. 4. Приложение для оценки полиса ОСАГО

Выводы. Представленные в таблице значения f-меры позволяют оценить применение алгоритмов бустинга и деревьев решений как потенциально успешный подход для классификации полисов ОСАГО. Разработанное приложение используется с февраля 2023 года в тестовом режиме в одном из филиалов страховой компании ОСК.

Таблица

Значения F-меры алгоритмов

Модель	F1-score
XGBClassifier	0.973
ExtraTreesClassifier	0.973
DecisiontreeClassifier	0.959
GradientBoostingClassifier	0.925
AdaBoostClassifier	0.854

Литература

1. Почему ОСАГО столько стоит: как формируется цена и какие критерии влияют сильнее всего [Электронный ресурс]. Режим доступа: https://www. banki.ru/ news/daytheme/?id=10976158 (дата обращения 08.02.2023).
2. Коэффициенты ОСАГО 2023 [Электронный ресурс]. Режим доступа: https://bip.ru/osago (дата обращения 08.02.2023).
3. Страховой стандарт: какие испытания приготовил МСФО 17 [Электронный ресурс]. Режим доступа: https://forbes.kz// finances/insurance/ strahovoy_ standart_ kakie_ ispyitaniya_ prigotovil_msfo_17 (дата обращения 16.02.2023).
4. 7 Over Sampling techniques to handle Imbalanced Data [Электронный ресурс]. Режим доступа: https://towardsdatascience.com/7-over-sampling-techniques-to-handle-imbalanced-data-ec51c8db349f (дата обращения 20.03.2023).
5. SMOTENC documentation [Электронный ресурс]. Режим доступа: https://imbalanced-learn.org/ stable/references/ generated/ imblearn.over_sampling.SMOTENC.html (дата обращения 21.03.2023).

УДК 004.9

БОТ-НАВИГАТОР ПО ПРОГРАММАМ ОБУЧЕНИЯ РЭУ ИМ. Г.В. ПЛЕХАНОВА

Савко С.В.[1] (ученик), **Виноградова А.П.**[1] (ученик), **Попкова Я.В.**[1] (ученик)
1 – Экономический лицей ФГБОУ ВО «РЭУ им. Г.В. Плеханова»
e-mail: likkera@ya.ru

Аннотация
Разрабатываемая система предназначена для потенциальных абитуриентов РЭУ им. Г.В. Плеханова и представляет собой сервис на базе привычного мессенджера Telegram, который предоставляет информацию о программах обучения, высших школах, а также о реальной студенческой жизни путем неформальной переписки с уже обучающимися студентами-волонтерами. Также система оказывает помощь в выборе направления обучения, проводя профориентационное тестирование и давая рекомендации на основе методов машинного обучения.

Ключевые слова
Чат-бот, машинное обучение, базы данных, профориентация, помощь абитуриентам.

Введение

Все абитуриенты сталкиваются с проблемой выбора направления в ВУЗе. Для осознанного выбора необходимо иметь как можно более точную и полную информацию. Помимо официальной информации о направлениях, факультетах и требованиях при поступлении, абитуриенты стремятся получить информацию напрямую от уже обучающихся студентов в рамках неформального общения в мессенджере.

Для помощи в выборе направления обучения используются профориентационные тесты, анализ результатов которых можно проводить с помощью методов машинного обучения. Этот подход популярен сегодня, и многие абитуриенты доверяют рекомендациям искусственного интеллекта.

Подобная информация присутствует в Интернете, например, на официальном сайте ВУЗа, в различных социальных сетях и профориентационных онлайн-тестах. Уникальность нашего решения заключается в предоставлении этих дополняющих друг друга возможностей в одном сервисе.

Чат-боты сегодня присутствуют почти во всех сферах жизнедеятельности, поэтому абитуриентам будет гораздо удобнее знакомиться с потенциальным местом получения высшего образования, пользуясь мессенджером, в частности, наиболее сегодня популярным, Telegram.

Для подтверждения актуальности нашего проекта было проведено исследование, включающее социологический опрос среди учеников 10-11 классов, и более 80% опрошенных выразили готовность воспользоваться подобным сервисом.

Целью работы является создание чат-бота на платформе Telegram для знакомства с направлениями обучения РЭУ им. Г.В. Плеханова, включая организацию неформального общения со студентами-волонтерами и профориентационную рекомендацию с использованием методов машинного обучения.

Задачи проекта

1. Провести анализ данных (изучить структуру образовательных программ РЭУ и методы машинного обучения).
2. Создать чат-бот в Telegram.
3. Создать профориентационный тест.
4. Создать модель машинного обучения на основе результатов теста.
5. Найти студентов-волонтеров из РЭУ им. Г.В. Плеханова для консультаций

абитуриентов.

6. Запустить чат-бот.

7. Внести корректировки в чат-бот, учитывая результаты проведенного пилотного использования.

Практическая значимость

Разрабатываемая система предназначена для абитуриентов, задумывающихся о выборе направления обучения и желающих получить как можно более полную информацию о ВУЗе.

Машинное обучение для классификации и прогнозирования

Машинное обучение (machine learning) - раздел искусственного интеллекта. Методы машинного обучения используются для распознавания образов, рукописных текстов, компьютерного зрения, прогнозирования, категоризация документов и многих других задач, решение которых возможно с помощью обучения на примерах.

Задача классификации – это определение принадлежности объекта наблюдения к определенному классу. При наличии множества объектов, у которых уже известна принадлежность к классам, можно построить алгоритм, который сможет определять, к какому из классов будет принадлежать новый объект. В терминах машинного обучения объекты с известной принадлежностью к классам называются обучающей выборкой, а задача определения принадлежности нового объекта к классу – классификацией. Модель, с помощью которой определяется принадлежность к классу, – классификатор [1].

Анализ результатов анкетирования, как правило, предполагает распределение множества ответов по известным классам и, таким образом, представляет собой задачу классификации [3, 4]. В частности, на основе результатов профориентационного теста абитуриента можно формировать рекомендации по выбору направления обучения. Для этого в качестве обучающей выборки берутся ответы студентов, уже успешно обучающихся в ВУЗе, а направления, на которых они обучаются представляют собой целевые значения.

Начальные данные для обучения модели можно получить с помощью профориентационного теста, например, на Google forms. Если в процессе эксплуатации система пополняется новыми данными и обратной связью, то эту информацию можно использовать для дообучения модели [2].

Оригинальность нашего решения состоит в интеграции в рамках одного сервиса для удобства пользователя взаимосвязанных возможностей, в том числе получение рекомендации с помощью машинного обучения. Также, в отличие от существующих решений [3, 4], в модуль машинного обучения были добавлены элементы дообучения модели новыми данными и учета обратной связи.

Сценарий использования системы

Рассмотрим сценарий использования системы (рис. 1). В системе есть две категории пользователей: абитуриенты и студенты. Вначале пользователи проходят регистрацию в системе и профориентационное анкетирование.

Результаты анкетирования студентов заносятся в базу данных и далее передаются в модуль машинного обучения в качестве обучающей выборки для дальнейшего повышения точности предсказания классификатора.

Абитуриент получает рекомендацию по выбору направления, полученную модулем машинного обучения. После этого абитуриент просматривает информацию об интересующем направлении, не обязательно рекомендованном, и указывает, со студентами каких высших школ, реализующих данное направление, он хотел бы связаться.

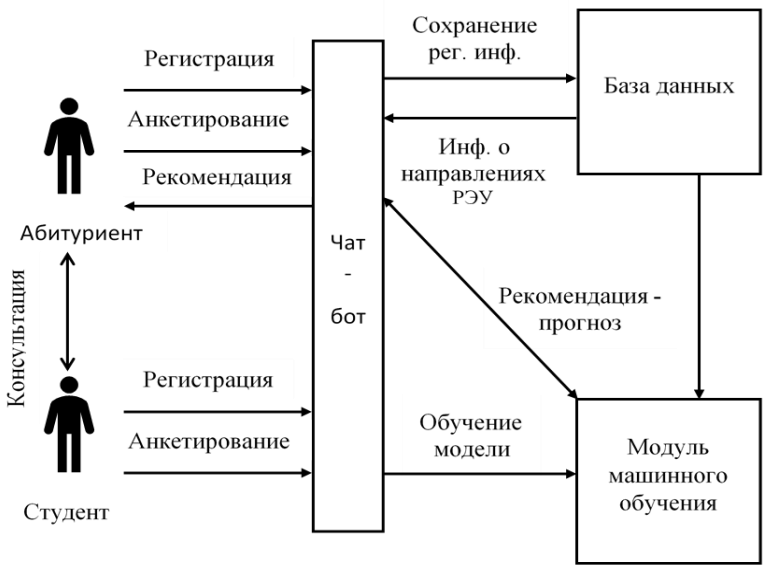


Рис. 1. Сценарий использования системы

После регистрации и выбора интересующих направлений обучения, информация об абитуриентах сохраняется в базе данных, а также направляется студентам-волонтерам для дальнейшего общения в мессенджере. Результаты анкетирования абитуриентов также сохраняются в базе данных и используются для дообучения модели классификатора, но с меньшим весом.

Архитектура системы

Архитектура системы состоит из интерфейса пользователя, реализованного в виде чат-бота, и самой информационной системы, состоящей из модуля машинного обучения и базы данных (рис. 2).

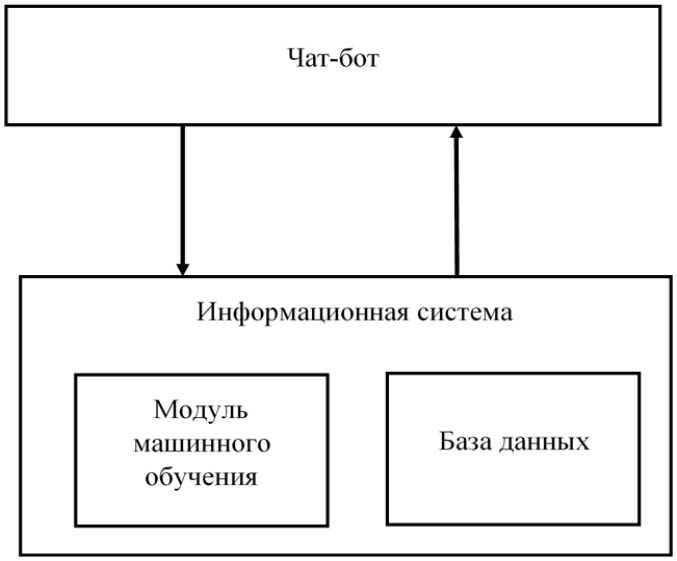


Рис. 2. Архитектура системы

Для реализации используется Python как язык разработки, так как в нем есть все необходимые для системы библиотеки: для реализации чат-бота на Telegram, для работы с базой данных, SQLite, множество библиотек для работы с данными и машинного обучения [5, 6].

Чат-бот

Основные функции чат-бота:

– предоставление информации об образовательных программах и структуре РЭУ им. Г.В. Плеханова;

– регистрация пользователей (абитуриентов и студентов) представлена на рисунке 3;

– анкетирование и рекомендации по выбору направления обучения представлено на рисунке 4 [3].

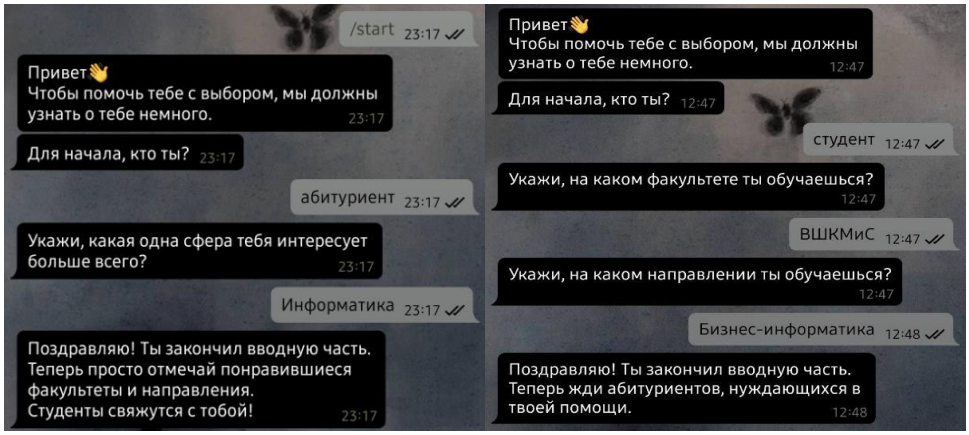


Рис. 3. Регистрация пользователей

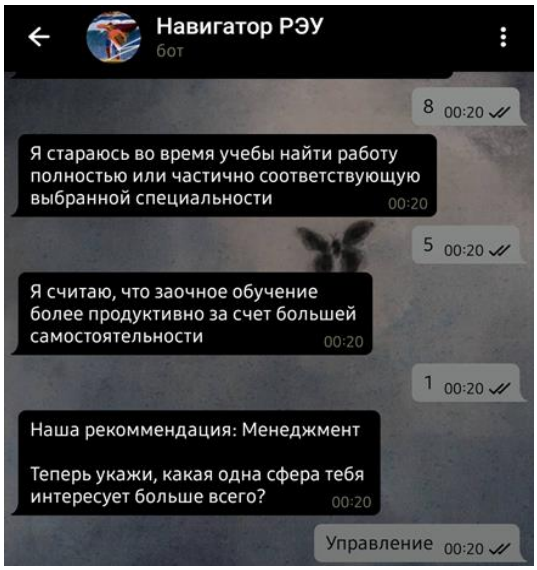


Рис. 4. Рекомендация по выбору направления

База данных

База данных, используемая в системе, содержит информацию об обучающих программах РЭУ, включая направления, высшие школы, реализующие их, необходимые предметы ЕГЭ.

Структура базы данных представляет собой набор связанных таблиц. Например, таблица «Направления» связана с таблицей «Высших школ» по ключу, что представлено на рисунке 5 [4].

Также база данных содержит информацию о зарегистрированных пользователях: абитуриентах и студентах. Результаты анкетирования сохраняются на файловой системе в csv-формате.

REA Имя таблицы: Branches WITHOUT ROWID

	Имя	Тип данных	Первичный ключ	Внешний ключ	Уникальность	Проверка	Не NULL	Сравнение	Generated	Значение по умолчанию
1	id	INTEGER								NULL
2	Name	CHAR								NULL
3	Faculty_id	INTEGER								NULL
4	Sphere	CHAR								NULL
5	Profiles	CHAR								NULL
6	EGE	CHAR								NULL

REA Имя таблицы: Faculties WITHOUT ROWID

	Имя	Тип данных	Первичный ключ	Внешний ключ	Уникальность	Проверка	Не NULL	Сравнение	Generated	Значение по умолчанию
1	id	INTEGER								NULL
2	Name	CHAR								NULL
3	Description	CHAR								NULL
4	Image									NULL
5	Short_name	CHAR								NULL

Рис. 5. Структура базы данных

Модуль машинного обучения

Анкетирование

Профориентационный тест был разработан на основе метода «Якоря карьеры» американского психолога, основателя научного направления «Организационная психология», Э. Шейна [7]. Анкета представляет собой набор из 32 вопросов, на каждый из которых можно ответить по шкале от 1 до 10. Суммируя ответы по определенным категориям, выбирается наиболее приоритетная категория с максимальной суммой и определяется психологическая предрасположенность пользователя.

Сбор данных

Предполагая, что похожие люди выбирают похожие профессии и направления обучения, можно давать рекомендации на основе оценок и личностных характеристик абитуриента, после прохождения профориентационного теста. Определение похожести векторов характеристик и распределение их по группам можно доверить искусственному интеллекту – это задача классификации в машинном обучении. Обучение, и затем прогнозирование, происходит с помощью алгоритма классификатора, представленного в таблице [1, 3].

Таблица

Представление результатов тестирования

Направление	x_1	x_2	x_3	…	x_n
Направление 1	4	5	3	…	7
Направление 2	5	5	4	…	10
Направление 3	3	5	8	…	8
Направление 4	4	6	2	…	0
Направление 5	5	4	3	…	8

В качестве обучающей выборки используются профориентационных анкеты, заполненные студентами и выпускниками РЭУ им. Г.В. Плеханова РЭУ с указанием целевого признака – направление обучения. Результаты анкетирования представляются

в виде целочисленного вектора свойств с целевым свойством «Направление».

Анкетирование для сбора обучающих данных реализовано на основе Google forms. Затем данные импортируются в csv формат и передаются в модуль машинного обучения. Студенты-волонтеры приняли участие в анкетировании еще на стадии разработки системы, таким образом были получены данные для начального обучения.

Во время эксплуатации системы, при регистрации, новые студенты-волонтеры также проходят профориентационный тест. Их ответы сохраняются в базе данных и используются для дообучения модели-классификатора.

Абитуриенты после анкетирования получают рекомендацию от модуля машинного обучения. Если рекомендация совпадает с их ожиданиями, и они выбирают предложенное направление, то ответы также сохраняются в базе данных и используются для дообучения модели, но с меньшим весом, чем ответы студентов.

Реализация

Для построения модели классификатора используются библиотеки Python Pandas и Sklearn.

Рассмотрим простейшую последовательность шагов для работы с результатами анкетирования. Сначала считываем данные в датасет.

```
import pandas as pd
df = pd.read_csv(‘data/anketa.csv’)
```

Далее проводим анализ данных. Пропусков в данных нет. Нормализация не требуется, так как размах всех признаков одинаков и равен 10. Целевая категориальная переменная представляет собой названия направлений и должна быть преобразована в числа с помощью обычной нумерации.

```
from sklearn.preprocessing import LabelEncoder
le = LabelEncoder()
df['target’]=le.fit_transform(df[‘target’])
```

Выделим из датасета целевую колонку:

```
X = df.drop('target', axis = 1)
y = df['target']
```

Выбираем для классификатора модель RandomForestClassifier, наиболее часто используемая для подобных задач классификации [1, 3]. В дальнейшем предполагается тестирование разных моделей машинного обучения, например, нейросетей, с целью улучшения точности предсказания [4].

Обучаем модель:

```
from sklearn.ensemble import RandomForestClassifier
model = RandomForestClassifier()
model.fit(X, y)
```

Чат-бот проводит анкетирование, формирует из ответов пользователя вектор значений, передает его модулю машинного обучения и получает рекомендацию.

```
model.predict(abiturient_df)
```

В результате работы над проектом был реализован бот-навигатор на платформе Telegram для знакомства с программами обучения РЭУ им. Г.В. Плеханова,

предоставляющий следующие возможности:

– информацию об образовательных программах и структуре РЭУ им. Г.В. им. Плеханова;

– организацию неформального общения со студентами-волонтерами;

– профориентационную рекомендацию по выбору направления с использованием методов машинного обучения.

Разработанная система предназначена, в первую очередь, для абитуриентов РЭУ им. Г.В. им. Плеханова. Но может быть легко адаптирована и обобщена для любого ВУЗа.

Бот-навигатор тестировался более, чем 50 пользователями: студентами и абитуриентами. По результатам тестирования были устранены ошибки. Сейчас навигатор находится в стадии пилотного использования.

Дальнейшее развитие системы предполагает следующее:

– размещение бота-навигатора на сервере;

– сопровождение и администрирование;

– сбор данных для модуля машинного обучения;

– исследование разных методов машинного обучения с целью улучшения точности предсказания.

Литература

1. Малышев В.В., Сливкин С.С., Рукавишников В.С., Базаркин Е.В. Применение методов машинного обучения для построения рекомендательной системы отбора анкет абитуриентов. [Электронный ресурс]. Режим доступа: https:// cyberleninka.ru/article/n/primenenie-metodov-mashinnogo-obucheniya-dlya-postroeniya-rekomendatelnoy-sistemy-otbora-anket-abiturientov (дата обращения: 29.10.22).
2. Воронина В.В., Михеев А.В., Ярушкина Н.Г., Святов К.В. Теория и практика машинного обучения. Учебное пособие. Ульяновск: УлГТУ. 2017.
3. Зубкова Т.М., Тагирова Л.Ф. Применение модели нейронных сетей для поддержки принятия решения абитуриента по выбору специальности. [Электронный ресурс]. Режим доступа: https:// cyberleninka.ru/ article/n /primenenie- modeli-neyronnyh- seteydlya- podderzhki- prinyatiya-resheniya- abiturienta-po- vyboru- spetsialnosti (дата обращения: 29.10.22).
4. Рытова Т.А. Применение искусственного интеллекта при обработке анкетных данных. [Электронный ресурс]. Режим доступа: https:// elar.urfu.ru/ bitstream/ 10995/64272/1/m_th_t.a.rytova_2018.pdf (дата обращения: 29.10.22).
5. Документация The Python 3.11.1 Tutorial [Электронный ресурс]. Режим доступа: https://docs.python.org/3/tutorial/index.html. (дата обращения: 15.10.22).
6. Документация библиотеки pyTelegramBotAPI [Электронный ресурс]. Режим доступа: https://pypi.org/project/pyTelegramBotAPI/ (дата обращения: 15.10.22).
7. Чикер В.А. Психологическая диагностика организации и персонала. СПб.: Речь. 2006.

УДК 51-76

РАЗРАБОТКА ИНСТРУМЕНТА ДЛЯ ИНТЕРАКТИВНОЙ ВИЗУАЛИЗАЦИИ И СКАФФОЛДИНГА ГЕНОМНЫХ СБОРОК ПО ДАННЫМ HI-C

Сердюков А.Н.[1] (магистрант), **Замятин А.А.**[1] (аспирант)
Научный руководитель – аспирант ФИТиП Замятин А.А.[1]
1 – Университет ИТМО
e-mail: 264264@niuitmo.ru, azamyatin@itmo.ru

Работа выполнена в рамках темы НИР №622267 «Разработка программного обеспечения для интерактивного скаффолдинга и валидации геномных сборок на основе данных Hi-C».

Аннотация
Целью проекта является разработка программного инструмента, позволяющего в реальном времени визуализировать карты Hi-C контактов и интерактивно выполнять операции скаффолдинга геномных сборок, сопоставимых по размеру с геномом человека. Создание набора программных библиотек и инструментов HiCT было начато в 2022 году. Данная работа посвящена созданию MVP-приложения, реализующего операции скаффолдинга, а также устранению недостатков, выявленных в интерфейсе приложения и в модели данных в процессе апробации HiCT.
Ключевые слова
Интерактивный скаффолдинг, сборка генома, анализ данных, модель данных, интерактивная визуализация, визуализация данных.

Скаффолдинг является последним этапом процесса геномной сборки и заключается в упорядочивании и ориентировании контигов – однозначно определённых последовательностей ДНК, полученных на выходе автоматического сборщика, в последовательности большего размера – скаффолды, которые должны соответствовать истинной последовательности нуклеотидов в молекуле ДНК. С появлением технологий секвенирования, таких как Oxford Nanopore и PacBio HiFi, позволяющих получать протяжённые области ДНК в одном прочтении, можно говорить о наступлении новой эры геномных сборок на уровне хромосомных скаффолдов, где один скаффолд почти полностью отражает последовательность всей молекулы ДНК (хромосомы). Крупные международные консорциумы, такие как The Human Pangenome Reference Consortium (HPRC) и The Vertebrate Genomes Project (VGP) на сегодняшний день заинтересованы в получении большого числа качественных сборок хромосомного уровня. Применение современных алгоритмов сборки для длинных прочтений позволяет получать протяженные контиги, каждый из которых определён однозначно, но неизвестен порядок их следования и ориентация. Для достижения хромосомного уровня сборки необходим этап скаффолдинга. Скаффолдинг по данным Hi-C на данный момент является одним из основных методов, финализирующих процесс геномной сборки. Hi-C – это метод молекулярной биологии, позволяющий получить информацию о взаимном расположении участков ДНК в трёхмерном пространстве [1]. Эти данные можно использовать в процессе скаффолдинга для того, чтобы правильно упорядочить и ориентировать контиги. Существует несколько автоматических скаффолдеров, таких как 3D-DNA, SALSA и YAHS, способных провести значительную часть скаффолдинга в полностью автоматическом режиме, однако для получения сборки наивысшего качества последний этап скаффолдинга на настоящий момент проводится и проверяется человеком [2], что значительно затрудняет общую автоматизацию процесса геномной сборки хромосомного уровня. Таким образом, актуальной является разработка ПО для интерактивного скаффолдинга геномных сборок, что в дальнейшем может открыть возможности большей формализации и автоматизации процесса, в том числе с помощью применения методов машинного обучения и искусственного интеллекта.

На момент создания инструмента HiCT в 2022 году проблема заключалась в отсутствии альтернатив среди ПО с открытым исходным кодом, эффективно решающего задачу интерактивного скаффолдинга для геномов, сравнимых по размеру с геномом человека, на основе данных Hi-C. Единственной альтернативой с открытым исходным кодом является инструмент JuiceBox Assembly Tool (JBAT) [3], разработанный в лаборатории Aiden Lab. Скриншот пользовательского интерфейса JBAT приведён на рисунке 1. Пользователями данного приложения был отмечен ряд недостатков, среди которых наиболее выделяются следующие:

- нерациональный расход оперативной памяти;
- ограниченный набор инструментов автоматического скаффолдинга (поддерживаются только 3D-DNA и SALSA2);
- долгий процесс финализации конечной сборки.

Несмотря на то что в последующих версиях инструмента JBAT некоторые из этих проблем были устранены, разработка альтернативного приложения с открытым исходным кодом, реализующего иной подход к модели хранения данных, является важной задачей. Для этого нами был разработан инструмент HiCT.

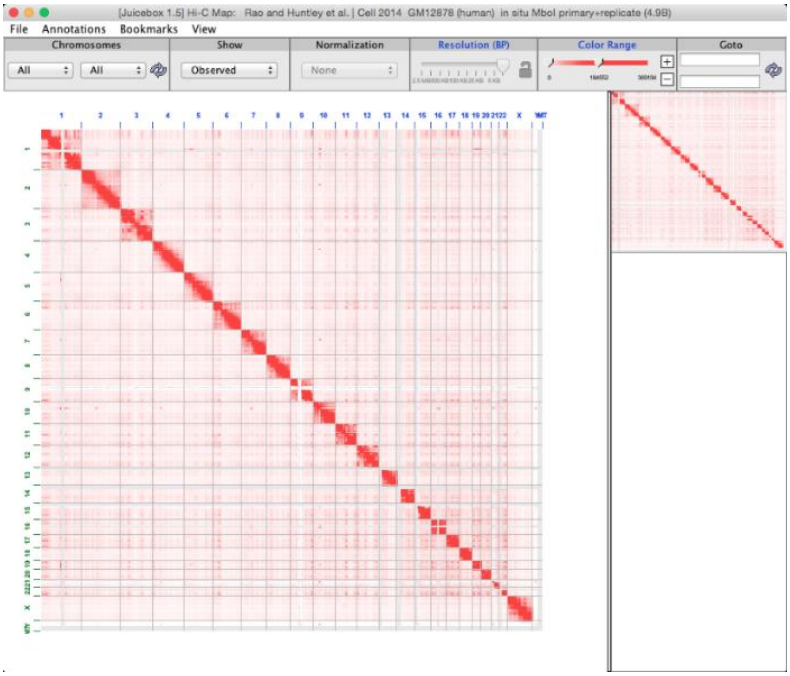


Рис. 1. Скриншот пользовательского интерфейса инструмента JBAT [3]

Матрица контактов Hi-C является симметричной разреженной матрицей. Вдоль её осей расположена текущая сборка, единицы измерения (строк и столбцов) называются *бинами*. При этом известно разрешение матрицы, то есть количество нуклеотидных оснований, объединённых в один бин. Основными операциями скаффолдинга [4] являются:

1. Разворот контига – операция, при которой изменяется ориентация контига в сборке. В матрице контактов Hi-C при этом происходит вертикальное отражение горизонтальной полосы, соответствующей данному контигу и горизонтальное отражение вертикальной полосы, соответствующей данному контигу, а подматрица на пересечении этих полос транспонируется.
2. Транслокация (перемещение) контига в сборке. Данная операция изменяет порядок следования контигов, а именно помещает выбранный контиг на указанную позицию в сборке. Соответствующие изменения отражаются в матрице контактов: отрезок строк и столбцов перемещается на указанную позицию.

3. Объединение контигов в скаффолд – после данной операции, указанный отрезок контигов в сборке группируется, образуя именованную группу – скаффолд.

4. Удаление скаффолда – данная операция удаляет скаффолд, снимая группировку с контигов, которые его образовывали.

Операции разворота и транслокации также могут применяться к скаффолдам: если при развороте отрезка контигов среди них оказались те, которые принадлежат скаффолдам, границы выделения автоматически расширяются так, чтобы покрыть все контиги из всех таких скаффолдов. При этом разворот отрезка контигов разворачивает каждый из контигов, а также меняет порядок их следования на противоположный. Транслокация отрезка контигов перемещает весь отрезок так, чтобы он начинался на указанной позиции, однако не изменяет порядок следования контигов в отрезке после переноса.

При разработке HiCT мы учитывали тот факт, что в процессе ручного скаффолдинга пользователю не требуется отображение всей карты контактов целиком, а лишь интересующий его участок. Более того, те контиги, длина которых не превосходит текущего разрешения (то есть, занимающие менее одного бина при текущем разрешении) можно скрывать, так как они не несут практически-значимой информации. Поэтому, при отрисовке карты в качестве единиц измерения используются пиксели, а не бины: длина контига в пикселях равна его длине в бинах, если контиг не скрыт, или нулю, если контиг скрыт. При этом сами операции скаффолдинга могут изменять даже те участки карты, которые не попадают в область видимости пользователя. Однако, важным свойством операций скаффолдинга является то, что они не изменяют значений на карте контактов, а лишь изменяют их положение. Следует также отметить, что несмотря на двухмерность карты контактов, состояние сборки (взаимное расположение контигов и их ориентация) является одномерными данными. Таким образом, достаточно разбить карту контактов на фрагменты небольшого размера, а также разработать способ адресации, позволяющий собрать карту контактов по текущему состоянию сборки и эффективно реализовать операции скаффолдинга в данной модели. Данный подход позволит загружать только те фрагменты, которые попадают в область видимости пользователя.

В инструменте HiCT данный подход реализован при помощи двух Декартовых деревьев по неявному ключу [6], а также слоя трансляции адресов. В процессе преобразования карты контактов в формат HiCT, исходная матрица равномерно разбивается на блоки – подматрицы размером 256×256, которые, в зависимости от числа ненулевых элементов в блоке, сохраняются в разреженном формате COO (координаты и значения ненулевых элементов внутри каждого блока) или в двухмерном представлении (если более половины элементов в блоке отличны от нуля). Каждый блок в матрице является пересечением двух полос (*stripe*) исходной матрицы – вертикальной и горизонтальной. Также вводится понятие единицы трансляции адресов (Address Translation Unit, ATU), описывающей подотрезок полосы: данный класс хранит в себе идентификатор полосы (stripe_id), начальный и конечный индексы, а также флаг разворота. По каждому контигу строится *дескриптор контига* (класс ContigDescriptor), хранящий его идентификатор (contig_id), имя, длины (в нуклеотидных основаниях и во всех разрешениях), флаги скрытия (на низких разрешениях короткие контиги может быть полезно скрывать), а также для каждого разрешения – упорядоченный список всех ATU, объединение которых даст отрезок контига в исходной сборке и массив префиксных сумм их длин.

Используя дескрипторы контигов, строится Декартово дерево по неявному ключу со случайными приоритетами. Каждая вершина данного дерева хранит дескриптор соответствующего ей контига, размер своего поддерева (в количестве вершин, нуклеотидных оснований, бинах и пикселях при данном разрешении), а также флаг разворота (для отложенных операций). В качестве неявного ключа используется позиция

контига в сборке, неявное вычисление которой происходит при помощи хранящихся в вершине размеров поддеревьев. При этом, все получающиеся неявные ключи согласованы: независимо от выбранной единицы измерения, взаимное расположение вершин в дереве остаётся одинаковым, а также в каждый момент времени контиги, соответствующие вершинам в порядке обхода дерева, образуют текущий порядок сборки. Аналогичное Декартово дерево используется для хранения информации о скаффолдах, однако каждая вершина в нём может как хранить соответствующий ей дескриптор скаффолда (в таком случае она отвечает скаффолду), так и не хранить дескриптор скаффолда (в таком случае, она отвечает отрезку контигов, не сгруппированных в скаффолд).

Использование Декартовых деревьев по неявному ключу со случайными приоритетами и отложенными операциями позволяет эффективно реализовать операции скаффолдинга, а также запрос произвольной подматрицы сборки за счёт эффективной реализации операции разделения дерева и выделения подотрезка.

Для того, чтобы получить подматрицу карты контактов Hi-C по координатам, выраженным в пикселях, необходимо получить список ATU, отвечающих отрезку сборки по вертикали и по горизонтали. Каждый из них можно получить при помощи выделения из Декартова дерева соответствующего подотрезка по координатам в пикселях. Затем, необходимо собрать ATU всех контигов, попавших в данный подотрезок, в один список, с учётом возможных отложенных операций разворота. При помощи двоичного поиска по массиву префиксных сумм длин удалить те ATU, которые не попадают в запрошенный подотрезок и, при необходимости, заменить первые и последние ATU, если границы запроса покрывают их частично. Далее, по двум спискам ATU необходимо построить результирующую матрицу, заполнив её пересечениями ATU по строкам и столбцам. Пересечение пары ATU по строке и столбцу можно найти по следующему алгоритму:

1. Загрузить блок исходной фрагментированной матрицы, соответствующий пересечению полос, идентификаторы которых хранятся в ATU.
2. Из полученной матрицы выделить только те строки и столбцы, которые находятся между граничными координатами, указанными в ATU.
3. При наличии флага разворота в ATU, выполнить отражение полученной матрицы по строкам и/или по столбцам.

Следует отметить, что среднее время работы операций выделения подотрезка в Декартовом дереве зависит от количества вершин логарифмически [6], а время работы обхода подотрезка линейно зависит от количества вершин, попавших в него. Таким образом, при фрагментации на блоки небольшого размера, достигается эффективность во времени работы предложенной реализации. Также, в силу одномерности текущего состояния сборки, достигается эффективность в использовании оперативной памяти.

Работа с предложенными моделью данных и форматом файла была реализована в виде набора программных инструментов HiCT. Для того, чтобы функционально-независимые части могли использоваться и разрабатываться независимо, при разработке HiCT был использован модульный подход: выделены три основных модуля, а также набор утилит. В таблице представлены основные характеристики модулей инструмента HiCT.

Основная библиотека HiCT (модуль HiCT_Library) реализует операции для работы с моделью данных и предложенным форматом файла HiCT, а именно: запросы произвольной прямоугольной области Hi-C карты при текущем состоянии сборки, операции скаффолдинга и операции экспорта сборки и геномного контекста области. Она разрабатывается с учётом планов по дальнейшей автоматизации процесса скаффолдинга, а также глубокого анализа Hi-C карт, в том числе с помощью применения методов ИИ. Также она упрощает интеграцию с существующими инструментами анализа данных на Python за счёт использования совместимых интерфейсов.

Таблица

Основные характеристики модулей инструмента HiCT

Модуль	HiCT_Library	HiCT_Server	HiCT_WebUI
Функциональное предназначение	Библиотека операций над файлами HiCT	Веб-сервер для взаимодействия библиотеки и веб-интерфейса	Пользовательский веб-интерфейс HiCT
Основной язык реализации	Python	Python	TypeScript
Использованные фреймворки и технологии	HDF5	Flask	Vue.js, OpenLayers

Модуль пользовательского интерфейса (HiCT_WebUI) был разработан в виде веб-интерфейса, так как этот подход позволяет сделать приложение кросс-платформенным (при наличии запущенного сервера, для работы с HiCT пользователю необходим лишь современный веб-браузер), а также за счёт использования хорошо зарекомендовавших себя существующих решений для отрисовки карт. Пользовательские элементы преимущественно реализованы с использованием библиотеки Bootstrap, интерактивная область Hi-C карты реализована при помощи картографической библиотеки OpenLayers [7]. Изначально OpenLayers разрабатывалась для отображения географических карт на разных разрешениях, однако в виду схожести данной задачи с задачей отображения карт контактов, её удалось применить для визуализации карт Hi-C. Для использования OpenLayers, необходимо реализовать запрос квадратной подматрицы текущей сборки в формате изображения PNG (*тайла*), что эффективно достигается при помощи библиотеки HiCT. Сжатие, хранение, кэширование и масштабирование тайлов не требуется реализовывать вручную, так как это эффективно реализовано в библиотеке OpenLayers.

Для связи веб-интерфейса с библиотекой HiCT был реализован сервер на базе фреймворка Flask, который позволяет веб-интерфейсу вызывать методы библиотеки HiCT. Также, именно на стороне сервера реализовано преобразование подматриц в изображения, что позволяет добавить последующие шаги в обработке за счёт использования сторонних библиотек. Скриншот текущей версии веб-интерфейса приведён на рисунке 2.

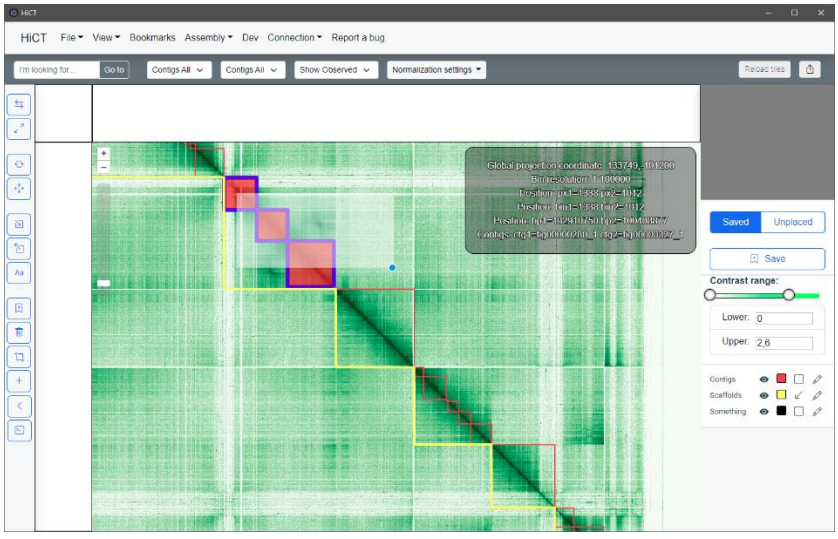


Рис. 2. Скриншот пользовательского интерфейса HiCT

Благодаря персистентному подходу в реализации Декартова дерева по неявному ключу с отложенными операциями, предложенная модель данных может использоваться в многопоточной среде. Однако, некоторые фундаментальные ограничения языка Python не позволяют в полной мере реализовать поддержку многопоточности в HiCT_Server и это ставит перед нами новые задачи по разработке.

Набор вспомогательных утилит включает в себя конвертер данных Hi-C из формата Cooler в формат HiCT. Cooler [5] был выбран ввиду его распространённости и универсальности для хранения карт контактов Hi-C, а также наличия множества инструментов для преобразования из других форматов хранения Hi-C карт. Это позволяет производить интерактивный скаффолдинг геномных сборок, полученных не только после автоматического скаффолдинга 3D-DNA и SALSA2, но и скаффолдерами нового поколения, такими как YAHS.

Данная реализация HiCT использовалась при сборке геномов девяти комаров семейства *Anopheles*, был составлен список замечаний к пользовательскому интерфейсу. В одном из девяти случаев исходные данные были сравнительно низкого качества, из-за чего был выявлен ряд недостатков в исходной версии модели данных при работе с большим числом коротких контигов. В текущей версии модели данных эти недостатки были устранены. Новая структура модели позволила существенно упростить и ускорить процесс конвертации данных из формата Cooler в формат HiCT. Программный интерфейс библиотеки HiCT унифицирован с Cooler в части запросов произвольной прямоугольной области для упрощения разработки и тестирования. В дальнейшем планируется создание режима «валидации» сборки, в котором будет отключён функционал скаффолдинга для оптимизации множественных запросов подматриц. Этот режим позволит упростить разработку расширений для HiCT, среди которых планируется реализация модуля для поиска геномных перестроек при помощи алгоритмов глубокого машинного обучения и ИИ с последующей его в веб-интерфейс.

Исходный код инструмента HiCT, документация, Docker-контейнеры и предварительно собранные версии приложений, а также инструкции по их использованию размещены в открытом доступе на платформе GitHub: https://github.com/ctlab/HiCT.

Литература

1. Lieberman-Aiden E. Comprehensive mapping of long-range interactions reveals folding principles of the human genome // Science. 2009. Т. 326. № 5950. С. 289–293.
2. Rhie A. Towards complete and error-free genome assemblies of all vertebrate species [Электронный ресурс]. 2020. URL: https://www.biorxiv.org/content/early/2020/05/23/2020.05.22.110833 (дата обращения: 12.01.2023).
3. Durand N.C. Juicebox Provides a Visualization System for Hi-C Contact Maps with Unlimited Zoom // Cell Syst. 2016. Т. 3. № 1. С. 99–101.
4. Genome Assembly Cookbook [Электронный ресурс]. URL: https://www.dnazoo.org/methods (дата обращения: 10.02.2022).
5. Abdennur N., Mirny L.A. Cooler: scalable storage for Hi-C data and other genomically labeled arrays // Bioinformatics. 2019. URL: https://doi.org/10.1093/bioinformatics/btz540 (дата обращения: 17.11.2022).
6. Декартово дерево по неявному ключу – Викиконспекты [Электронный ресурс]. URL: https://neerc.ifmo.ru/wiki/index.php?title=%D0%94%D0%B5%D0%BA%D0%B0%D1%80% D1% 82%D 0%BE% D0%B2%D0% BE_%D0%B4%D0% B5%D1%80 %D 0%B5%D0 %B2% D0%BE _%D0%BF% D0%BE_%D0 %BD%D0%B5 %D1%8F%D0 %B2%D0%BD %D0%BE %D0% BC% D1%83_%D0 %BA%D0%BB %D1% 8E%D1 %87%D1%83 (дата обращения: 10.02.2022).
7. OpenLayers [Электронный ресурс]. URL: https://openlayers.org/ (дата обращения: 12.01.2023).

УДК 004

ИСПОЛЬЗОВАНИЕ МОБИЛЬНЫХ ПРИЛОЖЕНИЙ ДЛЯ ОТКАЗА ОТ ВРЕДНЫХ ПРИВЫЧЕК

Сташишин А.И.[1] (студент)
Научный руководитель – д.э.н., профессор Максимова Т.Г.[1]
1 – Университет ИТМО
e-mail: alex.stashishin@gmal.com

Работа выполнена в рамках темы НИР «Определение необходимого функционала для разрабатываемого приложения».

Аннотация
В настоящее время, смартфоны позволяют использовать различные приложения. Появляются различные мобильные медицинские методы, нацеленные на улучшение физической подготовки, здорового питания и привычек, таких как отказ от курения, уменьшение употребления алкоголя и поддержание психического здоровья. В данном исследовании изучается, насколько эффективны такие приложения в изменении поведения людей и какие функции должны использоваться в таких медицинских методах и исследованиях. В данном исследовании мы использовали международную базу PabMed определяются функции, которые необходимы для того, чтобы пользователи отказывались от вредных привычек.
Ключевые слова: эффективность; поведение в отношении здоровья; рейтинг; пропаганда здорового образа жизни; неинфекционное заболевание; мобильные приложения; смартфон; мобильный телефон; цифровое здоровье; изменение поведения, обзор.

У многих современных людей существует потребность избавиться от вредных привычек. Однако не все могут справиться самостоятельно, и не все мобильные приложения, предназначенные для борьбы с вредными привычками, помогают пользователям. Более восьмидесяти процентов граждан с зависимостями не имеют возможности лечиться в стационарном режиме. Однако количество пользователей мобильных телефонов с расширенным функционалом с каждым годом увеличивается [1]. Также увеличивается среднее время использования мобильных приложений, поскольку это является удобным и простым способом взаимодействия с смартфоном [2]. Мобильное здравоохранение может улучшить доступ к медицинской помощи. Эти технологии открывают значительные возможности для доступа к людям, нуждающимся в медицинской помощи, там и тогда, когда это может быть затруднительно на практике. Поэтому существует необходимость проанализировать какими функциями и особенностями должно обладать мобильное приложение, чтобы иметь высокую эффективность в борьбе с вредными привычками.

Целью исследования является определение функций, которые необходимы в мобильных приложениях для того, чтобы пользователи смогли эффективно отказываться от вредных привычек. Для достижения цели, необходимо провести поиск, фильтрацию и анализ публикаций по биомедицинским исследованиям. Необходимо определить наиболее эффективные способы использования мобильных приложений для отказа от вредных привычек.

Использовалась международная база данных PubMed поскольку она представляет собой электронно-поисковую систему публикаций медицинских исследований. Данная база данных является интересной поскольку хорошо разработанных мобильных приложений на данный момент много, но для высокой эффективности нужно учитывать человеческие особенности поведения.

Выявление функций, которые могут повысить эффективность вмешательства, может помочь в разработке приложений для изменения поведения в отношении здоровья и помочь в оценке сложных вмешательств. Были проанализированы публикации,

индексируемые в базе PubMed в январе 2023 года. При использовании тематических фильтров и поискового запроса (((mobile app) OR (mobile application)) AND (bad habits)) OR (((mobile app) OR (mobile application)) AND (bad habits) OR (smoking OR Alcoholism OR addiction)) в результате изучения названий и аннотаций статей были выделены обзоры, соответствующие цели исследования. Исследования были включены, если оценивали вмешательства, реализуемые с помощью приложений. При анализе высокую эффективность продемонстрировали функции получение уведомлений, отслеживания своих результатов, мотивация и поддержка.

Процесс выбора публикаций для проведения исследования:

- во время первичного обзора было найдено 3450 публикаций;
- с использованием фильтров 126 публикация;
- глубокий анализ был проведен над 12 публикациями (рисунок).

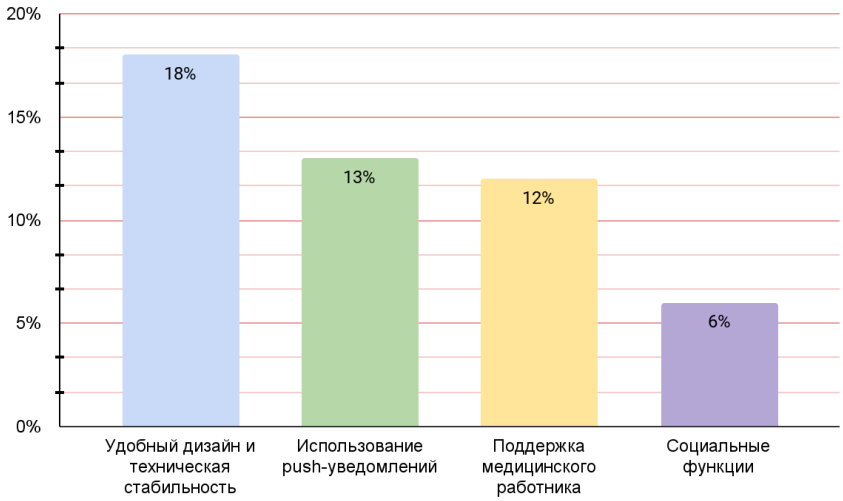


Рисунок. Функции, которые влияют на результативность пользователя

Исследование показало, что для эффективного изменения привычек пользователей, необходима комбинация функций, к которой относятся:

1. Удобный дизайн и техническая стабильность являются основными параметрами, которые пользователи мобильных приложений ожидают в продукте. Удобный дизайн обеспечивает легкость в использовании и позволяет пользователям быстро выполнить необходимые действия. Техническая стабильность же гарантирует бесперебойную работу приложения и делает его более надежным и доверительным в глазах пользователей. Оба этих аспекта обеспечивают позитивный опыт работы с приложением и повышают вероятность того, что пользователи будут продолжать пользоваться приложением. Кроме того, удобный дизайн позволяет пользователям быстро находить нужные функции, а стабильность приложения гарантирует его бесперебойную работу [4, 6, 7]. Это особенно важно для медицинских приложений, где ошибки могут привести к серьезным последствиям для здоровья пациента.
2. Использование push-уведомления могут быть очень полезными для мобильных приложений, так как они могут напоминать пользователям о важных событиях, встречах, приемах лекарств. Кроме того, push-уведомления могут быть использованы для предоставления информации о важных обновлениях приложения и новых функциях [2, 3].
3. Поддержка медицинского работника является очень важным фактором для пользователей мобильных медицинских приложений. Пользователи ожидают от приложения быструю и грамотную помощь медицинских специалистов, включая консультации, рекомендации. Кроме этого, крупные медицинские приложения включают базы данных с болезнями, препаратами, что позволяет использовать приложение

независимо от местонахождения. Для оказания более эффективной помощи медицинскому работнику, необходимо, чтобы приложение создавало графики, отображающие изменения поведения и самочувствия пациента в течение определенного периода времени. Это поможет медицинскому работнику более точно оценить состояние пациента и принять соответствующие меры для улучшения его здоровья или поддержки [2].

4. Социальные функции также способствуют созданию позитивного опыта пользователей мобильных приложений. Большинство пользователей хотят иметь возможность общаться с другими пользователями, делиться своими достижениями, узнавать интересные и полезные новости из различных областей [5].

Также было выявлено, что преимущества мобильных медицинских методов в изменении поведения пользователей заключается в гибкости и доступности приложений. Они могут использоваться в любое время и в любом месте, по сравнению с традиционными методами, которые требуют более жесткого графика или определенного окружения. Это делает мобильные медицинские методы более доступными для большего количества пользователей и повышает их эффективность в изменении здорового поведения. Выделены факторы, влияющие на приверженность к предполагаемому использованию. Эти факторы были дополнительно классифицированы [1, 4].

Некоторые из публикаций включали в себя исследовательскую работу, целью которой было выяснить мнения испытуемых про приемлемость мобильных вмешательств, а также получить дополнительные идеи для дальнейших исследований и разработок, направленных на улучшение этой приемлемости. Дополнительно мы провели качественные исследования, которые позволили изучить мнения и опыт участников относительно проводимых мероприятий. Мы проанализировали имеющиеся документы и выделили ключевые темы, которые были сопоставлены с результатами исследований. В результате были выявлены следующие общие темы:

1) у участников имеются опасения относительно безопасности, конфиденциальности и прочей защищенности проводимых мероприятий через смартфон и как именно хранятся их данные. Важно убедиться, что мероприятия проходят на безопасной платформе и конфиденциальная информация зашифрована [2, 6];

2) необходимо создать интересный и удобный формат проведения мероприятий. Важно найти баланс между комфортом и целью мероприятия, чтобы участники получили максимальную пользу [2, 5, 7];

3) в некоторых случаях участникам может потребоваться удаленная поддержка. Важно убедиться, что система поддержки работает качественно, корректно и доступна всякий раз, когда возникает необходимость [1, 3];

4) разница в технической грамотности, а также различные технические проблемы, которые могут существовать, ведь они могут повлиять на приемлемость мероприятия. Поэтому необходимо разработать, внедрить, а также объяснить участникам инструкции по использованию технологии и помочь им преодолеть технические трудности [1];

5) потенциальное влияние и желание пользователя принимать активное участие остается еще одним фактором, который нужно учитывать. Важно убедиться, что мероприятие представляет действительную пользу для каждого участника и дополнительно мотивировать их участие [4].

Некоторые участники интерактивных сессий могли испытывать трудности с пониманием технических инструкций или использованием сложных функций мобильных приложений. Это могло приводить к недопониманию и ошибкам, которые могли повлиять на их удовлетворенность опытом использования мобильных приложений для интерактивных сессий. Для того, чтобы оптимизировать этот процесс и сделать его более удобным для пользователей, необходимо обеспечить поддержку пользователей и обучающие материалы для помощи им в понимании технических

инструкций и использовании функций мобильных приложений. Также необходимо устранять технические проблемы и дефекты мобильных приложений для обеспечения надежного и быстрого функционирования [7].

Были высказаны тревоги относительно влияния психологического состояния на участие в интервенции [2]. Дополнительно можно сказать, что помимо настроения, психологическое состояние участников может включать и другие факторы, такие как уровень тревожности или привычность к социальным взаимодействиям. Необходимо учитывать все эти нюансы при работе с участниками проекта, чтобы обеспечить им подходящую мотивацию и комфортную среду для участия. Кроме того, определенные методы интервенции могут быть более предпочтительны для участников в зависимости от их текущего психологического состояния. Так, например, для тех, кто испытывает высокий уровень тревожности, может быть эффективна более индивидуальная, личностно ориентированная интервенция, в то время как для тех, кто склонен к социальному изоляционизму, может быть полезной групповая динамическая терапия. Все эти факторы следует учитывать при разработке и реализации проекта, чтобы добиться максимально полезного и продуктивного воздействия на участников.

Мы считаем, что приложения, которые предлагают даже небольшую пользу для здоровья, все же могут быть ценным вмешательством в общественное здравоохранение, если охват населения достаточно высок.

Проведена выборка и анализ способов, которыми можно влиять на пользователя. Исследований из России не найдено, что обосновывает необходимость разработки мобильных приложений, которые будут обладать функциями повышающие результативность отказа от вредных привычек, с обязательной проверкой их эффективности. Если в приложении пользователь будет получать уведомления, сможет отслеживать количество дней своего воздержания, а также будет получать поддержку и мотивацию, то он с большей вероятностью достигнет своей цели. Данное исследование подчёркивает необходимость тесного сотрудничества разработчиков с врачебным сообществом и потребителем, чтобы добиться больших результатов.

Направления будущих исследований:

1. Тщательные исследования для проверки эффективности приложений и их функций.
2. Проведении оценочных исследований в России для поиска потребности в альтернативном способе воздействия на пациента.
3. Продолжить исследование эффективных способов воздействия и благоприятные функции для пользователей.

Литература

1. Rubaab Bahadoor, Jean-Marc Alexandre, Lucie Fournet, Thibaut Gellé, Fuschia Serre, Marc Auriacombe Inventory and Analysis of Controlled Trials of Mobile Phone Applications Targeting Substance Use Disorders: A Systematic Review. [Электронный ресурс]. Режим доступа: https://pubmed.ncbi.nlm.nih.gov/33692708/. (дата обращения: 02.02.2023).
2. Fiona H McKay, Annemarie Wright, Jane Shill, Hugh Stephens, Mary Uccellini Using Health and Well-Being Apps for Behavior Change: A Systematic Search and Rating of Apps [Электронный ресурс]. Режим доступа: https:// pubmed. ncbi. nlm. nih.gov/31274112/ (дата обращения: 03.02.2023).
3. Madison Milne-Ives, Ching Lam, Caroline De Cock, Michelle Helena Van Velthoven, Edward Meinert Mobile Apps for Health Behavior Change in Physical Activity, Diet, Drug and Alcohol Use, and Mental Health: Systematic Review [Электронный ресурс]. Режим доступа: https://pubmed.ncbi.nlm.nih.gov/32186518/ (дата обращения: 04.02.2023).
4. Jing Zhao, Becky Freeman, Mu LiCan Mobile Phone Apps Influence People's Health Behavior Change? An Evidence Review. [Электронный ресурс]. Режим доступа:

https://pubmed.ncbi.nlm.nih.gov/27806926/ (дата обращения: 05.02.2023).
5. Babak Tofighi, Chemi Chemi, Jose Ruiz-Valcarcel, Paul Hein, Lu Hu Smartphone Apps Targeting Alcohol and Illicit Substance Use: Systematic Search in in Commercial App Stores and Critical Content Analysis [Электронный ресурс]. Режим доступа: https://pubmed.ncbi.nlm.nih.gov/31008713/ (дата обращения: 06.02.2023).
6. Jackson M. Steinkamp, Nathaniel Goldblatt, Jacob T Borodovsky, Amy LaVertu, Ian M Kronish, Lisa A Marsch, Zev Schuman-Olivier Technological Interventions for Medication Adherence in Adult Mental Health and Substance Use Disorders: A Systematic Review [Электронный ресурс]. Режим доступа: https://pubmed.ncbi. nlm.nih.gov/30860493/ (дата обращения: 07.02.2023).
7. Robert Jakob, Samira Harperink, Aaron Maria Rudolf, Elgar Fleisch, Severin Haug, Jacqueline Louise Mair, Alicia Salamanca-Sanabria, Tobias Kowatsch Factors Influencing Adherence to mHealth Apps for Prevention or Management of Noncommunicable Diseases: Systematic Review [Электронный ресурс]. Режим доступа: https://pubmed.ncbi.nlm.nih.gov/35612886/ (дата обращения: 02.02.2023).

УДК 004.62

РАЗРАБОТКА АЛГОРИТМА ДЛЯ РАСШИФРОВКИ И ОПРЕДЕЛЕНИЯ АББРЕВИАТУР В ТЕКСТЕ

Тирский Б.Д.[1], Рунушкина О.Р.[1], Верко Р.А.[1]
Научный руководитель – к.п.н., доцент Авксентьева Е.Ю.[1]
1 – Университет ИТМО

Аннотация
Данная статья посвящена разработке алгоритма для расшифровки и определения аббревиатур в тексте. Так как большинство документов хранятся в электронном виде и для оптимизации размеров используются различные сокращения и аббревиатуры в том числе, поэтому задача достаточно актуальная. Помимо дословной расшифровки аббревиатур, требуется еще понять сам смысл, который несут в себе аббревиатуры, а это уже нетривиальная задача.
Ключевые слова
Корпус текстов, аббревиатура, семантика, эквивалентность, словарь, API.

Задача расшифровки аббревиатур стала достаточно актуальной с появлением электронной документации, так как объемы текстов увеличились и, вручную обрабатывать такие корпуса трудозатратно, поэтому появились мысли о том, чтобы оптимизировать данный процесс. Одна из основных проблем – это сложность обработки естественных языков, то есть требуется проделать большое количество процедур по приведению слов к нормальной форме, например удалению знаков препинания и многие другие [1]. Помимо всего этого требуется точно определить семантическую составляющую у слов и аббревиатур.

Существует несколько методов по расшифровке аббревиатур:

1. Анализ контекста – это выделение семантической составляющей из аббревиатуры исходя из слов, которые окружают саму аббревиатуру, чаще всего использует небольшую ширину контекстного окна, например 3 или максимум 5 слов вокруг аббревиатуры, но чем больше ширина окна, тем больше времени требуется на анализ.
2. Онлайн словари – существуют ресурсы, например Acronym Finder и Abbreviations.com, которые позволяют расшифровать аббревиатуры, но словари приспособлены только английские аббревиатуры и для русского языка они не подходят.
3. Машинное обучение – это, наверное, самый популярный метод, который используют большинство компаний, так как он достаточно эффективный и не требует больших затрат в плане вычислительной мощности конкретно для данной задачи.

Самым первым пунктом предстояло выявить в выборке аббревиатуры, реализовывалось это с помощью регулярных выражений, так как при использовании регулярных выражений можно было отобрать и комбинированные аббревиатуры, которые состоят из букв как в большом, так и в маленьком регистре.

В данной работе использовался комбинированный метод контекстного анализа и онлайн словарь sokr.ru [2], помимо всего использовалась статистическая мера для оценки важности слова TF-IDF для применения контекстной семантической составляющей в расшифровке аббревиатур.

Для решения поставленной задачи можно было бы использовать библиотеку nltk в языке программирования Python, но реализуемый в библиотеке метод зарекомендовал себя не лучшим образом, так как словарь аббревиатур nltk содержит очень ограниченное количество расшифровок и появляется потребность в ручном их добавлении. После расширения с помощью онлайн словаря удалось увеличить количество расшифровок практически до ста процентов [3].

Существует ряд сервисов, которые подразумевают использование сторонних API для расшифровок аббревиатур, например, WolframAlpha, с помощью которого можно получить расшифровку своей аббревиатуры при запросе через клиент, но у него также есть недостаток в небольшом объеме словаря для определений аббревиатур на русском языке [4].

Ниже описана общая схема алгоритма, на которой видно, что на выборке из словаря происходит обучение модели и проверка полученных ответов, на этой же модели в дальнейшем можно определять и расшифровывать аббревиатуры (рис. 1).

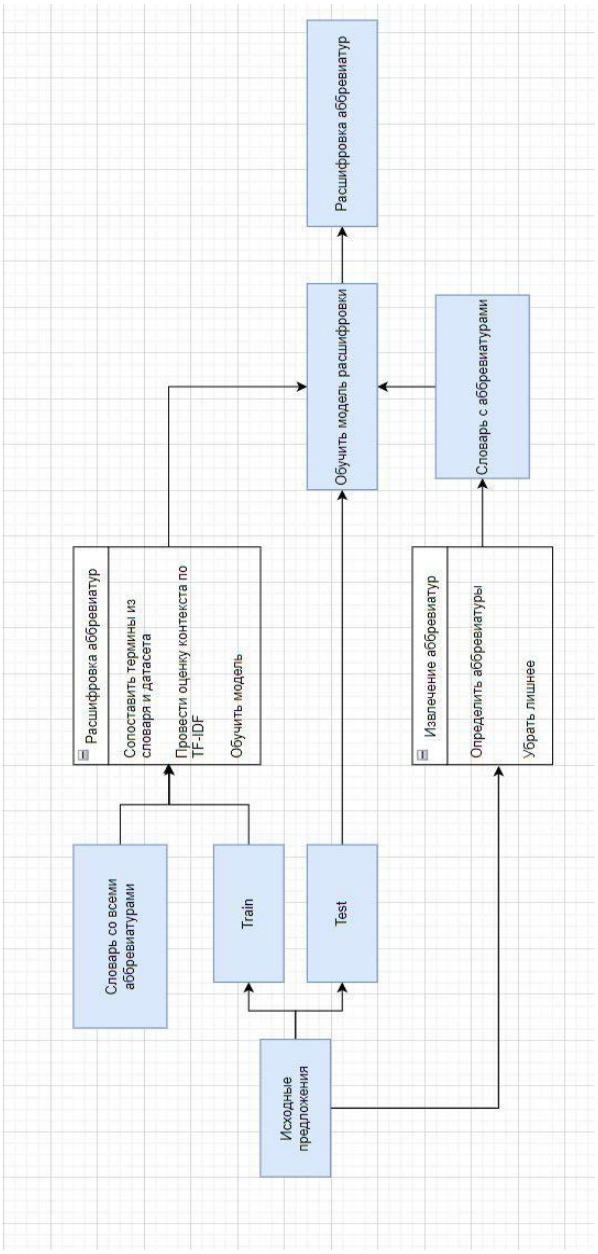


Рис. 1. Общая схема алгоритма

Корпус текстов представляет собой ряд предложений, которые делятся на тестовый и обучающий датасеты, далее из обучающих данных выделяются аббревиатуры и контекст, с помощью словаря, который был взят с ресурса sokr.ru, и меры TF-IDF с шириной контекстного окна 5 расшифровываются аббревиатуры, после обучения модели работа алгоритма проверяется на тестовой выборке, после проверки можно применять к обученной модели уже свои наборы текстов, для расшифровки аббревиатур [5].

Прежде чем использовать словарь, требовалось его преобразовать к нужному виду, так как данные передаются в формате html, нужно было написать парсер, который бы на выходе демонстрировал расшифровки аббревиатур в формате словаря [6]. Парсер был реализован на языке программирования Python с помощью библиотеки BeautifulSoup. После парсинга страницы html были получены различные варианты расшифровок аббревиатур, а также комментарии, которые могут помочь с определением контекста того или иного определения.

Ниже представлен сайт со словарем (рис. 2).

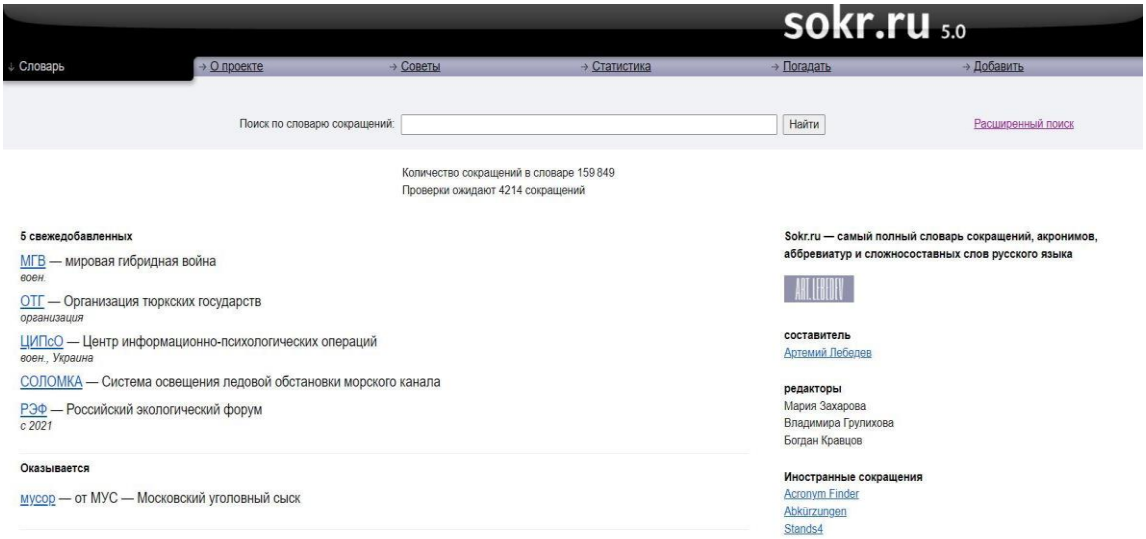


Рис. 2. Сайт со словарем

Рассмотрим подробнее алгоритм, изначально слова представляются в векторном виде для того, чтобы вычислить частоту вхождения для каждого слова, затем получается список наиболее важных слов [7]. Отличие данного метода в том, что захватывается весь список аббревиатур, который есть в словаре, словарь тем временем всегда дополняется неологизмами.

Первым шагом был подготовлен набор текстовых данных, содержащий аббревиатуры, которые нужно расшифровать.

Далее определены для каждой аббревиатуры все возможные варианты ее расшифровки, все варианты расшифровок берутся из онлайн-словаря sokr.ru.

После этого использовались библиотеки для обработки текста, такие как NLTK или SpaCy, для предобработки данных, включая лемматизацию, токенизацию, стемминг, удаление стоп-слов, пунктуации и других ненужных символов [8]. Данный этап необходим для того, чтобы не оценивать различные служебные символы, например предлоги, союзы, частицы и т. д [9].

Следующим шагом было превратить подготовленный текст в численный массив, который будет использоваться для нахождения наиболее подходящих расшифровок. В этом случае воспользуемся методом TfidfVectorizer из библиотеки scikit-learn для получения векторов функций для каждого документа. Создается матрица весов для определения значимости каждого символа в предложении.

И наконец обучается модель машинного обучения на наборе данных, который подготовлен заранее, используя вектора функций, созданные для текстовых данных с аббревиатурами. Получается так, что данные делятся на два типа, на одних обучается сама модель, на других модель тестируется.

Ниже на рисунке 3 представлены примеры предложений, которые подаются для анализа.

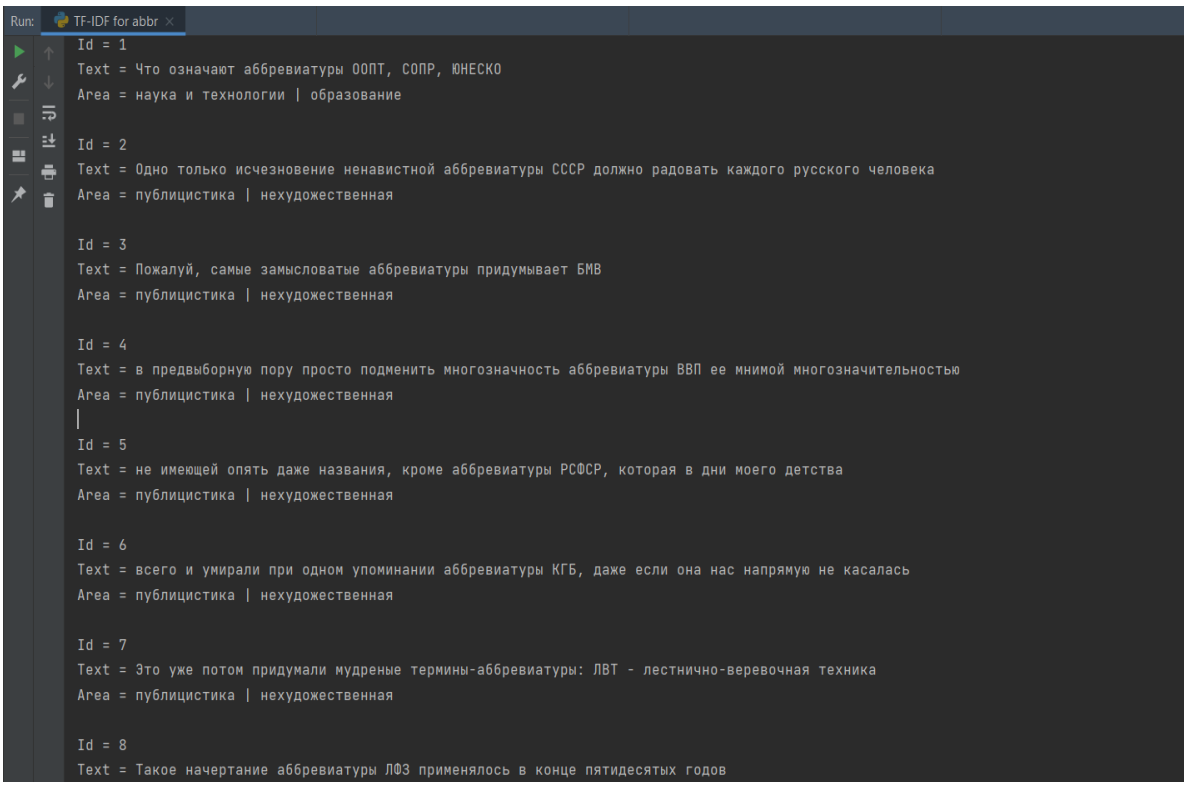


Рис. 3. Примеры предложений

Вычисление точности определения аббревиатур в корне зависит от способа, который используется для расшифровки.

Например, в данной статье использовался онлайн-словарь для расшифровки аббревиатур, а также метод машинного обучения, следовательно точность определения может быть рассчитана на основе того, сколько определений из словаря соответствуют данной аббревиатуре, используемой в конкретном контексте, также точность может быть определена по сравнению результатов, полученных автоматически с правильными ответами (результатами, полученными вручную).

Пример расчета точности определения аббревиатур:

$$Accuracy = \frac{TrueDef}{Abbreviations}, \quad (1)$$

где Accuracy – точность, TrueDef – правильные определения аббревиатур, при сравнении с предполагаемыми правильными расшифровками, Abbreviations – общий список всех аббревиатур.

Посчитаем точность классическим методом, с помощью встроенных функций в Python. Было получено значение точности, которое рассчитывалось по формуле (1) Accuracy = 89.1% что является неплохим результатом по сравнению с алгоритмом, реализованным в библиотеке nltk, который показал Accuracy = 83.6 %. Разница значений получается примерно 0,6 процента. Несмотря на то, что в эффективности библиотека nltk незначительно уступает, она не покрывает все возможные аббревиатуры из выборки, из-за чего в целом снижается точность оценки.

В результате проведенного анализа существующих методов по расшифровке аббревиатур был выбран наиболее оптимальный с точки зрения скорости работы, точности оценки, а также полноты выборки аббревиатур. Был реализован алгоритм с использованием онлайн словаря, статистической меры TF-IDF и контекстного метода для более точной оценки семантической составляющей аббревиатуры, который показал лучший результат решения поставленной задачи по сравнению с существующими.

Литература

1. Шилов И.М. Автоматическое выявление и расшифровка аббревиатур и сокращений в тексте [диссертация из интернета]. Санкт-Петербург: Санкт-Петербургский государственный университет (СПбГУ). 2016. С. 7–9. Режим доступа: https:// nauchkor.ru/ pubs/ avtomaticheskoe- vyyavlenie- i-rasshifrovkaabbreviatur-i-sokrascheniy-v- tekste-587d36365f1be77c40d58984. (дата обращения: 17.02.2023).
2. Словарь сокращений. Режим доступа: https://sokr.ru/. (дата обращения: 17.02.2023).
3. Manning C. Introduction to information retrieval. 2008. С. 114–115.
4. Mushahid Khan, Andriy Miranskyy. String Comparison on a Quantum Computer Using Hamming Distance. 2021. С. 1-19.
5. Левенштейн В.И. Двоичные коды с исправлением выпадений, вставок и замещений символов. Доклады Академий Наук СССР. 1965. С. 845–848.
6. Winkler W.E. String Comparator Metrics and Enhanced Decision Rules in the Fellegi-Sunter Model of Record Linkage, Proceedings of the Section on Survey Research Methods, American Statistical Association. 1990. P. 354–359.
7. Damerau-Levenshtein distance [Электронный ресурс]. URL: https:// en.wikipedia. org/wiki/Damerau%E2%80%93Levenshtein_distance. (дата обращения: 17.02.2023).
8. Bloom D.A. 2000. “Acronyms, abbreviations and initialisms.” BJU International 86(1).
9. Колшанский Г.В. Контекстная семантика. 2002. 126 с.

УДК 004.457

РАЗРАБОТКА ПРОЕКТА АВТОМАТИЗАЦИИ РАЗВЕРТЫВАНИЯ ИНФРАСТРУКТУРЫ ДЛЯ РЕШЕНИЯ ЗАДАЧ МАШИННОГО ОБУЧЕНИЯ

Тихонов Л.А.[1] (студент)
Научный руководитель – преподаватель практики Брюханов К.В.[1]
1 – Университет ИТМО

Аннотация
В этой статье мы рассмотрим проект разработки автоматизации развертывания инфраструктуры для решения задач машинного обучения. Этот проект позволяет объединить несколько шагов развертывания инфраструктуры в единую автоматизированную процедуру. И в результате работы инфраструктуры мы должны достичь сокращения времени работы алгоритмов машинного обучения.
Ключевые слова
MLOps, devops, машинное обучение, автоматизация. spark, hadoop.

В связи с быстрым развитием технологий машинного обучения и анализа данных, значительной частью процесса работы с данными становится их обработка и подготовка к анализу. Для этого необходима инфраструктура, которая позволяет настроить окружение для запуска и тестирования моделей машинного обучения.

В настоящее время возможно развернуть инфраструктуру как на серверах, так и на облачных платформах. Однако, процесс развертывания инфраструктуры вручную может занять много времени, и подвержен ошибкам при выполнении нескольких шагов вручную [1].

Ключевое преимущество конвейеров машинного обучения заключается в автоматизации этапов жизненного цикла модели. Когда становятся доступными новые обучающие данные, должен быть запущен рабочий процесс, который включает проверку данных, предварительную обработку, обучение модели, анализ и развертывание.

Мы наблюдали, как слишком много рабочих групп, занятых анализом данных, вручную выполняют эти шаги. Это обходится очень дорого и часто является источником ошибок. Вот некоторые из преимуществ конвейеров машинного обучения.

1. Возможность сосредоточиться на новых моделях, а не на поддержке существующих.
2. Автоматизированные конвейеры машинного обучения освободят специалистов по обработке данных от необходимости поддерживать существующие модели.
3. Упрощение процесса развертывания: благодаря использованию автоматических инструментов, упрощается процесс развертывания необходимых компонентов, исключается ручная настройка и управление, что позволяет сократить время, затрачиваемое на настройку и ускорить запуск проекта.
4. Улучшение безопасности: автоматическое развертывание позволяет повысить уровень безопасности системы за счет применения стандартизированных методов и защиты от ошибок и проникновений.
5. Уменьшение затрат: автоматизация процессов значительно упрощает работу ведущих специалистов и сокращает их время на выполнение повторяющихся задач. Это позволяет снизить оплату труда и уменьшить затраты компании на поддержание инфраструктуры, что особенно актуально в условиях ожесточенной конкуренции.

В проект разработки включены набор инструментов и технологий MLOps [2], которые позволяют автоматизировать процесс развертывания инфраструктуры, включая настройку серверов, настройку окружения, настройку параметров машинного обучения,

создание окружения для запуска и тестирования моделей машинного обучения, а также управление версиями.

Поскольку объем данных и моделей машинного обучения в управлении много, важно хранить их и работать с ними эффективно и безопасно. Ключом к решению этих проблем является автоматизация, поэтому был выбран стек технологий, состоящий из Hadoop, Spark, Python и bash.

Hadoop это инструмент с HDFS – распределённая файловая система для хранения данных на различных узлах. В неё встроена система дублирования данных, чтобы обеспечить их надёжность и сохранность даже при отказе отдельных серверов. Здесь хранятся неструктурированные данные, извлекать которые можно не просто запросами, но и специальными средствами и инструментами [3].

Spark – это платформа, которая без проблем работает с Hadoop, позволяя взаимодействовать как с данными, так и с методами машинного обучения, а также обеспечивая подготовку и мониторинг данных [4].

Python, стандартный язык сценариев, позволяет подключать Spark и файлы с алгоритмами машинного обучения.

Сценарии Bash также используются для эффективной работы на системном уровне, включая установку программного обеспечения и развертывание инфраструктуры.

Переходя к самой инфраструктуре, первым шагом будет непосредственно установка всех компонентов программного обеспечения на сервера или удаленные хосты.

После того как все установлено и работоспособность проверена, мы можем приступить к работе с данными. Сначала осуществляется сбор и загрузка данных в хранилище Hadoop. Затем данные подготавливаются в хранилище посредством распределения и группировки с последующим формированием выборок данных. Далее за работу берется Spark где он используется для настройки работы с хранилищем и подготовки алгоритмов машинного обучения.

Следующий этап включает в себя обучение модели с использованием данных из Hadoop. После обучения модель проходит тестирование с использованием тестов написанных на Python и метрик используемых в Spark. Если модель соответствует требованиям, она переходит к этапу публикации; в противном случае она возвращается к этапу настройки, где текущие параметры подвергаются изменениям с помощью скриптов (рис. 1).

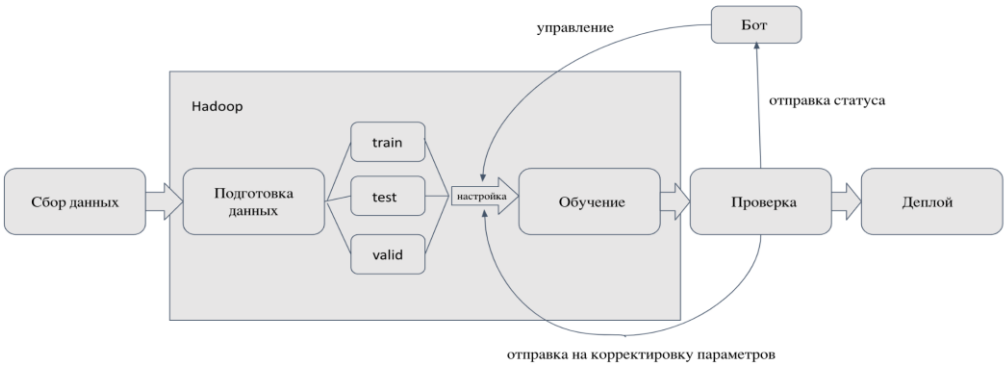


Рис. 1. Схема разработанной инфраструктуры

На протяжении всего процесса обучения и проверки дополнительный мониторинг осуществляется Телеграмм ботом, который присылает обновления о текущем состоянии модели и статусе проверок.

Бот написан на языке Python и библиотеке aiogram. Он осуществляет подключение к нашей модели которая находится в конвейере и считывает основные параметры, статусы. Эти значения передаются клиенту посредством API запросов (рис. 2).

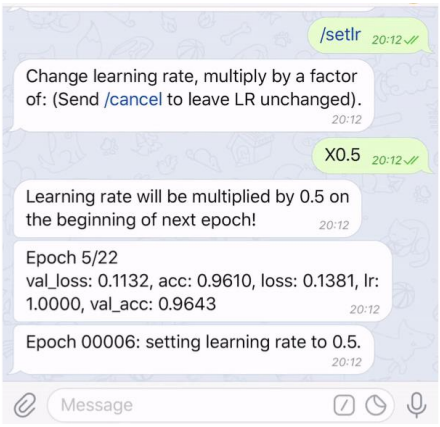


Рис. 2. Настройка параметров

Отличительной особенностью этого бота является возможность вмешиваться в процессы, которые происходят в инфраструктуре. А именно изменить параметры обучения, остановить обучение, вывести графики и наблюдать за качеством обучения посредством сообщений со статистикой (рис. 3).

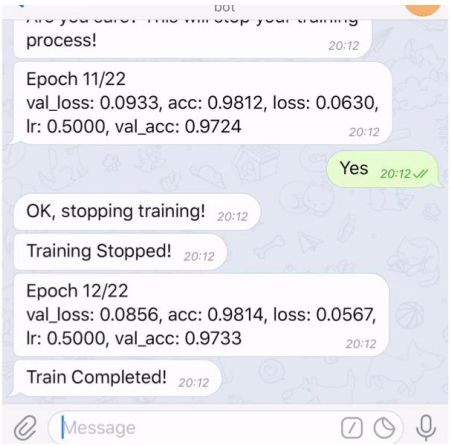


Рис. 3. Остановка работы

Перейдем к заключительному этапу. Это тестирование на реальных данных.
Для проведения оценочного тестирования был взят датасет состоящий из 1000 объектов который загрузился за 30 секунд при использовании HDFS, без нее загрузка того же самого объема данных заняла около двух минут.

После того с данными разобрались приступил к работе алгоритм машинного обучения. Результаты которого представлены ниже.

```
До:

CPU times: user 14.1 s, sys: 1.56 s, total: 15.7 s
Wall time: 1min 2s


После:

CPU times: user 11.6 s, sys: 1.55 s, total: 13.1 s
Wall time: 58.6 s
```

Рис. 4. Результаты работы

Алгоритмы машинного обучения на инфраструктуре Spark работают значительно быстрее за счет использования оперативной памяти вместо дискового хранилища, а также за счет использования интеллектуальной оптимизации запросов, такой как трансляция запросов в более эффективный формат выполнения.

Без инфраструктуры машинное обучение может работать быстрее или медленнее в зависимости от сложности алгоритма и доступных вычислительных ресурсов. Обычно скорость работы на одной машине бывает быстрее, чем на кластере из-за отсутствия задержек ввода-вывода и обработки запросов между узлами. Однако, машинное обучение без инфраструктуры ограничено доступным объемом памяти и процессорной мощности, что может быть недостаточно для обработки больших объемов данных или сложных алгоритмов.

В результате применения проекта для развертывания инфраструктуры для решения задач машинного обучения удалось сократить затраты времени и уменьшить риск ошибки при выполнении нескольких шагов ручного развертывания. Кроме того, проект позволяет быстро расширять инфраструктуру для работы с данными и моделями машинного обучения, что важно для исследовательских команд и малых бизнесов. А также следует отметить, что использование автоматизации и технологий, таких как Hadoop, Spark и Python, оказалось необходимым для эффективного управления большими объемами данных и моделями машинного обучения. Используя эти инструменты, компании могут оптимизировать свои процессы и обеспечить безопасность и эффективность своих систем управления данными.

Литература

1. Брюханов К.В., Петров В.В., Авксентьева Е.Ю. Анализ существующих проблем в организации непрерывной доставки и развертывании программного продукта // Современная наука: актуальные проблемы теории и практики. Серия: Естественные и технические науки. 2020. № 5. С. 79-82.
2. Руководство по MLOps. [Электронный ресурс]. Режим доступа: https://medium.com/sciforce/mlops-comprehensive-beginners-guide-c235c77f407f (дата обращения: 03.02.2023).
3. Документация по Hadoop. [Электронный ресурс]. Режим доступа: https://hadoop.apache.org (дата обращения: 28.01.2023).
4. Документация по Spark. [Электронный ресурс]. Режим доступа: https://spark.apache.org (дата обращения: 20.01.2023).

УДК 004.891

ПРОГНОЗИРОВАНИЕ ТРЕНДОВ В РАЗРЕЖЕННЫХ ВРЕМЕННЫХ РЯДАХ С ПРИМЕНЕНИЕМ НЕЙРОННЫХ СЕТЕЙ

Улизько М.В.[1] (аспирант), **Добренко Д.А.**[2] (сотрудник), **Голубев А.А.**[3] (сотрудник)
Научный руководитель –к.т.н., доцент Гусарова Н.Ф.
1 – Университет ИТМО
2 – ГУП «Петербургский Метрополитен»
3 – ООО «МЕТЭКО»
e-mail: m@ulizko.ru

Аннотация
В данной работе рассмотрено практическое применение байесовских нейронных сетей для решения задачи прогнозирования в условиях малых выборок. Потребность в прогнозе на малой количестве данных имеется во всех сферах научных исследований. При этом, традиционные модели и инструменты показывают результат, недостаточный для анализа. В работе будет рассмотрен ряд возможных причин низкой производительности нейронных сетей в задаче прогнозирования временных рядов.

Ключевые слова
Временные ряды, нейронные сети, прогнозирование, малые выборки, анализ данных.

Прогнозирование малых выборок сегодня используется широко используется в научной деятельности. Однако, в рассматриваемых работах авторы сводят задачу прогнозирования к линейной регрессии. Классические модели, такие как ARIMA, доказали свою эффективность, но не могут быть использованы на малом количестве данных. В том числе, стоит проблема предобработки данных и перестройки модели под новые данные [1, 4].

Временной ряд – собранный в разные моменты времени статистический материал о значении каких-либо параметров исследуемого процесса.

Каждый временной ряд содержит два элемента:

1. Значения времени.
2. Соответствующие им значения уровней ряда.

Прогнозирование временных рядов – это метод прогнозирования будущих событий статистического материала.

В классических задачах анализа данных считается, что различные наблюдения независимы. В нашем случае, при прогнозировании предполагается, что прошлые значения ряда содержат информацию о его поведении в будущем.

Для прогноза разреженных временных рядов предлагается использовать Байесовскую нейронную сеть. Байесовская сеть – модель нейронной сети, где связь между парой нейронов задается вероятностью по формуле Байеса [3]. На рисунке изображено сравнение байесовской и обычной нейронной сети.

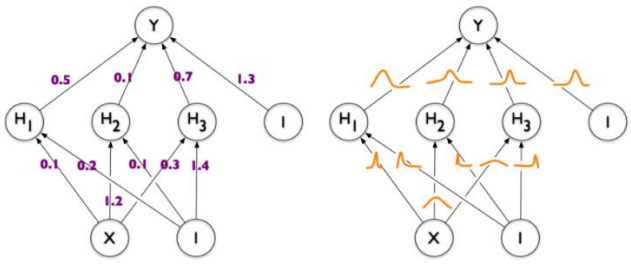


Рисунок. Модель байесовской нейронной сети

Байесовские нейронные сети используют вероятностные слои, которые фиксируют неопределенность в отношении весов и активаций, и обучаются с использованием байесовского вывода. Байесовские сети относятся к группе алгоритмов, описывающих формализованные взаимосвязи [2]. В рамках интеллектуального анализа данных с помощью формулы Байеса можно более точно рассчитать вероятность, используя как ранее известную информацию, так и данные новых наблюдений.

Рассмотрим главные преимущества Байесовской сети:

1. Значения неизвестных переменных обрабатываются значительно быстрее при наличии заданных зависимостей в модели между переменными.
2. Байесовские нейронные сети являются легко интерпретируемыми на этапе прогностического моделирования.

Для оценки качества работы и сравнения производительности с другими моделями был использован датасет с Kaggle «Student Grade Prediction». В нем собраны данные об учениках двух португальских школах.

В качестве целевой переменной предлагается использовать G3 – оценку за финальный экзамен. Был произведен предварительный анализ данных и обработка датасета. Исходный набор данных содержал 398 записей. Далее набор данных был разрежен. После отбора признаков датасет содержит следующий признаки:

1. G3 - финальный экзамен (numbers: from 0 to 20, output target).
2. Failures – количество неудовлетворительных оценок (numeric: n if 1<=n<3, else 4).
3. Medu - образование матери (numeric: 0 - none, 1 - primary education (4th grade), 2 5th to 9th grade, 3 secondary education or 4 higher education).
4. Higher – желание получить высшее образование (binary: yes or no).

Крайним значением для получения релевантных результатов стало 39 записей. В качестве оценки было произведено тестирование на базовых алгоритмах а затем на Байесовской сети. В качестве оценки использовались метрики MAE, RMSE и log like. Обучающая и тестовая выборка были разделены в соотношении 80-20.

Набор данных изначально содержал 398 записей. В качестве оценки было произведено тестирование на базовых алгоритмах а затем на Байесовской сети. В таблице 1 представлены результаты.

Таблица 1

Оценка работы алгоритмов на наборе данных с 398 записями

Алгоритм	MAE	RMSE	loglike
Linear Regression	3.485	4.432	-
Random Forest	3.804	4.731	-
SVM	3.549	4.581	-
Gradient Boosted	3.571	4.500	-
Baseline	3.789	4.825	-
Bayes Neural Network	1.102	1.257	-1.127

Как видно из таблицы 1, байесовская нейронная сеть показала лучший результат. Далее набор данных был разрежен. Крайним значением для получения релевантных результатов стало 39 записей. Байесовская сеть показала вновь показала отличный результат, когда базовые алгоритмы показали низкую точность. Для устранения

проблемы переобучения использовались методы регуляризации L1 и L2 [3]. Результаты продемонстрированы в таблице 2.

Таблица 2

Оценка работы алгоритмов на наборе данных с 39 записями

Алгоритм	MAE	RMSE	loglike
Linear Regression	9.17	11.475	-
Random Forest	4.23	5.744	-
SVM	5.911	6.761	-
Gradient Boosted	6.2	7.362	-
Baseline	4.521	5.763	-
Bayes Neural Network	1.244	1.346	-1.500

На тренировочном датасете была применена и настроена Байесовская нейронная сеть. Размерность датасета была уменьшена в 10 раз, при этом потеря в качестве прогнозирования составила 7%. Модель обучена и настроена совместно с ГУП "Петербургский Метрополитен" и ООО «МЕТЭКО» для дальнейшего использования аналогичной модели для предприятий после того, как будет настроена система для сбора данных для анализа.

Литература

1. Chimmula, Vinay Kumar Reddy, and Lei Zhang. "Time series forecasting of COVID-19 transmission in Canada using LSTM networks." – Chaos, Solitons & Fractals 135. 2020: 109864.
2. Кустицкая Т. А. (2019, сентябрь). Прогнозирование успешности обучения студента с помощью байесовской сети. – Информатизация образования и методика электронного обучения: Материалы III Междунар. науч. конф. Красноярск: Сибирский федеральный университет. pp. 257-262.
3. Wang R., Ye S., Li K. and Kwong S. 2021. Bayesian network-based label correlation analysis for multi-label classifier chain. Information Sciences 554. pp. 256-275.
4. Kareem S.W., Alyousuf F.Q.A., Ahmad K., Hawezi R. and Awla H.Q. 2022. Structure Learning of Bayesian Network: A Review – QALAAI ZANIST JOURNAL, 7 pp. 956-975.

УДК 004.93’12

НАБЛЮДЕНИЕ ЗА ПАЦИЕНТОМ С ПРИМЕНЕНИЕМ ТЕХНОЛОГИЙ ОКРУЖАЮЩЕГО ИНТЕЛЛЕКТА И УСТРОЙСТВ ИНТЕРНЕТА ВЕЩЕЙ

Фам Т.А.[1] (аспирант), **Субботин А.Н.**[2] (аспирант)
Научный руководитель – к.т.н., доцент Жукова Н.А.[1]
1 – Университет ИТМО, Санкт-Петербург
2 – Университет СПбГЭТУ «ЛЭТИ»
e-mail: ptanh@vku.udn.vn

Аннотация
Применение технологий Интернета Вещей (Internet of Things - IoT) в области здравоохранения способствует улучшению повседневной работы в больницах, особенно в сфере наблюдения и ухода за пациентами. В последние годы наблюдается постоянно возрастающий интерес к технологиям создания интеллектуальных палат, что обусловлено их высокой востребованностью в медицинских учреждениях, также при уходе за пациентами на дому. Использование существующих устройств и систем IoT для построения умных палат в рамках создания систем повседневного интеллекта (Ambient Intelligence) - практичное, недорогое и быстрое решение, отвечающее современным требованиям медицины. Их построение и использование не требует установки нового оборудования, позволяет экономить финансовые средства.
Ключевые слова
Системы окружающего интеллекта, интеллектуальная палата, распознавание лиц, эмоций, позы.

Введение

С постоянным развитием аппаратных и программных технологий, когда почти все устройства становится возможным подключить к локальным и глобальным сетям различными способами (Интернет, Wi-Fi, инфракрасный порт... и т.д.), расширяется область применения IoT, все большее число устройств оказываются связанными между собой. Internet of Medical Things (IoMT) представляет собой технологию, предполагающую применение Интернета Вещей в области здравоохранения. Технологии IoT в области здравоохранения способствуют улучшению повседневной работы в больницах, особенно в сфере ухода за больными. В настоящее время "интеллектуальная палата" является не только тенденцией, но и необходимой потребностью в медицинских учреждениях, а также потребностью при уходе за больными вне стационара. После операции за пациентами необходимо регулярно наблюдать, оценивать состояние здоровья, болевые ощущения, а также контролировать непрерывность их сна. Кроме того, необходимыми требованиями, предъявляемыми к интеллектуальной палате, являются: мониторинг и обнаружение травм в результате падений, помощь пациентам в общении с медицинским персоналом. Однако, использование современного, специализированного, интегрированного, многофункционального оборудования не представляется возможным для многих больниц по причине его высокой стоимости и сложности внедрения. Следовательно, применение существующих устройств и построение систем на основе IoT является практичным, недорогим и быстрым решением для построения умной палаты.

Технология окружающего интеллекта и интеллектуальная палата

В интеллектуальной палате с использованием установленных в ней медицинских приборов осуществляется сбор данных о состоянии пациентов [1]. Благодаря этому можно повышать качество ухода за больными, предоставляя интеллектуальные коммуникационные решения для работников сферы здравоохранения или других организаций сопровождающего профиля. Некоторые больницы и медицинские организации по всему миру пытаются превратить свои медицинские учреждения в

интеллектуальные с применением новейших систем видеонаблюдения, регистрации событий, путем реконфигурации систем и модернизации устаревших устройств [2].

За последние годы системам медицинской поддержки пациентов уделяется много внимания и исследований [3]. Эти системы в основном предназначены для отслеживания и обнаружения падений, ведения журналов активности [4]. В ряде случаев используются носимые специализированные устройства для наблюдения за пациентами. Существующие системы для мониторинга и поддержки ухода за пациентами были проанализированы и оценены их возможности, что позволило определить необходимые и полезные функции этих систем, которые могут быть реализованы при создании интеллектуальной палаты для пациента [5].

Предлагаемая система интеллектуальной палаты

Разработка системы для интеллектуального наблюдения в палате за пациентом предполагает применение камер видеонаблюдения и смартфона. Интеллектуальная палата может применяться в медицинских организациях, в которых нет дорогостоящего оборудования, для семейного наблюдения, когда нет необходимости в использовании профессионального оборудования.

Описание структуры системы интеллектуальной палаты

Структура системы представлена на рисунке 1. Она включает в себя систему общих камер, подключенную к серверу с помощью существующей инфраструктуры системы безопасности, и систему персональных камер, подключенную к серверу с помощью телефонной сети или Wi-Fi соединения. Обе системы камер подключаются к общему серверу управления интеллектуальной палатой.

Описание диаграммы информационного потока

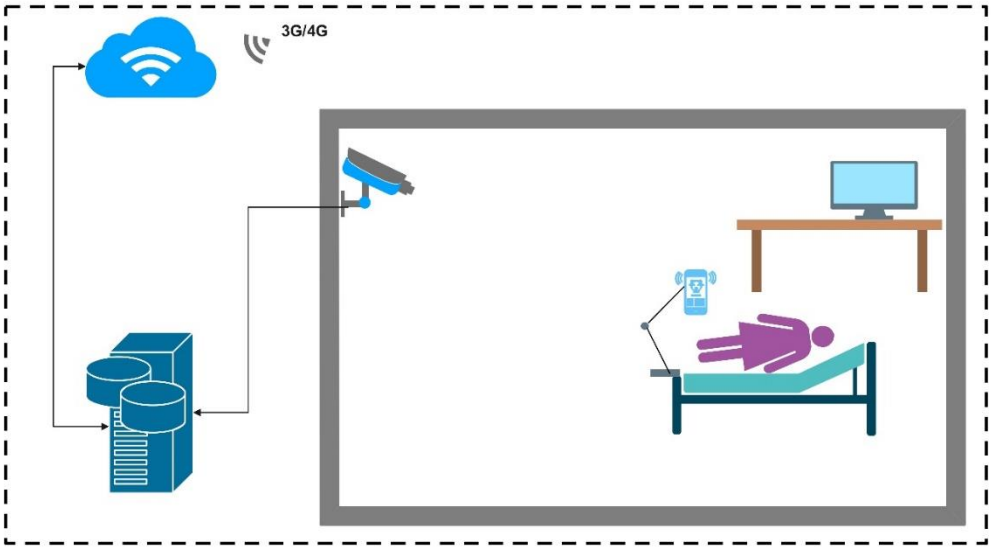


Рис. 1. Структура интеллектуальной палаты с AmIs-IoT

Диаграмма потока данных показана на рисунке 2, она состоит из двух входных потоков данных. Во-первых, это данные с общих камер, которые являются элементами системы безопасности. Они устанавливаются на высоте, которая позволяет камере охватить всю комнату, не заслоняя обзор. Полученные данные передаются непосредственно в базу данных системы. После обработки изображений, в частности, распознавания человека, его лица, позы, результаты передаются в приложения, установленные на смартфоне пользователя.

Второй поток данных поступает с персональных камер, используются фронтальные камеры смартфонов, которые размещаются вблизи места нахождения пользователя, могут захватить изображение его лица.

Из-за ограничения памяти смартфона, ограничения интернет-соединения и с учетом способности смартфонов обрабатывать данные, при получении изображения система немедленно выполняет функцию обнаружения лица и определяет по нему эмоциональное состояние человека. Результаты отображаются в интерфейсе приложения, а также отправляются на сервер для хранения с использованием интернет-соединения смартфона.

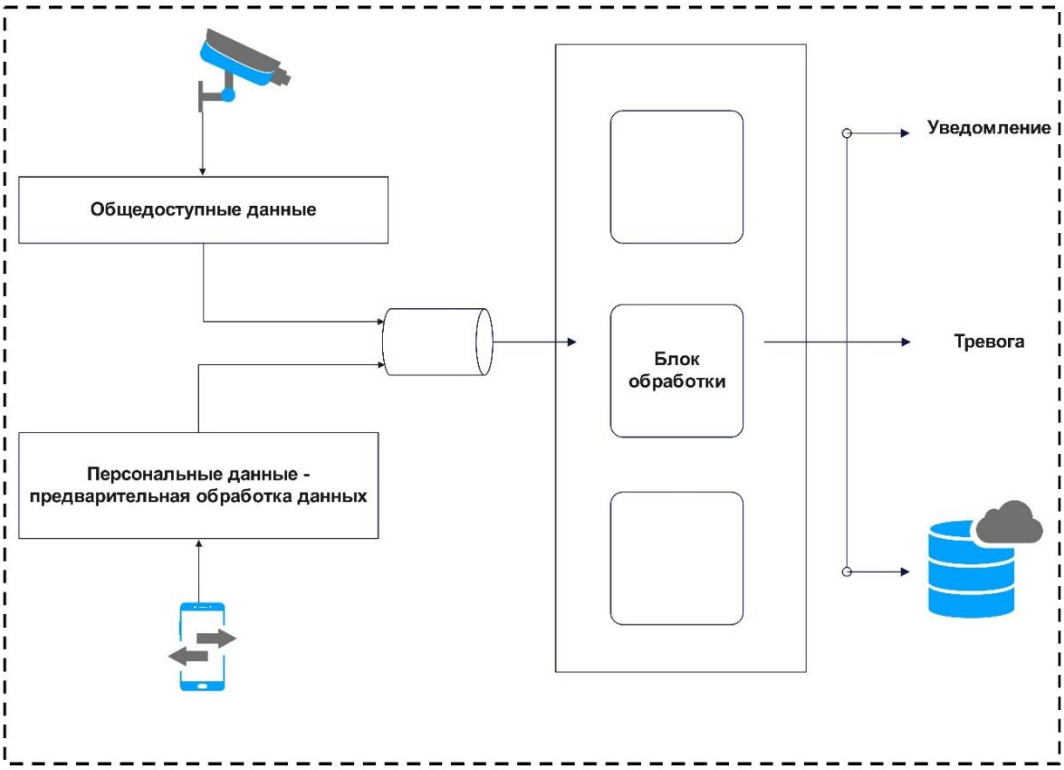


Рис. 2. Диаграмм потоков данных системы интеллектуальной палаты с AmIs-IoT

Описание функций системы интеллектуальной палаты:

– функция обнаружения лица: для сокращения вычислений и увеличения скорости обработки система использует методы на основе факторной модели лица для обнаружения лица человека на изображении.

– функция определения эмоционального состояния человека: после обнаружения лица выполняется функция определения эмоционального состояния.

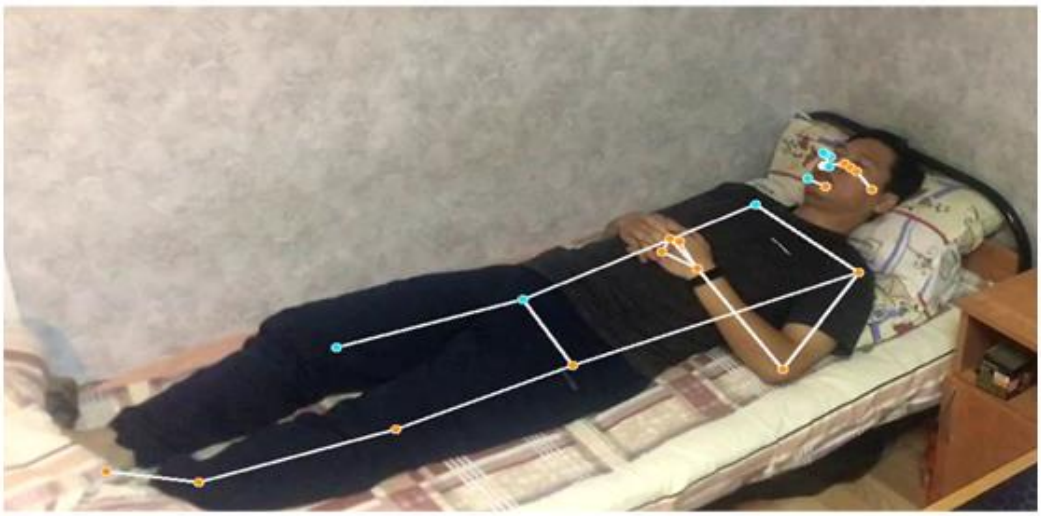
– функция определения положения скелета человека: при получении данных с камер безопасности, сначала выполняется функция определения скелета человека. Для определения положения скелета система использует решение MediaPipe Pose. Пример распознанного по изображению положения скелета показан на рисунке 3.

– функция распознавание позы человека: после завершения процесса обнаружения человеческого скелета обрезанное изображение отправляется на сервер распознавания для распознавания позы. Для обнаружения были выбраны четыре позы: человек сидит, человек стоит, человек лежит и человек отдыхает/спит. Чтобы определять позы, разработана и обучена пользовательская модель с помощью Image AI. Для обучения специализированной модели использовалось около 500 изображений для каждой позы. Для тестирования каждой позы использовался набор данных из 100 изображений для каждой позы.

Выводы

Предлагаемая система предполагает использование существующих устройств (камер наблюдения, подключенных к интернету, серверных систем получения данных, камер смартфонов и пр.), построение системы не требует установки дорогостоящего оборудования (серверов видеообработки, высококачественных видеокамер и пр.). Система позволяет использовать возможности различных существующих устройств. Камера смартфона может выполнять обработку данных и имеет ограничения на передачу данных по сети, поэтому сначала данные обрабатываются (предварительная обработка), а затем результаты их обработки передаются в центральную систему, которая размещается на сервере в облаке. Камеры, обеспечивающие безопасность, не могут выполнять предварительную обработку данных, но у них хорошее соединение с интернетом и они имеют возможность передавать данные на сервер для их обработки. Система использует возможности распространенных устройств и средств обработки данных для создания умной палаты с полным набором функций, которые определяются современными требованиями к умной палате. Такая система может быть применена в медицинских учреждениях в стационаре и амбулаторно на дому (рис. 5).

a)



b)

Рис. 3. Распознавание скелета человека и определение позы человека по изображению (a) Входное изображение с общей камеры, (b) Результаты распознавания скелета человека

Литература

1. Cook D.J., Augusto J.C., Jakkula V.R. Ambient intelligence: Technologies, applications, and opportunities //Pervasive and mobile computing. 2009. Т. 5. №. 4. С. 277-298.
2. Haque A., Milstein A., Fei-Fei L. Illuminating the dark spaces of healthcare with ambient intelligence //Nature. 2020. Т. 585. №. 7824. С. 193-202.
3. Baker C.R. A system architecture for ambient intelligent environments: дис. Master's thesis, December, EECS Department, University of California,.Berkeley. 2006.
4. Ariunbold Y., Brito S., Leong A. FallDetectNet: A Computer Vision Platform for Fall Detection.
5. Igual R., Medrano C., Plaza I. Challenges, issues and trends in fall detection systems //Biomedical engineering online. 2013. Т. 12. №. 1. С. 66.

УДК 004.624

АНАЛИЗ ВЗАИМОДЕЙСТВИЯ РАСПРЕДЕЛЕННЫХ СЕРВИСОВ С ИСПОЛЬЗОВАНИЕМ ТЕЛЕМЕТРИИ

Фролов М.А.[1] (студент)
Научный руководитель – PhD, доцент Аксенов В.Е.[1]
1 – Университет ИТМО
e-mail: mk17.ru@mail.ru

Аннотация
В работе проведен анализ взаимодействия распределенных сервисов Яндекса с использованием данных трассировок. Для этого были применены методы обработки больших данных mapreduce На основе существующих алгоритмов анализа данных предложены и реализованы в виде прототипов модели вычислительных сетей для анализа больших данных.
Ключевые слова
Трассировки, распределенные системы, большие данные, анализ производительности и эффективности.

При работе с распределенными системами большую сложность доставляет поиск проблем во взаимодействии их компонентов. Примерами таких проблем могут быть ошибки и отказы сервисов, их перегруженность и большая задержка ответов. Каждый такой сбой может нести серьезные убытки для компании, поэтому сейчас для их предотвращения и быстрого исправления требуется специальный отдел менеджеров по чрезвычайным ситуациям (рис. 1).

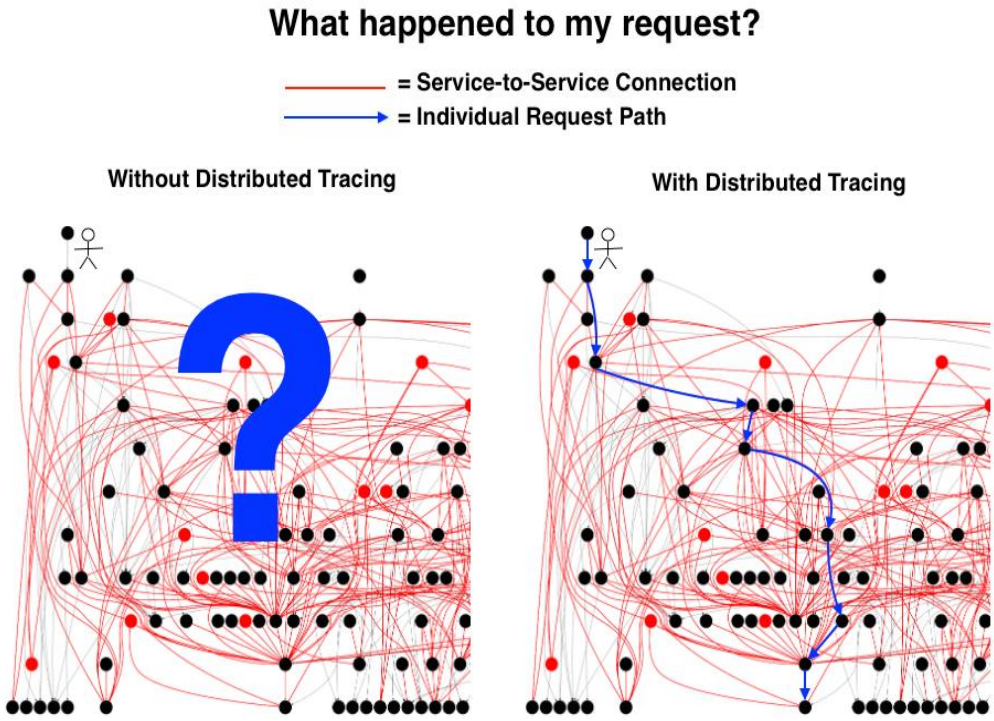


Рис. 1. Поиск в распределенных системах

В компании Яндекс была создана система мониторинга *Juggler*, позволяющая быстро реагировать на проблемы, предупреждать сбои и получать уведомления при критическом изменении важных для данного сервиса метрик.

Для анализа логов и трассировок в России и за рубежом активно применяется сервис *Jaeger*, предоставляющий функционал работы с трассировками и логами: поиск

и просмотр информации о цепочке вызовов функции, о времени, ошибках и параметрах вызова.

Однако, несмотря на использование этих решений, разработчики часто сталкиваются со сложностями в поиске проблем взаимодействия сервисов, в особенности при рассмотрении ошибок, происходящих в *межсервисном* взаимодействии. У разработчиков возникает запрос на динамический анализ работы системы, например рассмотреть ситуацию *до* и *во время* поломки [2, 3].

Предлагаемое мною решение – новый подход к анализу взаимодействия различных сервисов, где они представляются в виде графа сервисов и операций с дополнительной статистикой по каждой операции, что помогает быстро находить сбои и наглядно наблюдать изменения во взаимодействии нескольких сервисов в разные промежутки времени.

Построение графа взаимодействия сервисов состоит из трех основных этапов.

На первом этапе производится обработка и извлечения информации из трассировок и логов сервиса. Данные представляются в стандартном формате распределенного трейсинга и мониторинга OpenTelemetry, в котором каждый вызов функции это один спан с информацией о времени вызова, названии сервиса, информации о родительском спане, информации об успешности вызова операции. Из-за большого количества данных о трассировках они представлены в Protobuf формате, позволяющем эффективно их передавать и хранить.

Однако для того, чтобы воспользоваться данными нужны инструменты для анализа больших данных, большой объем хранилища, а также методы для извлечения полезной информации и вывода ее в удобном для пользователей виде. Данными возможностями обладает система Yandex Table (аналог Hadoop), которая позволяет эффективно обрабатывать большие объемы данных. В результате создается таблица с основной информацией об одной операции взаимодействия двух сервисов: названия сервисов, название проекта, время начала/конца операции, код ответа (рис. 2).

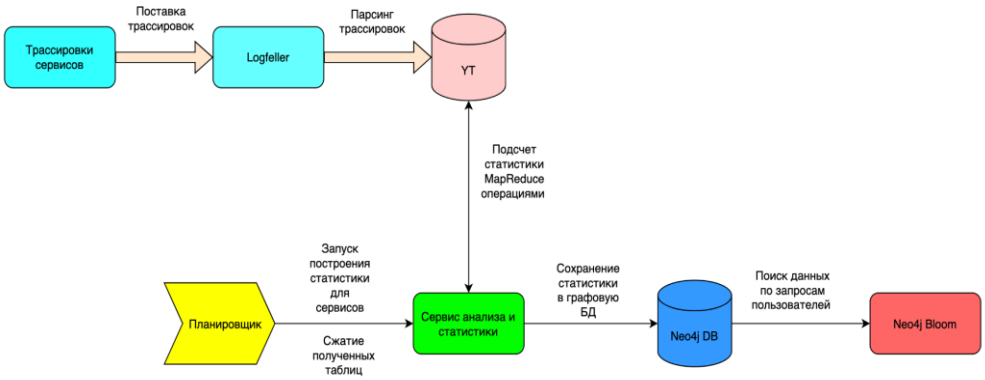


Рис. 2. Организация работы сервиса

Вторым этапом является запуск задачи по подсчету статистики для данного сервиса в интервале времени. Для этого планировщик раз в пять минут запускает MapReduce операцию, получая на выходе несколько таблиц с данными о каждой операции: количество, процент ошибок, статистику по видам ошибок, квантили для времени отклика 90/95/99, количество операций в секунду и другие. Данная таблица прекрасно подходит для подтверждения тех случаев, которые будут найдены на третьем этапе (рис. 3).

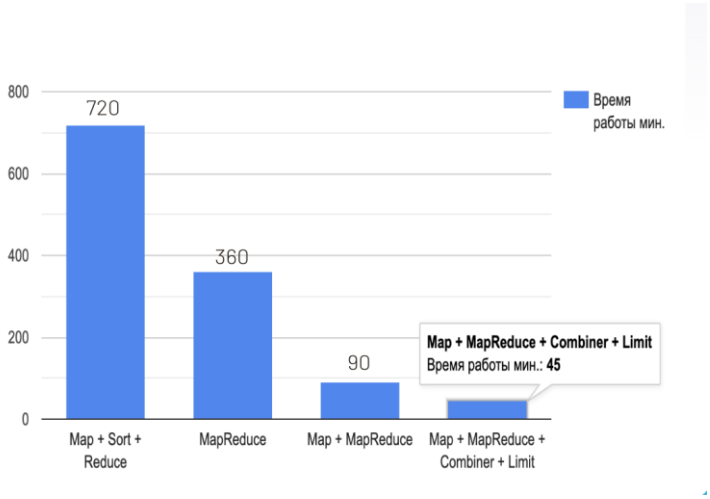


Рис. 3. Оптимизация MapReduce операций

Третий этап – построение графического интерфейса. Представим наши данные в виде графа, где вершинами будут сервисы, а ребрами операции из статистики, рассчитанной на втором этапе. Наиболее удачным хранилищем для таких данных является графовая база данных Neo4j, которая не только позволяет эффективно хранить данные, но также осуществлять отображения данных в виде графа (рис. 4) [1].

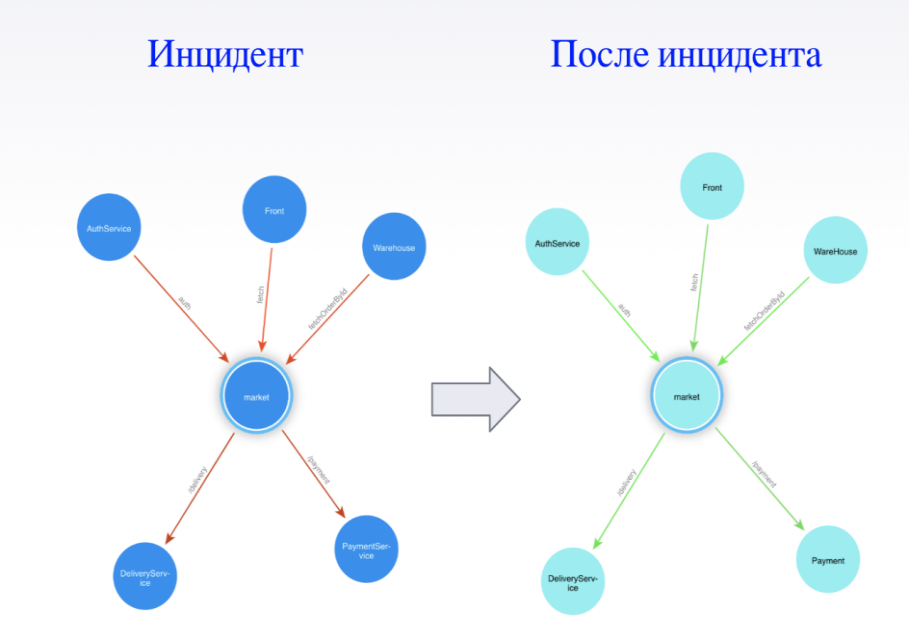


Рис. 4. Графы операций для сервисов до/после инцидента

Графический интерфейс позволяет отображать 3 вида графов:

1) сервисы и операции для одной трассировки;
2) взаимодействие сервисов по каждой операции;
3) граф сервисов, где операции объединены в одно ребро.

Есть классификация графов по тем данным, которые мы отображаем:

1) ошибки и их виды 4xx, 5xx;
2) квантили времени.

Цветом на графе отображается одна из статистик, так, например, можно отобразить операции с большим процентом ошибок красным, а операции с малым процентом ошибок зеленым (рис. 5).

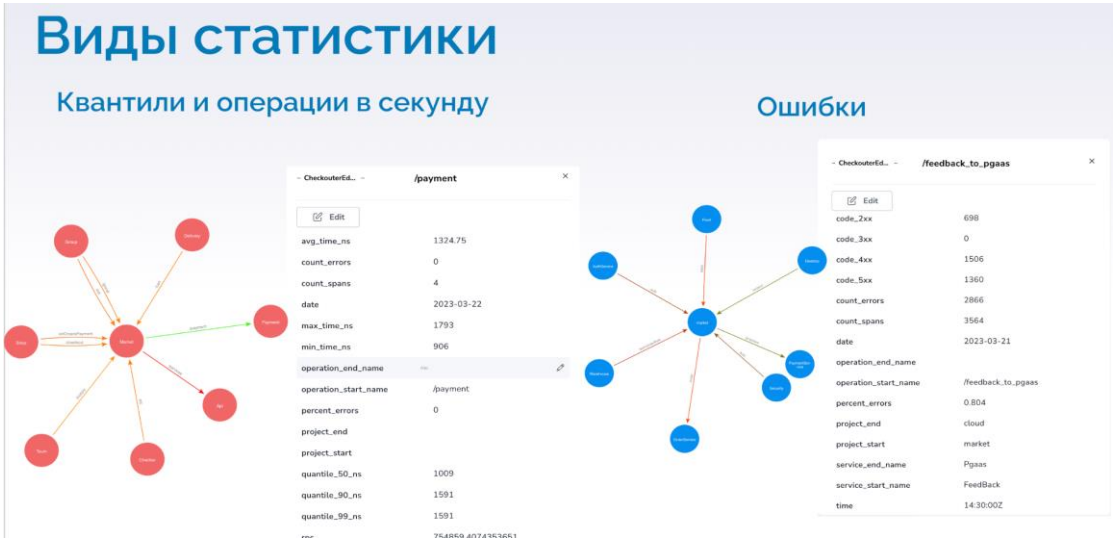


Рис. 5. Виды статистики

Данное разделение позволяет структурировать тот большой объем данных, который содержится в трассировках, и представить его в читаемом виде. Так как данные хранятся в графовой базе данных, можно быстро и эффективно искать и преобразовывать данные, и у пользователей появляется возможность самим настраивать метрики для своего графа.

Основные варианты использования данного решения:

1. Построение схемы взаимодействия сервисов.
2. Отслеживание проблем во взаимодействии сервисов.
3. Автоматическое определение критического пути.
4. Сравнительный анализ изменения взаимодействия двух сервисов в разные промежутки времени (в момент/до сбоя или до/после выхода новой версии приложения: сравнить на каких операциях процент ошибок и время стало больше).
5. Нахождение всех операций и сервисов с заданными ограничениями (например, много ошибок или большое время работы).
6. Сбор информации о взаимодействии сервисов и ее агрегация (время ответа, процент ошибок, количество байт, операций в секунду, квантили для времени отклика).

С помощью представленного решения были найдены и проанализированы несколько проблем взаимодействия сервисов. Одним из примеров использования данного сервиса было нахождение большого количества ошибок во взаимодействии фронтового и бэкендового сервисов. Проблема заключалась в том, что фронтенд обращался к тем сервисам бэкенда, с которыми у него не должно было быть взаимодействия в данном пользовательском сценарии. В результате были заведены задачи в соответствующие команды, и проблема была устранена.

Литература

1. Формат данных Opentelemetry // Opentelemetry URL: https://opentelemetry.io/ (дата обращения: 05.02.2023).
2. Графовая база данных Neo4j // Neo4j URL: https://neo4j.com/ (дата обращения: 01.01.2022).
3. Остин Паркер Распределенные трассировки на практике. 2020. С. 45-67.

УДК 004.89

ОПТИМИЗАЦИЯ НЕЙРОСЕТЕВЫХ АЛГОРИТМОВ ПУТЕМ АВТОМАТИЗАЦИИ ВЫЧИСЛЕНИЙ ПОКАЗАТЕЛЕЙ ГРАДИЕНТНОГО СПУСКА И ФУНКЦИИ АКТИВАЦИИ

Черников К.М.[1] (школьник)
Научный руководитель – старший преподаватель Сухаренко Д.В.[1,2]
1 – Самарский региональный центр для одаренных детей
2 – ФГБОУ ВО «Самарский государственный социально-педагогический университет»
e-mail: kirill242006@yandex.ru, sukharenko@srcod.ru

Аннотация
В работе были рассмотрены существующие алгоритмы прикладных решений, реализованных с помощью нейронных сетей, проведен анализ ключевых параметров нейронных сетей (НС), реализована собственная парадигма оптимизации НС, протестирован оригинальный подход к применению активационной функции, проведено тестирование подхода на различных данных (бинарных и небинарных), в том числе на аномальных.

Ключевые слова
Нейронная сеть, активационная функция, этапы обучения алгоритма, оптимизация.

Искусственная нейронная сеть представляет собой систему нейронов (простых процессоров), связанных и взаимодействующих друг с другом. Их синаптическая связь может задаваться некоторой формулой, зависящей от значений простейших нейронов в ней, которые, в свою очередь, зависят от активационной функции. Соответственно одной из основных задач оптимизации НС является выбор активационной функции.

Функция активации нормализует значения нейрона относительно типов входных и выходных данных. Также она нужна для введения реальных свойств в нейронные сети. В искусственно созданных нейронных сетях функции активации различны [1]. Активационные функции используют на каждом слое: входном, скрытым и выходном. Для каждого слоя есть своя функция, производная от которой показывает градиентный спуск для каждого слоя отдельно. В зависимости от слоя и его назначения выбирается оптимальная функция активации. Сначала решается, каким будет значение активации, а затем используется функция, наиболее подходящая для подобранных значений. Необходимо нормализовать значение каждого нейрона в слое относительно нейронов в прошлом слое, чтобы после подать корректный ответ дальше [2]. Если не использовать функцию активации, то вся нейронная сеть сведётся к линейной регрессии, имеющую ограниченный характер, что не всегда удовлетворяет условиям применения НС [3].

Данная работа основана на принципе получения активационной функции путем интегрирования оптимальной функции градиентного спуска;

Решение было разделено на несколько модулей:

- вычисление наименьшей возможной ошибки для данной сети путем обучения сети;
- сравнительный анализ градиентного спуска при различных функциях относительно минимальной ошибки;
- аппроксимация производных различных вариантов активационных функций с целью приблизиться к значению динамики ошибки;
- вывод оптимальной относительно точности, времени и количества итераций активационной функции для каждого слоя, исходя из градиентного спуска.

При рассмотрении некоторых вариантов активационных функций (рис. 1), реализованы три различные архитектуры построения с применением подобных функций и протестированы на одинаковых данных (таблица 1) [4].

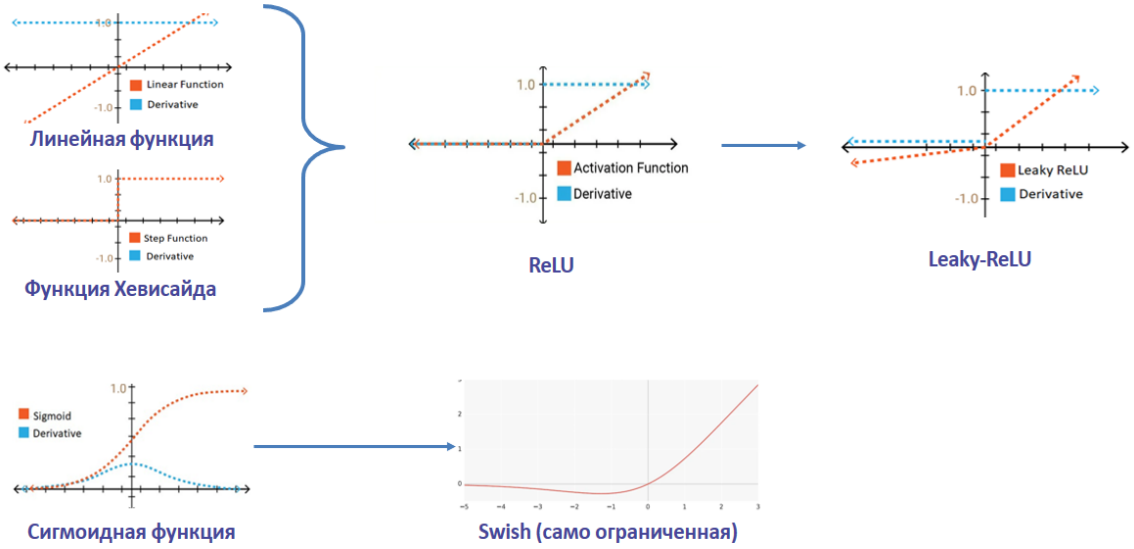


Рис. 1. Часто применяемые функции активации [5, 6]

Идея первой архитектуры тестовой нейронной сети была следующей:

- варьировать количество слоёв от одного до четырёх (наиболее подходящее значения для наборов данных с маленькой/средней размерностью);
- варьировать количество нейронов в слоях до n*2(одинаковое количество во всех слоях) где n- количество входящих данных;
- количество эпох обучения – 10 (Эмпирическим методом в ходе тестирования на собственном датасете установлено, что это минимальное количество эпох необходимое для полноценного обучения сети).

Были проведены эксперименты и тестирования, определяющие оптимальность использования представленной выше архитектуры. Затем, данная архитектура была протестирована для каждой активационной функции отдельно. Ниже представлен анализ всех функций, которые были апробированы.

Таблица 1

Анализ функций

	Softmax	Relu	Sigmoid	Tanh
Кол-во слоёв	2	1	1	2
Кол-во нейронов в слоях	50	50	350	50
Минимальная ошибка	3181.6	2845	3084.8	3084

Вторая архитектура тестовой нейронной сети для данного набора данных была изменена:

- от четырёх до пяти слоёв;
- количество нейронов в слоях варьируется;
- количество эпох и прочие характеристики сохранились.

После выполнения экспериментов были получены результаты, приведенные в таблице 2.

Таблица 2

Анализ функций

	Softmax	Relu	Sigmoid	Tanh
Кол-во слоёв	Не имеет значения	3	3	3
Кол-во нейронов в 1 слое	Не имеет значения	100	400	350
Кол-во нейронов во 2 слое	50	150	400	400
Кол-во нейронов во 3 слое	50/300	400	50	50
Кол-во нейронов в 4 слое	-/50	-	-	-
Минимальная ошибка	3181.6	1158.3	3084.2	3084

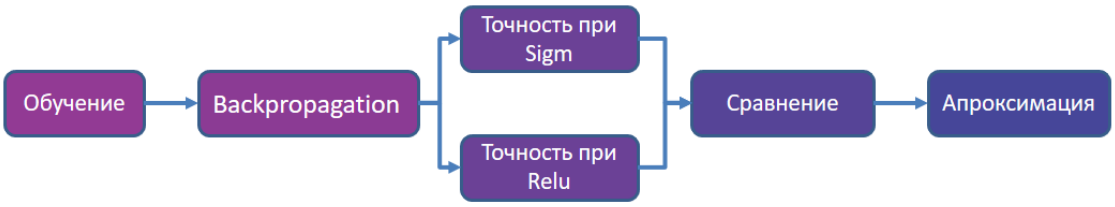
Для третьей архитектуры было принято решение изменить данные, и немного варьировать значения на определённых слоях, результат представлен ниже в таблице 3.

Таблица 3

Анализ функций

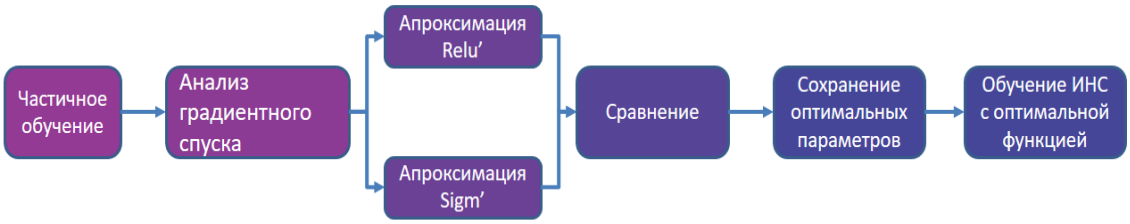
	Softmax	Relu	Sigmoid	Tanh
Кол-во слоёв	4	3	3	3
Кол-во нейронов в 1 слое	50	100	400	350
Кол-во нейронов во 2 слое	50	150	400	400
Кол-во нейронов во 3 слое	300	400	50	50
Кол-во нейронов в 4 слое	50	-	-	-
Минимальная ошибка	0.33	0	0.072	0.029

После построения данных архитектур были рассмотрены варианты иного подхода к определению активационной функции и дальнейшей работы с её результатами, представленные ниже на рисунках 2-4.



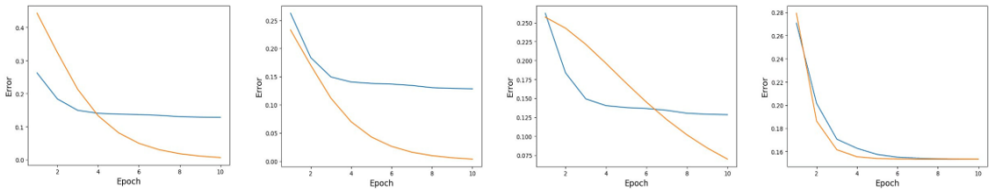
	Обычный метод	Ансамблевый
Время	2,4	1.4
Точность	0.689	0.784
Количество итераций	20	10

Рис. 2. Ансамблевый метод и сравнительная таблица



	Обычный метод	Обратный
Время	2,4	1.98
Точность	0.689	0.834
Количество итераций	20	10

Рис. 3. Обратный метод и сравнительная таблица



	Общий вид	Верхняя граница	Уменьшение шага	Нижняя граница
Ошибка	0.10920709690424221	0.09386031341750668	0.04396389659484821	0.0088876984647505
Формула	$y = \frac{e^{-Bx}}{(1+e^{-Bx})^2}$	$y = \frac{e^{-Bx}}{n(1+e^{-Bx})^2}$	$y = \frac{e^{-Bx}}{n(1+e^{-Bx})^2}$	$y = \frac{e^{-Bx}}{n(1+e^{-Bx})^2} + L$
Параметры	B=0.5; -0.5	B=0.5; -0.5 n=1,8867	B=0.2; -0.2 n=1,9841	B=1.2;-1.2 n=1.978 L=0.1204

Рис. 4. Этапы реализации функции

Тестирования были проведены на собственном наборе данных, характеристики которого представлены ниже (табл. 4).

Таблица 4

Asteroids	
Количество столбцов входящих данных	17
Размерность входящих данных	1075 элементов + возможность создания синтетических данных
Тип выходящих данных	Бинарные и не бинарные

Сравнение изменения ошибки при использовании стандартных функций и аппроксимированных функций показало, что ошибка, достигнутая в последней эпохе при измененной функции не столь заметно отличаются от стандартных значений. Несмотря на это, она всё же меньше, а также достигается за меньшее количество эпох, что особенно четко видно на первых эпохах. Значения градиента на первых эпохах сильно разнятся, что при потенциальном исследовании может положительно сказаться на затрачиваемых вычислительных мощностях. На рисунке 5 представлен график зависимости ошибки от эпохи при использовании стандартной и измененной функции (оранжевый график - график динамики ошибки при аппроксимированной функции, синий - динамика ошибки при использовании стандартных функций).

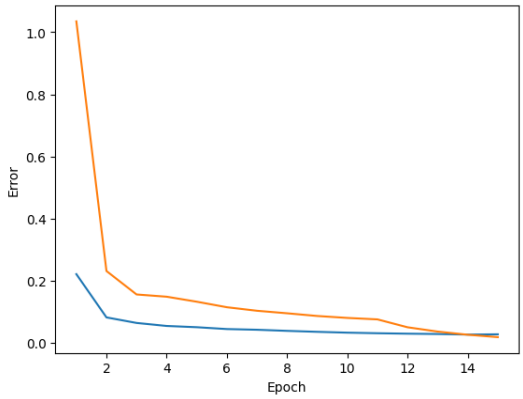


Рис. 5. График ошибки

На основе проведённого тестирования на нескольких выборках данных, было выявлено, что благодаря разработанному методу была достигнута точность подбора активационной функции равная, или даже превышающая результаты подбора человеком, что сделает нейронные сети более автономными и уменьшит время работы по подбору этих функций.

Данный метод имеет приемлемые для прикладного применения характеристики, что актуализирует создание библиотеки для этого метода, благодаря которому каждый сможет освободить себя от ручного процесса подбора активационных функций.

Ознакомится со всеми исходниками исследования и результатами можно по ссылке: https: //drive. google. com/drive/ u/1/folders/ 1O0q ELhy UtuQ S5MPW wRVH7N nZbmvZtrm.

Литература

1. Бураков М. В. Нейроные сети и нейроконтроллеры: Учебное пособие [Текст]. СПб.: Изд-во ГУА. 2013. 284 с.
2. Павленко А. Типы нейронных сетей. Принцип их работы и сфера применения [Электронный ресурс] // OTUS [сайт]. URL: https://otus.ru/nest/post/1263/ (дата обращения: 24.09.2022).
3. Инициализация веса в нейронных сетях: путь от основ к каймингу [Электронный ресурс] // Машинное обучение, нейронные сети, искусственный интеллект [сайт]. URL: https:// machinelearningmastery.ru/ weight-initialization-in- neural-networks-a-journey-from-the-basics-to-kaiming-954fb9b47c79/ (дата обращения: 11.10.2022).
4. Нейронные сети и компьютерное зрение - 55 урок. Многоклассовая классификация. Софтмакс. [Электронный ресурс]. URL: https: //www.youtube. com/ watch?app =desktop&v=UHUso2Fd_Ec (дата обращения: 12.09.2022).
5. Неявные нейронные представления с периодическими функциями активации [Электронный ресурс] // Хабр: [сайт]. URL: https:// habr. com/ru/ company/ lanit/blog/ 548250/ (дата обращения: 26.09.2022).
6. Почему инициализировать веса нейронной сети одинаковыми значениями (например, нулями) – это плохая идея [Электронный ресурс] // Хабр: [сайт]. URL: https://habr.com/ru/post/592711/ (дата обращения: 17.10.2022).

УДК 004.4

ПРОЕКТИРОВАНИЕ МОДУЛЯ ПОЛЬЗОВАТЕЛЕЙ В СИСТЕМЕ УПРАВЛЕНИЯ УЧЕБНЫМИ ПРОЕКТАМИ

Шутов Д.Э.[1] (студент)
Научный руководитель – к.т.н., доцент Горлушкина Н.Н.[2]
1 – Университет ИТМО
e-mail: danya.shutov@inbox.ru

Аннотация
В работе было проведено исследование существующих систем-аналогов, с помощью которых можно организовывать и управлять проектами. Были выявлены преимущества и недостатки этих систем, и на основе этого сформулированы функциональные требования к проектируемому модулю для собственной информационной системы. Принято решение использовать для проектирования модуля управления пользователями паттерн создания модулей на платформе Odoo.

Ключевые слова
Управление проектами, информационная система, организация проектов, управление пользователями, Odoo.

Введение

На учебных дисциплинах зачастую у студентов присутствует командная работа над проектами. В настоящее время актуальны проблемы, связанные с организацией и реализацией проектов в рамках учебной деятельности, присутствуют проблемы у преподавателей с отслеживанием статусов всех проектов, которые разрабатываются на дисциплине, получением отчетности от команд, приглашение большого количества студентов в систему и просмотр статусов их участия на учебных проектах.

Эффективная реализация образовательных проектов требует системы, которая может помочь организовать и управлять различными аспектами проекта, от планирования до выполнения. Информационная система, разработанная для поддержки управления образовательными проектами, может помочь решить эти проблемы, предоставив централизованную платформу для выполнения и мониторинга проектов. Такая система может облегчить сотрудничество между заинтересованными сторонами, предоставлять обновления в режиме реального времени о ходе проекта и обеспечивать принятие решений на основе данных.

Системы управления проектами являются важными инструментами для любой команды или организации, которым необходимо отслеживать задачи, сроки и ресурсы. Был проведен анализа платформ, на которых есть возможность организовать процесс разработки проекта, были выявлены преимущества и недостатки этих платформ [1, 2].

Среди недостатков особо выделяются:

1) нет возможности просматривать резюме бакалавров;
2) информация об опыте участниках описана очень кратко;
3) нет возможности экспорта/импорта пользователей;
4) нет возможности массового приглашения пользователей в систему;
5) нет возможности создавать необходимые пользователей в системе;
6) ограничение по количеству пользователей (студентов);
7) закрытый исходный код.

Преимущества:

1) позволяет студентам создавать свои проекты и набирать команду для реализации;

2) имеется возможность поучаствовать в любом проекте, которые есть на платформе после одобрения руководителя;

3) учёт времени и трудозатрат на поставленные задачи.

Основная часть

Было проведено небольшое исследование в виде глубинного интервью, на котором были опрошены магистранты и преподаватели. Целью опроса было выявление используемых систем для управления проектами на данный момент, узнать преимущества и недостатки этих платформ. Опрос показал такие результаты:
Основные используемые систем на данный момент:

1) trello;
2) telegram;
3) gitlab.

Преимущества платформ:

1) простата их использования;
2) удобное распределение задач посредством досок;
3) распределение ролей, выбор сроков, описание заданий;
4) код и задачи находятся в одном пространстве, что сокращает путь до проверки корректности результата;
5) удобное, понятное и установленное у всех приложение.

Основные недостатки:

1) нет возможности создания резюме для студентов;
2) нет возможности ознакомиться с описанием всех проектов в одном месте;
3) отсутствуют чаты внутри платформы;
4) нет возможности утверждения и отслеживания статусов по всем проектам.

На основании проведенного исследования в виде глубинного интервью и анализа систем-аналогов были сформулированы требования по функционалу проектируемого модуля управления пользователями в системе управления учебными проектами:

1) создание резюме студентов;
2) возможность экспорта/импорта пользователей;
3) возможность массового приглашения пользователей в систему;
4) возможность создавать необходимые роли для пользователей в системе;
5) чаты между пользователями;
6) отслеживание статусов по участию студентов в проектах (интеграция с модулем проекты).

На основании обзора систем-аналогов и принятого решения о самостоятельной разработке для реализации системы была выбрана платформа Odoo. Odoo – это набор бизнес-приложений с открытым исходным кодом, которые охватывают все потребности: CRM, электронная коммерция, бухгалтерский учет, инвентаризация, точки продаж, управление проектами и т. д. Уникальное ценностное предложение Odoo заключается в том, чтобы быть одновременно очень простым в использовании и полностью интегрированным. В Odoo хорошая платформа для написания почти любого бизнес-приложения любой сложности и позволит разработать любую функциональность применив наименьшее количество ресурсов, так как можно использовать уже реализованные пользовательские плагины.

Архитектура Odoo спроектирована так, чтобы быть модульной, гибкой и масштабируемой, что позволяет легко настраивать и расширять ее для удовлетворения потребностей любого бизнеса [3].

Odoo следует многоуровневой архитектуре, что означает, что презентация, бизнес-логика и хранилище данных разделены. В частности, используется трехуровневая архитектура.

Для проектирования модуля управления пользователями используется паттерн создания модулей на платформе Odoo. Odoo модули могут, как добавлять абсолютно новую бизнес-логику, так и расширять существующую [4]. Платформа Odoo требует

структуру следующих пакетов: модели (models), вид (views), безопасность (security), всплывающие окна (wizards), данные (data), отчет (report), сервисы (services), статические файлы (static) (изображения, CSS или javascript-файлы, используемые веб-интерфейсом или веб-сайтом), тест (test), инициализация модулей (__init__.py), файл манифеста (__manifest__.py). Управление пользователями (participants) будет осуществляться с помощью пакета модели (models).

Необходимо создать модель lp_resume.py, где будут указаны все поля, которые присутствуют в резюме. В резюме можно указать опыт работы, навыки, области интересов. С помощью этих параметров магистрантам будет легче найти подходящего студента на проект.

Также необходимо создать модель для приглашения пользователей в систему lp.project.invitation. С помощью этой модели будет возможность приглашать пользователей в систему для дальнейшего участия.

Также бакалавры смогут откликаться на понравившийся для них проект. Откликнуться один пользователь сможет только два раза с двумя приоритетами: первый приоритет и второй приоритет.

После того как бакалавр начнет участие в определенном проекте, у него будет доступ к канбан-доске, где он сможет заводить задачи и менять их статус. Так, руководитель проекта сможет оценивать прогресс выполняемых задач.

Так как в системе присутствуют несколько ролей, то необходимо создать файл ir.model.access.csv. В этом файле предоставляются доступы к разным ролям к тому или иному функционалу.

Выводы

Разработка информационной системы для управления образовательными проектами может улучшить процессы управления проектами, позволяя преподавателям упорядочивать задачи, отслеживать прогресс и принимать решения на основе данных. Однако успех системы зависит от конкретных потребностей, целей и доступных ресурсов. Важно вовлекать заинтересованные стороны в процесс разработки, чтобы убедиться, что система соответствует их ожиданиям. Были проведены глубинные интервью с магистрантами и преподавателями, которые являются основными пользователями системы. Выявлены недостатки существующих решений и сформулированы функциональные требования к модулю управления пользователями. В ходе анализа существующих решений были определены системы, которые используются для организации и ведения проектов. Были выделены ключевые недостатки этих аналогов и сформулированы функциональные требования для модуля по управлению пользователями. В результате работы был спроектирован модуль управления пользователями в системе управления учебными проектами.

Литература

1. Коваленко С.А., Гребенникова Н.И., Малиновкин В.А., Скочко П.С. Разработка электронной платформы для организации образования онлайн на основе компетентностного подхода // Воронежский государственный технический университет. 2021. №3. С. 25–31.
2. Официальный сайт ITMO.Future. [Электронный источник] Режим доступа: https://future.itmo.ru/ (дата обращения: 15.02.2023).
3. Odoo 16 Documentation [Электронный ресурс] URL: https://www. odoo.com/ documentation/16.0/ (дата последнего обращения: 18.03.2023).
4. Создание модуля / Open odoo.ru. [Электронный источник] Режим доступа: https:// doc. open-odoo.ru/ developer/11.0/ru/howtos/backend.html (дата обращения: 15.02.2023).

УДК 004.922

ТЕХНОЛОГИИ СИМУЛЯЦИИ ПОВЕРХНОСТИ ОКЕАНА В РЕЖИМЕ РЕАЛЬНОГО ВРЕМЕНИ В ИГРОВЫХ ДВИЖКАХ

Юрьев М.Р.[1]
Научный руководитель – аспирант Юрьев Р.Н.[2]
1 – ГБОУ Гимназия 642 “Земля и вселенная”
2 – Университет ИТМО
e-mail: mitrofan3452@gmail.com

Аннотация
В работе рассмотрены основные технологии симуляции океанической поверхности при помощи входных параметров на основе сложения волн Герстнера и быстрого преобразования Фурье в совокупности с многопоточными вычислениями. Так же в работе задействовано использование уровней детализации и тесселляции и приведен пример реализации при помощи движка Unity 3D с измерением производительности.
Ключевые слова
Волна Герстнера, быстрое преобразование Фурье, многопоточные вычисления, тесселляция, уровни детализации.

Введение

Симуляция реалистичного океана – одна из наиболее сложных задач в 3D графике. Данная статья описывает некоторые из наиболее продвинутых и эффективных техник симуляции океана, доступных на данный момент. В качестве примера описан пример реализации при помощи игрового движка Unity 3D.

Волны

Когда мы встречаем слово «океан», в нашем сознании рисуются разные картины, но их, как правило, объединяют две вещи: вода и волны. Симуляция реалистичного океана без волн невозможна. Нередко кажется, что разного рода шумы могут помочь избежать применения волн. Для этого обращаются к шуму Вороного или к направленным шумам. Однако эта идея не верна в корне, так как ни одним из этих шумов не возможно создать такой же или более правдоподобный рельеф водной поверхности, чем при использовании волн.

Волны Герстнера

Синусоидальные волны, которые мы взяли в качестве примера, не подходят для описания реалистичных океанических волн. Океанические волны очень сложны, однако есть и упрощенные решения, такие как волны Герстнера (Gerstner Wave/ Trochoidal Wave). Несмотря на то что они были изобретены еще в 19 веке, они до сих пор остаются простыми и актуальными для реалистичной симуляции океанических волн [1].

$$x = x_0 - \frac{f}{k} A sin(f x_0 - wt),$$

$$y = A cos(f x_0 - wt),$$

$$k = 2\pi/\lambda,$$

f в данном случае является векторной величиной. Этот параметр также называют вектором волны (wavevector), из себя он представляет горизонтальный вектор, который указывает направление движения волны; k – магнитуда этого же вектора.

Сложение волн

Исходя из опыта очевидно, что поверхность океана является не одной волной, а их совокупностью. Реализовать подобное можно сложением волн. Сложение проводится путем последовательной передачи результатов изменения поверхности от одной волны Герстнера к другой. Чем больше волн мы сложим, тем более сложный и реалистичный рельеф получим.

Применение быстрого преобразования Фурье

Вычислительные мощности современных устройств не дают возможности рассчитывать тысячи и больше волн одновременно. В этом может помочь БПФ (быстрое преобразование Фурье). Для удобства представим, что у нас есть N – количество точек на поверхности и M - количество сложенных волн. Количество операций в таком случае будет пропорционально M * N, поскольку мы рассчитываем каждую волну в каждой точке.

Мы заинтересованы в том, чтобы M было максимально большим. Реализация подобных вычислений в вершинном шейдере (Vertex Shader) попросту невозможна.

Наиболее удобное решение для этого – использование быстрого преобразования Фурье [2].

$$X_k = \sum_{n=0}^{N-1} x_n \cdot e^{-\frac{i2\pi}{N} kn},$$

ДПФ

$$X_k = \sum_{n=0}^{N-1} x_n e^{-i2\pi kn/N} \qquad k = 0, \ldots, N-1,$$

БПФ

Внешне формула БПФ крайне похожа на дискретное преобразование Фурье, однако имеет некоторые отличия. Преимущество БПФ заключается в возможности производства раздельных вычислений, таким образом, мы можем добиться сложности O(NlogN) в отличие от ДПФ O(N^N) [3]. В совокупности с возможностью распараллеливания процессов это значительно повышает производительность. Важно заметить, что преимущество в производительности будет сохраняться только при условии соблюдения ограничений по размеру сетки (рис. 1).

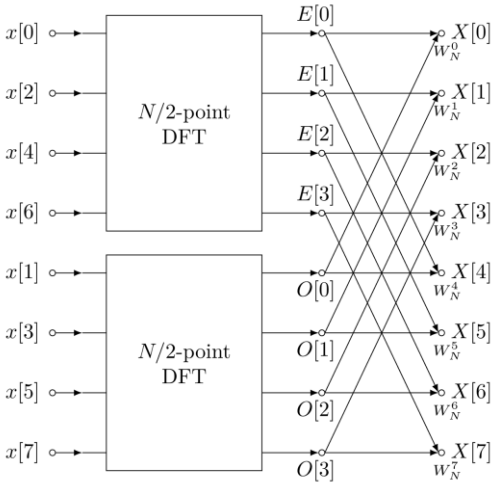


Рис. 1. Структурная формула БПФ [2]

Океанографические спектры

Внешний вид океана зависит от многочисленных физических параметров, которые на него воздействуют. Чтобы внести воздействие подобных факторов в нашу симуляцию нужно использовать океанографические спектры, которые основываются на реальных измерениях.

Уровни детализации

Для экономии ресурсов при визуализации поверхности следует использовать различные уровни детализации (LOD — Level of detail) [4]. Менее детализированные уровни должны находиться дальше от условного зрителя, в то время как более детализированные ближе к нему. Наиболее удобна будет система таких уровней, в которой центральный уровень представляет из себя квадрат, а остальные создают квадратные рамки вокруг него (рис. 2).

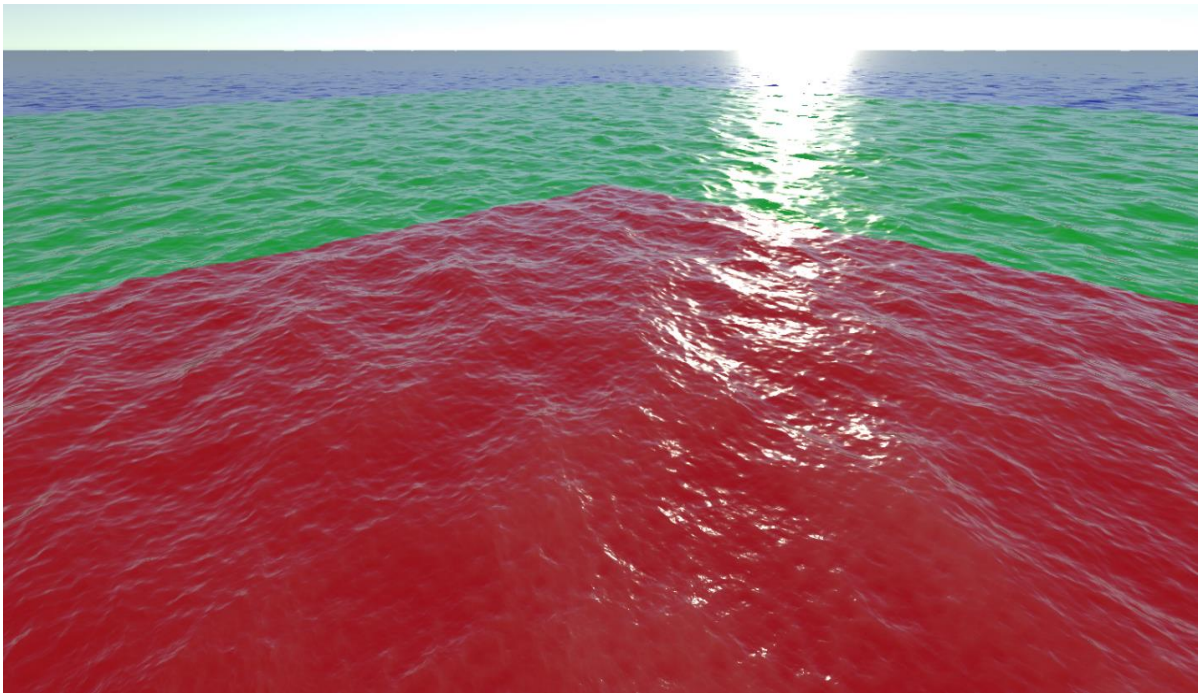


Рис. 2. Визуальная демонстрация уровней детализации при помощи разных цветов

Тесселляция поверхности

Для симуляции более детализированной и сглаженной поверхности возможно использование тесселляции. Суть процесса в подразделении поверхности непосредственно в графическом шейдере, то есть вычисления будут производиться на видеокарте, а не на процессоре.

Итоги

Реализовав вышеописанное, удалось создать настраиваемый модуль на базе Unity 3D (рис. 3), симулирующий правдоподобную океаническую поверхность на основе физических параметров. С точки зрения производительности удалось добиться частоты 60 кадров в секунду (при размере сетки 256*256) NVIDIA GeForce GTX 1080 Ti. Одним из возможных применений данного решения может быть использование его при геопозиционировании [5], для реализации потребуется воспользоваться симуляцией физики плавающих тел, описанной в одной из работ группой Безгодова Алексея Алексеевича (рис. 4) [6, 7].

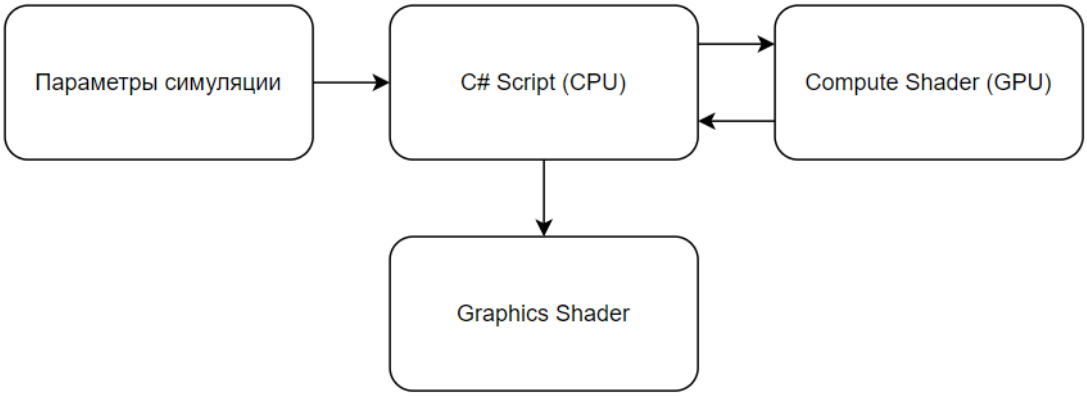


Рис. 3. Схема реализации симуляции в движке Unity 3D

Рис. 4. Внешний вид симулируемой поверхности океана

Литература

1. [Электронный ресурс]. Режим доступа: https:// en.m. wikipedia. org/wiki/ Trochoidal_ wave (дата обращения: 25.04.2023).
2. Jerry Tessendorf. Simulating ocean water. Simulating Nature: Realistic and Interactive Techniques. SIGGRAPH. 1. 2001.
3. [Электронный ресурс]. Режим доступа: https:// en. wikipedia.org/ wiki/ Fast_ Fourier_transform (дата обращения: 25.04.2023).
4. [Электронный ресурс]. Режим доступа: https:// ru. wikipedia. org/ wiki/ Level_ of_Detail (дата обращения: 25.04.2023).
5. Патент № 2712645 C1 Российская Федерация, МПК G06F 1/00. Геопозиционная мишень: № 2019100489: заявл. 04.01.2019: опубл. 30.01.2020 / Юрьев Р.Н.
6. Безгодов А.А. Визуализация динамики морских объектов в широкоэкранных системах виртуальной реальности / Безгодов А.А., Загарских А.С., Бухановский А.В. // Научно-технический вестник Санкт-Петербургского государственного университета информационных технологий, механики и оптики. 2011. № 3(73). С. 84-88.
7. Bezgodov A. A practical GPU-accelerated method for the simulation of naval objects on irregular waves / A. Bezgodov, A. Karsakov // European Association for Computer Graphics - 37th Annual Conference, EUROGRAPHICS 2016 - Posters: 37. Lisbon, 09–13 мая 2016 года. Lisbon. 2016. P. 15-16. DOI 10.2312/egp.20161044.

ЦИФРОВАЯ УРБАНИСТИКА

УДК 711.1:624.042.41

ОЦЕНКА ВЛИЯНИЯ ВЕТРОВОГО РЕЖИМА НА ОПЫТ ИСПОЛЬЗОВАНИЯ ТЕРРИТОРИИ ЖИЛЫХ ДВОРОВ САНКТ-ПЕТЕРБУРГА

Алексеевская А.А.[1] (студент)
Научный руководитель – к.соц.н., доцент Недосека Е.В.[1]
1 – Университет ИТМО
e-mail: aa.alexeevskaya@gmail.com

Аннотация
Статья посвящена оценке влияния ветра на опыт использования общественных пространств. Проведен анализ территории жилых дворов: измерена скорость ветра, опрошены пользователи, построена CFD-модель. Сопоставлено субъективное восприятие с фактической скоростью ветра, получены диапазоны комфортных значений. Приведено сравнение параметров исследуемых дворов. Обоснована необходимость разработки рекомендаций по регулированию скорости ветра в проектируемых в рамках проекта планировки территории общественных пространств.

Ключевые слова
Общественное пространство, опыт использования, проект планировки территории, скорость ветра, социологический опрос, компьютерное моделирование, CFD-анализ.

В настоящее время существует проблема сильных ветров на территории открытых общественных пространств. Ветровой комфорт является одним из ключевых факторов, оказывающих влияние на опыт использования – восприятие и ответные действия пользователя при взаимодействии со средой [1]. территории жилых дворов в частности и общественных пространств в целом. На сегодняшний день ветровой режим – ветровые условия, формируемые в конкретной местности, определяющиеся совокупностью характеристик: повторяемостью, направлением и скоростью ветра [2] не регулируются законодательно [3]. Таким образом, целью исследования стал сбор и обработка данных о существующем опыте в организации ветровых потоков для прогнозирования неблагоприятного ветрового режима на территории общественных пространств. На ветровой режим территории влияют решения, принимаемые на стадиях территориального планирования: генеральный план, правила землепользования и застройки, документация по планировки территории и архитектурное проектирование [4]. Наиболее подходящим из обозначенных документов для регулирования ветрового режима территорий на данный момент выступает проект планировки территории. Поэтому для проектирования новых кварталов целесообразно изучать опыт уже сложившихся общественных пространств в рамках реализованных проектов планировки территории, что будут далее рассмотрены в работе. В рамках исследования была разработана методика оценки влияния ветра на опыт использования общественных пространств. Методика включает в себя: выбор и анализ территории исследования; проведение фактических измерений скорости ветра; опрос пользователей общественного пространства; компьютерное моделирование ветрового режима объекта. Для исследования были выбраны современных жилые кварталы, размещенные: в сером поясе, у залива и на периферии города. Для их сравнения принимается один вид общественного пространства, наиболее распространенный в новой застройке – жилой двор. В статье будут продемонстрированы результаты исследования на примере жилых кварталов в сером поясе и у залива.

В первом случае территория исследования – жилой двор в квартале серого пояса, расположенный в ЖК «София» Фрунзенского района муниципального образования №75 на юге города Санкт-Петербург, реализованный в рамках проекта планировки и межевания территории в период с 2011 по 2018 год. На территории двора были

произведены измерения скоростей ветра по всей площади двора равномерно в 12 точках. Определены точки с максимальными, минимальными и средними скоростями. Далее в точках с характерными значениями скоростей ветра был произведен опрос пользователей. Результаты социологического опроса представлены на рисунке 1, а описание данных в таблице 1.

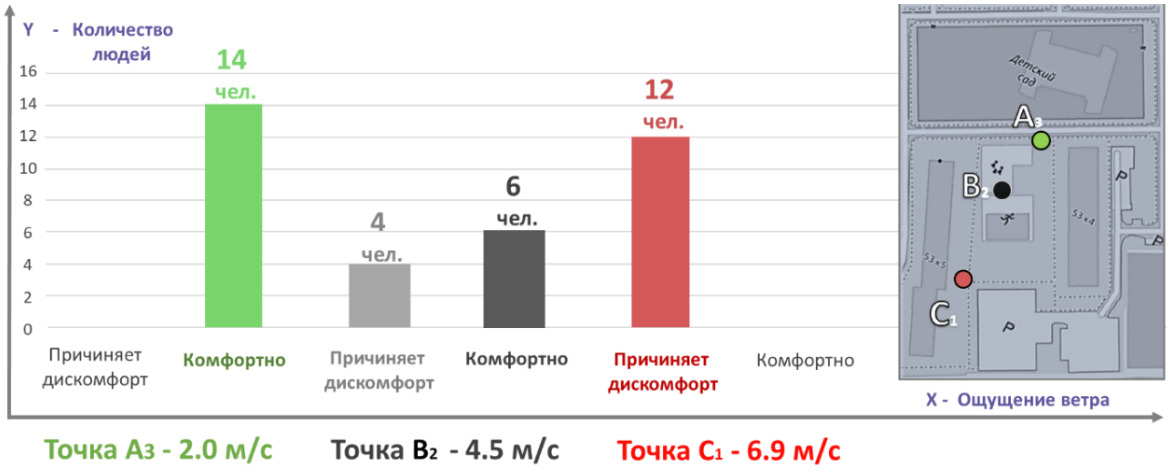


Рис. 1. Сводная столбчатая диаграмма социологического опроса в характерных точках исследуемого пространства со схемой размещения во дворе ЖК «София»

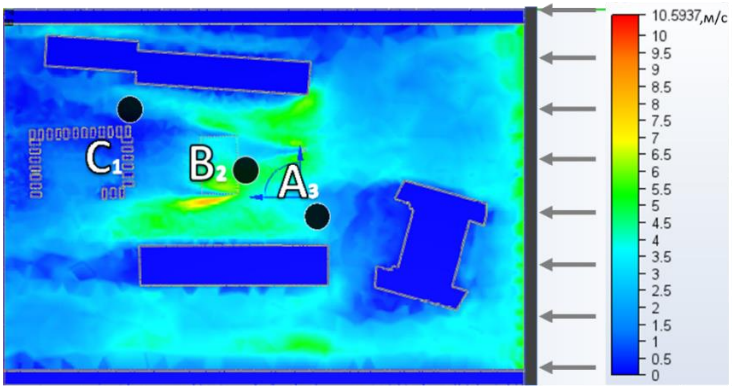
Таблица 1

Результаты социологического опроса пользователей в ключевых точках для строчной застройки жилого комплекса «София»

Точки ключевых значений	Описание размещения и результатов опроса в характерных точках
Точка А3 – с низкой скоростью 2 м/с	Расположена на тротуаре у одного из входов на детскую площадку. 12 из 14 человек отметили ветер не несущий дискомфорт и не оказывающий влияния на время пребывания в данной зоне
Точка В2 – со средней скоростью 4,5 м/с	Находится на детской площадке. 6 из 10 человек отметили, что такая скорость ветра не причиняет дискомфорта, однако оказывает существенное влияние на время нахождения во дворе. Другие опрошенные сочли не задерживаться в этом месте надолго. Все опрошенные в данной точке согласились с утверждением, что при меньшей силе ветра проводили бы на площадке больше времени
Точка С3 – с самой высокой скоростью 6,9 м/с	Размещена на тротуаре у выхода с парковки. 7 из 12 человек отказались давать интервью, мотивировав свой отказ нежеланием останавливаться из-за сильного ветра. Все опрошенные сказали, что ветер причиняет сильный дискомфорт, вызывает желание покинуть это место

Во время измерений также проводилось наблюдение за пользователями жилого двора. Для интерпретации значений как фактических измерений, так и для компьютерного моделирования был выявлен комфортный диапазон значений для пребывания пользователей общественных пространств в зимнее время года. Исходя из полученных данных установлено, что комфортное значение скорости ветра – не более 4.5 м/с. Далее была построена твердотельная модель и произведен CFD-анализ по методу конечных элементов исследуемого двора. Скорость по средним значениям замеров на границе двора с наветренной стороны и направление ветра были взяты из

экспериментальных данных. Итогом моделирования стала схема движения воздушных потоков при северном направлении ветра со скоростью 5,0 м/с на высоте около 1,5 м представлена на рисунке 2.



Точка	Измерения, м/с	Модель, м/с
A3	2,0	2,2
B2	4,5	5,0
C1	6,9	6,5

Рис. 2. Схема движения воздушных потоков с табличными значениями сравнения скоростей фактических измерений с полученными при моделировании в ЖК «София»

Во втором случае территория исследования – жилой двор был выбран в квартале у залива, расположенный в ЖК «Golden city» Василеостровского района Морского муниципального образования на западе города Санкт-Петербург, реализованный в рамках проекта планировки и межевания территории в период с 2018 по 2021 год. Для территории двора были произведены измерения скоростей по аналогии с первым случаем в 9 точках. Определены точки с максимальными, минимальными и средними скоростями ветра, произведен опрос пользователей в таких точках. Результаты социологического опроса представлены на рисунке 3, а описание полученных данных в таблице 2.

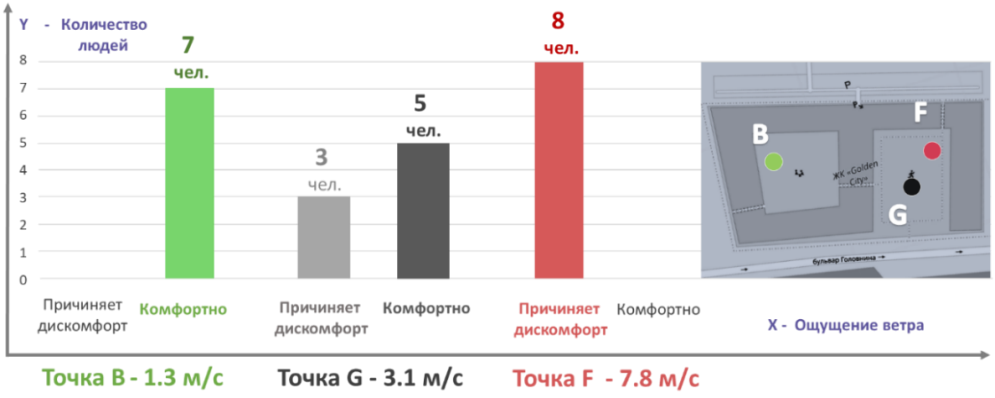


Рис. 3. Сводная столбчатая диаграмма социологического опроса в характерных точках исследуемого пространства со схемой размещения во дворе ЖК «Golden city»

В результате измерений и опроса пользователей о ветровой обстановке на территории исследования были выявлены комфортные диапазоны значений в зимнее время года с учетом повышенной влажности района исследования на территории острова. – не более 3.0 м/с. Также было произведено моделирование и CFD-анализ исследуемого общественного пространства. Параметры ветра: скорость и направление ветра были взяты из экспериментальных данных. Итогом моделирования фактических измерений для исследуемого объекта стала схема движения воздушных потоков при северном направлении ветра со скоростью 5,0 м/с на высоте около 1,5 м представлена на рисунке 4.

Таблица 2

Результаты социологического опроса пользователей в ключевых точках для периметрального типа двора в жилом комплексе «Golden city»

Точки ключевых значений	Описание размещения и результатов опроса в характерных точках
Точка B – с низкой скоростью 1,3 м/с	Расположена во внутреннем углу двора, в центре беседки. 6 из 8 опрошенных отметили почти полное отсутствие ветра, несмотря на ее фрагментарное решетчатое ограждение по периметру
Точка G – со средней скоростью 3,1 м/с	Размещена в зоне спортивных тренажеров, в центре правой части двора. 5 из 8 опрошенных указали, что скорость ветра не причиняет дискомфорта во время занятия спортом даже зимой, но оказывает существенное влияние на продолжительность. Другие же сочли такую скорость некомфортной и выразили нежелание заниматься спортом в этом месте долго
Точка F – с самой высокой скоростью 7,8 м/с	Находится по оси арки, достаточно крупного размера 6 на 4,5(H) м на транзитном маршруте. Многих пользователей невозможно было опросить из-за скорости их перемещения в данном месте. Все 8 опрошенных однозначно сказали, что ветер причиняет сильный дискомфорт

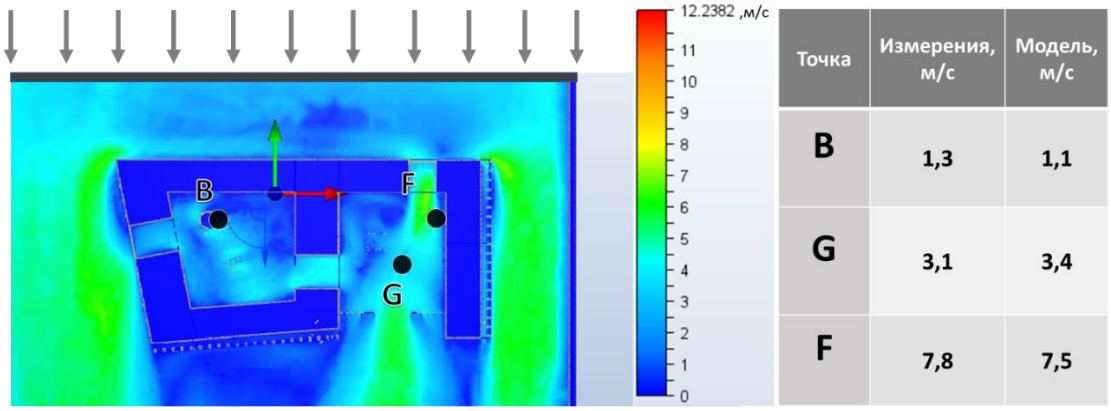


Рис. 4. Схема движения воздушных потоков с табличными значениями сравнения скоростей фактических измерений с полученными при моделировании в ЖК «Golden city»

Результаты анализа ветровых режимов двух жилых дворов Санкт-Петербурга для наглядного сравнения сведены в таблицу 3.

Таблица 3

Сравнение параметров жилых дворов Санкт-Петербурга

Параметры исследуемой территории	ЖК «София»	ЖК «Golden city»
Год постройки	2018(ППТ с 2011 до 2018 гг.)	2021(ППТ с 2018 до 2021 гг.)
Размещение квартала	В зоне бывшего серого пояса	На намывной территории
Окружение квартала	Застроенное жилыми микрорайонами	Малозастроенное жилыми комплексами
Местоположение	Фрунзенский район	Василеостровский район
Тип застройки	Строчная, 1-3 секций	Периметральная, 9 секций
Высота/ этажность	70 и 75 м/ 24 и 25 этажей	20 и 56 м/ 7 и 18 этажей

продолжение таблицы

Параметры исследуемой территории	ЖК «София»	ЖК «Golden city»
Направление ветра	Северное	Северное
Температура/ влажность	-4°С /64%	0°С/73%
Комфортный диапазон скоростей	0 – 4,5 м/с	0 – 3,0 м/с

Два жилых двора, расположенные в центре кварталов, имеют на своей территории как открытые участки с наиболее сильными ветрами: в виде заужений между зданиями, арок и зон с перепадом высот, так и участки с ветровыми тенями во внутреннем пространстве двора, образованные за счет периметрального типа застройки и угловых стыковок строений. При фактическом измерении скорости ветра параллельно проводилось включенное наблюдение пользователей во время использования исследуемых пространств. Таким образом, были выявлены зоны с определенным видом активности, соотносившимся со скоростью ветра в этих местах. Стоит отметить короткое время пребывание в обоих случаях для зон спортивных тренажеров, на транзитных маршрутах и по оси арок, где скорость ветра местами превышала 5 м/с, однако во внутренних углах секций скорость сохранялась до 2 м/с. Результаты моделирования в двух случаях подтвердили возможность применения компьютерного моделирования для прогнозирования неблагоприятных зон ветрового режима общественного пространства. Эмпирические значения скоростей ветра в обоих случаях отличаются от значений модели незначительно, максимальное отклонение которых составило не более 10%. Однако, для двух разных районов Санкт-Петербурга были установлены разные комфортные диапазоны скоростей ветра в зимний период, что свидетельствует о необходимости учета ветрового режима для разных типов местности и сезонов года.

Решением проблемы регулирования ветрового комфорта может стать создание рекомендаций по учету ветрового режима на территориях открытых общественных пространств, которые будут разработаны на основе данных об их использовании, измерений параметров ветра и с применением методов компьютерного моделирования в целях улучшения их опыта использования. Документом для регулирования ветрового режима как территории жилых дворов, так и других видов общественных пространств может стать документация по планировке территории совместно с вышестоящими документами территориального планирования и документацией по архитектурному проектированию.

Проекты планировки и межевания территории имеют детальный уровень планирования и регламентируют пространственные характеристики территории. На данном этапе разработки градостроительной документации возможен учет точного местоположения, конфигурации, расположение в пространстве и высотность зданий окружения, проведение комплексного анализа для разных времен года. Документация по планировке территории – основа для комплексного и устойчивого развития при формировании городской структуры и является реализацией генерального плана. Таким образом, проекты планировки и межевания территории выделяют основные элементы планировочной структуры и задают направления для пространственного проектирования среды, что может учесть ветровой режим разрабатываемого пространства.

Литература

1. ГОСТ Р ИСО 9241-210—2012 Эргономика взаимодействия человек-система. Часть 210: Человеко-ориентированное проектирование интерактивных систем.

2. Серебровский Ф.Л. Аэрация жилой застройки [Текст] / Серебровский Ф.Л., канд. архитектуры, доц. Москва: Стройиздат. 1971. 112 с.
3. Градостроительный кодекс Российской Федерации" от 29.12.2004 N 190-ФЗ (ред. от 06.12.2021) // Официальный интернет-портал правовой информации. [Электронный ресурс]. Режим доступа: http://pravo.gov.ru/proxy/ips/?docbody&nd=102090643 (дата обращения: 10.04.2023).
4. Руководство по оценке и регулированию ветрового режима жилой застройки /ЦНИИП градостроительства. М.: Стройиздат*. 1986. 59 с.

УДК 004.00

СОЗДАНИЕ ЕДИНОЙ УНИВЕРСАЛЬНОЙ БАЗЫ ДАННЫХ ТЕРРИТОРИЙ ПРИКРЕПЛЕНИЯ ШКОЛ К ЖИЛЫМ ДОМАМ В ГОРОДАХ РОССИИ

Апухтина Ю.В.[1] (магистрант)
Научный руководитель – к.т.н., доцент Золин А.Г.[1]
1 – Университет ИТМО
e-mail: yu.apukhtina@gmail.com

Аннотация
В работе рассмотрена структура и описано применение пространственной базы данных прикрепления жилых домов к школам, являющегося частью механизма обеспечения доступности общего образования в городах России, для работы алгоритма оптимизации территорий прикрепления и их визуализации как части интеллектуальной системы поддержки принятия решений по управлению социальной инфраструктурой.

Ключевые слова
Доступность образования, пространственный анализ, территориальное прикрепление школ, микроучастки, задача комбинаторной оптимизации.

Территориальный принцип прикрепления (зонирования) является одним из основных механизмов обеспечения доступности общего образования в большинстве крупных городов мира, включая города России; это востребованный механизм - школы по месту жительства посещают более 80% школьников. В РФ территориальная доступность школьных мест обеспечивается совокупность нормативных актов: федеральный Порядок приема на обучение по общеобразовательным указывает на требование к органам власти регулировать доступность через механизм «прикрепления» домов к школам и указывает на место жительства как на третий приоритет приема. Соответственно, актами органов власти на государственном или муниципальном уровне, в зависимости от распределения полномочий, описываются территории, жителям которых предоставляется преимущество при зачислении в одну или несколько школ на этих территориях. При этом принцип и структура самого зонирования никак не регламентированы и обычно основаны на сложившейся практике муниципалитета. В виде пространственных данных эти сведения на уровне нормативных актов не представлены, цифровые модели территорий прикрепления в уже разработанных государственных и муниципальных геоинформационных ИС отсутствуют.

При этом в настоящее время в крупных городах России наблюдается неравномерный баланс нагрузки на школьные здания: в развивающихся микрорайонах школы переполнены, в стагнирующих недозаполнены. Решение задачи оптимизации социальной инфраструктуры с помощью информационных систем имеет примеры на материале стран Европы, Африки, школьных округов в штатах США [1-3]. В России, несмотря на очевидную социальную важность, проблема оптимизации пространственной организации системы общего образования исследовалась мало; в немногих примерах [4] проблема зонирования затрагивается косвенно. В рамках разработки алгоритма оптимизации территорий прикрепления мы описываем пространственную базу данных, связывающую жилые дома со школами, ее применение для работы алгоритма и предлагаем ее использование как центра агрегации данных о прикреплении из городов РФ и основы для введения доказательных политик пространственной организации школьных систем.

В настоящее время в РФ данные о территориях прикрепления публикуются муниципальными органами в виде текстовых документов, в которых в таблицах и списках приводится соответствие «школа-адрес жилого дома» (или несколько школ-адрес жилого дома). Адрес домов, к которым прикрепляется школа, может быть приведен с точностью

не только до дома, но и до улицы и других единиц территориального деления, включая садовые товарищества, коттеджные поселки, внутригородские поселки без муниципального управления и т.д. Пример такого нормативного акта приведен на рисунке 1.

ПЕРЕЧЕНЬ

территорий, закрепляемых за муниципальными общеобразовательными учреждениями муниципального образования «город Екатеринбург»

Наименование территориальной единицы	Жилые дома	Номер муниципального общеобразовательного учреждения
1	2	3
Академический район		
40 лет Октября СНТ	Все дома	16, 19, 23, 25, 31, 79
50 лет октября СНТ	Все дома	16, 19, 23, 25, 31, 79
Автомобильного транспорта СНТ	Все дома	16, 19, 23, 25, 31, 79
Автотранспортник СНТ	Все дома	16, 19, 23, 25, 31, 79
Академика Парина ул. (нечетная сторона)	Все дома	16, 19, 23, 25, 31, 79
Академика Парина ул. (четная сторона)	№ 4, № 4/2	123, 181
	От № 38 до конца улицы	16, 19, 23, 25, 31, 79
Академика Сахарова пр. (нечетная сторона)	№ 27 – № 31б	123, 181
	№ 37 – № 45	16, 19, 23, 25, 31, 79, 123, 181
	От № 47 до конца улицы	16, 19, 23, 25, 31, 79
Академика Сахарова пр. (четная сторона)	От № 62 до конца улицы	16, 19, 23, 25, 31, 79

Рис. 1. Пример постановления о прикреплении одного из городов России
Источник: данные администрации Екатеринбурга

Также адрес может быть приведен с точностью до квартиры – некоторые большие МКД могут быть разделены по разным зонам прикрепления.

В системах образования европейских стран, США, Канады, Австралии принято описание территорий прикрепления в виде явно заданных границ зон, покрывающих, как правило, всю территорию населенного пункта или школьного округа без промежутков. Такой подход позволяет эффективно доносить до родителей данные о зонировании. Он помогает не только описывать имеющееся распределение домов, но и быстро прогнозировать изменения при появлении новых жилых адресов, включая строительство новых домов (рис. 2, 3).

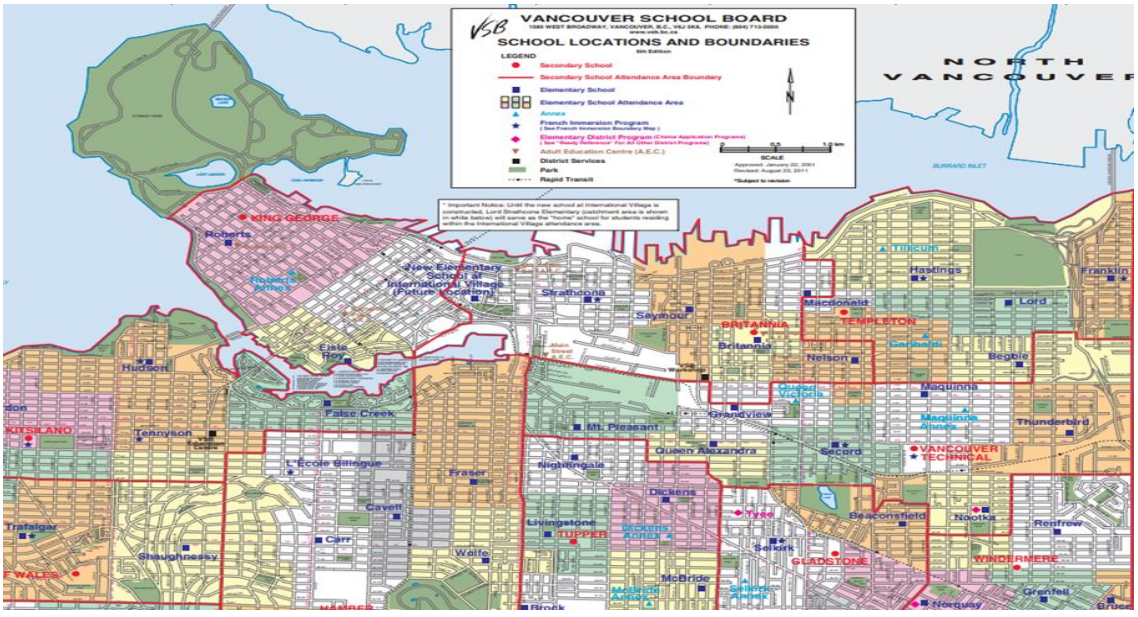


Рис. 2. Карта школьного зонирования Ванкувера (фрагмент). Разноцветные полигоны отображают зоны для начальной школы, красные линии — зоны для средней школы

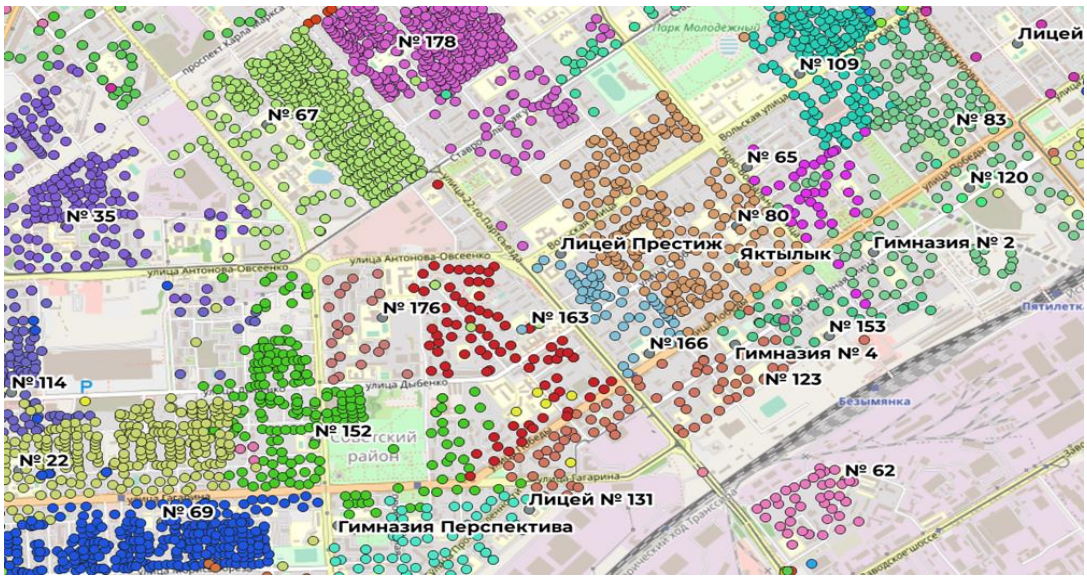


Рис. 3. Карта школьного зонирования Самары (фрагмент).
Точки одного цвета отображают дома, прикрепленные к школе

Типы зонирования могут быть выделены не только по методу описания, но и по структуре отношения школа-дом.

В настоящее время, по данным муниципальных нормативных актов и основанных на них описаний территорий прикрепления, в РФ можно выделить несколько типов таких отношений (рис. 4).

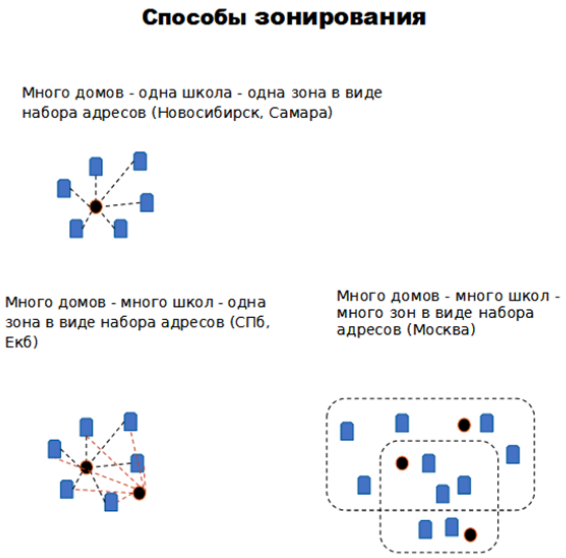


Рис. 4. Типы зонирования по типам отношений школа-дом

Алгоритм оптимизации прикрепления является частью интеллектуальной системы поддержки принятия решений (ИСППР) по управлению сетью школьных зданий. Система позволяет рассчитывать прикрепление домов к школам (формирование школьных зон), оценивать его параметры, прогнозировать изменение нагрузки при появлении новых жилых домов, оценивать последствия открытия и закрытия школьных зданий.

Основой ИСППР является алгоритм распределения жилых домов по школам с учетом мощности школьных зданий и пешеходной доступности.

Для системы общего образования города в целом оптимальным решением задачи является нахождение такого прикрепления домов к школам, при котором дети могут

добраться в школу пешком за разумное время, а мощности зданий школ используются максимально эффективно: не допускается ни переполнение (вторая смена, превышение допустимой площади в расчете на одного учащегося), ни недозагрузка зданий, влекущая за собой неэффективное использование ресурсов и потенциальное закрытие школы.

Таким образом, параметрами, определяющими оптимальность прикрепления дома к школе, являются расстояние от дома до школы и сохранение численности учащихся в школе в границах отклонения от заданного значения, определяемого площадью здания и показателем сменности.

Целевая функция состоит в максимизации числа детей, обучающихся в ближайшей к дому школе и в минимизации отклонения от заданных границ мощности.

Процесс оптимизации обеспечивается двухфазным алгоритмом аллокации на основе жадного алгоритма.

На каждой итерации алгоритма создаются кластеры домов по числу школ. Для каждого дома в таком кластере эта школа является ближайшей. Затем дома получают ранги (веса) по близости к школе от 1 до N по числу домов.

Алгоритм сначала обходит все дома с рангом 1 и прикрепляет их к ближайшей школе, затем дома с рангом 2 и так далее до N. Выполняется действие, позволяющее достичь локального оптимума – прикрепляется ближайший дом из числа тех, для кого школа является ближайшей. Если бы каждый дом удалось прикрепить к ближайшей к нему школе, задача определить максимум детей в ближайшие к дому школы была бы решена; однако этому мешают ограничения вместимости. Если в школе заканчиваются места, дом получает статус неприкрепленного на данной итерации.

Во второй и последующих итераций из списка школ удаляются школы, в которых закончились места. Неприкрепленным домам назначаются ранги в зависимости от расстояния от них до тех школ, которые теперь являются к ним ближайшими. Дома прикрепляются к ним по очереди, пока в этих школах еще есть места. Алгоритм завершает работу, когда все дома распределены или когда закончились школы.

После прохода всех домов всех рангов должны образоваться M кластеров по числу школ, с хотя бы одним домом в каждом кластере, так как для каждой школы существует хотя бы один дом, для которого эта школа является ближайшей. Доля детей, прикрепленных к ближайшей для каждого из них школе в рамках ограничений ее мощности, составит, соответственно, степень достижения заявленного значения целевой функции.

Если все школы обойдены (число итераций равно числу школ), но из-за ограничений мощности остались нераспределенные дома, события развиваются по двум сценариям, (алгоритм):

1. ИС просит пользователя увеличить число мест в школах. После получения нового ограничения цикл итераций производится заново – Этап №1 повторяется.
2. Пользователь просит ИС распределить дома по школам с произвольным оптимальным превышением их мощности. Происходит переход к Этапу №2, являющемуся расширением Этапа №1. На нем алгоритм дополняет местами школы исходя из максимальной близости и ограничения вместимости, проверяя соответствие критериям оптимальности.

Вариация алгоритма может применяться для распределения не домов, а детей по школам с учетом приоритетов родителей. Такой вариант необходим для реализации подхода к зонированию «много школ-много домов». В этом случае на большой территории искусственно создается большая зона с несколькими школами, а на этапе приемной кампании заявления распределяются исходя из параметров близости, мощности зданий и приоритетов заявителей.

В любом случае по итогам расчета формируется «зона» - именованная совокупность домов, прикрепленных к одной или нескольким школам. Она может быть

представлены в виде пар множества домов или в виде полигона, охватывающего все отнесенные к зоне здания.

Для алгоритма распределения необходимы географические координаты домов и школьных зданий, данные о количестве школьников в жилых домах (данная характеристика является оценочной, вероятностной), данные о вместимости школьных зданий. Такие данные обусловливают пространственный характер базы данных, а наличие разных схем формирования территорий прикрепления – соответствующую логику отношений.

На основании анализа типов зонирования и структуры данных, описывающих приемные кампании в школы, школьные здания и жилые дома, нами предлагается следующая схема базы данных. Подобный подход применяется в базе данных проекта SABINS [5]. В нашем случае в нем учитывается специфика российской системы образования: необходимость использовать уникальные идентификаторы школ (ИНН) и эталонные адреса и структуру адресов из Федеральной информационной адресной системы.

Элементами пространственной организации в базе данных являются координаты точек и полигонов зданий и земельных школьных корпусов и комплексов, координаты точек и/или полигоны жилых домов/комплексов, полигоны территорий (зон), на которых все дома прикреплены к школе или группе школ. Атрибутивной информацией являются данные о мощности школьного корпуса, о численности населения в уже введенных в эксплуатацию жилых домах, планируемые и строящиеся жилые дома, данные об улично-дорожной сети. База данных, реализованная на PostgresSQL/PostGIS, включает в себя следующие сущности: таблица школ как юридических лиц; таблица школьных зданий; таблица зон, связанных со школами; таблица адресов жилых домов, связанных с зонами; таблица сведений о населении в жилых домах; таблица итогов приемных кампаний прошлых лет; таблица планируемого строительства. Такая организация данных позволяет быстро вносить изменения в прикрепление, моделировать зонирование для вариантов прикрепления жилого дома к одной или нескольким школам, использовать данные об улично-дорожной сети для оценки путей от домов к школьным корпусам. В целом база данных позволяет проводить исследования пространственной организации систем общего образования, оценку ее влияния на доступность школьных мест, оценку безопасности и других параметров, характеризующие эффективность общего образования, а также включать школы в исследования территорий в целом.

Основная логическая схема базы данных приведена на рисунке. Она позволяет учесть схему зонирования «многие ко многим» и различные типы представления данных о зонировании – явные границы и списки адресов. Граф, включающий эти элементы, приведен на рисунке 5.

Рис. 5. Схема связи школ и домов

Фрагмент логической схемы БД приведен на рисунке 6.

Рис. 6. Логическая схема базы данных (фрагмент)

Такая схема (нормализованная по 1НФ, 2НФ, 3НФ) позволяет хранить как исходные данные для работы алгоритма распределения, так и данные, получаемые алгоритмом. Она помогает формировать и визуализировать территории прикрепления как в виде отдельных точек домов, размеченных по принадлежности к зоне, так и в виде полигонов, границы которых создаются вручную или методами пространственного анализа по данным граничных точек.

Предлагаемая база данных создана и использована для исследования пространственной организации системы общего образования г.о. Самара. Алгоритм оптимизации применен на выделенном фрагменте территории города, сформированы оптимальные варианты прикрепления, произведен прогноз изменения нагрузки на школьные корпуса при вводе в эксплуатацию новых МКД. При агрегации данных из городов РФ база данных предлагается в качестве централизованного источника данных для администраций муниципалитетов и для региональных ГИС ОГД для включения в оценку мощностей школьных зданий и нагрузки на них. Создание цифровой пространственной модели систем школьного образования позволяет моделировать оптимальное зонирование; прослеживать взаимосвязи между динамикой населения на территории и нагрузкой на школьных здания и в дальнейшем доказательно аргументировать необходимость строительства новых или закрытия старых школьных зданий.

Литература

1. Chen M., Thill J.-C., Delmelle E. iGLASS: An Open Source SDSS for Public School Location-Allocation. - GeoComputational Analysis and Modeling of Regional Systems. - Springer. 2018. p. 325.
2. DeVerteuil G. Reconsidering the Legacy of Urban Public Facility Location Theory in Human Geography. - Progress in Human Geography. 2000. Vol. 24. Iss. 1. P. 47–69.

3. Macharia Peter & Moturi Angela & Mumo Eda & Giorgi Emanuele & Okiro Emelda & Snow Robert & Ray Nicolas. Modelling geographic access and school catchment areas across public primary schools to support subnational planning in Kenya. Children's Geographies. 2022.
4. Турутина Т.Ф. Заблотская Я.А. Анализ территориальной доступности и экологически благоприятного функционирования школ в городской среде на примере Свердловского района Красноярска. Экология урбанизированных территорий. 2021.
5. Saporito S., Van Riper D., Wakchaure A. Building the School Attendance Boundary Information System (SABINS): Collecting, Processing, and Modeling K to 12 Educational Geography. Journal of the Urban and Regional Information Systems Association. 2013. 25(2). p. 49-62.

УДК 565.02

ЦИФРОВЫЕ МЕТОДЫ ВЗАИМОДЕЙСТВИЯ ПРИ СОУЧАСТВУЮЩЕМ ПРОЕКТИРОВАНИИ В ПЛАНИРОВАНИИ СИСТЕМ ОБЩЕСТВЕННОГО ТРАНСПОРТА

Бумагин Д.Г.[1]
Научный руководитель – ассистент ИДУ Соловьева В.В.[1]
1 – Университет ИТМО
e-mail: bumagin2000@inbox.ru

Аннотация
Цифровизация механизма соучаствующего проектирования – новый и быстроразвивающийся тренд. Он может значительно облегчить процесс анализа выявленных идей и увеличить число вовлечённых в соучаствующее проектирование людей. В данной работе рассмотрены основные цифровые методы соучаствующего проектирования при планировании систем общественного транспорта, а также выявлены наиболее целесообразные этапы их внедрения с учётом жизненного цикла проекта.

Ключевые слова
Транспортное планирование, соучаствующее проектирование, общественный транспорт, городская мобильность, цифровизация транспорта, цифровые методы взаимодействия.

При проведении соучаствующего проектирования в планировании систем общественного транспорта важно вовлекать заинтересованные стороны не на отдельно взятых этапах проекта. Не менее важно, чтобы это делалось неформально. Соучастие важно на протяжении всего жизненного цикла проекта: от предпроектного анализа существующей маршрутной сети, до тестирования и получения обратной связи по конечному реализованному продукту. Ниже, в таблице, представлен механизм вовлечения заинтересованных сторон в соотношении с жизненным циклом проекта.

Таблица

Механизм вовлечения заинтересованных сторон в соотношении с жизненным циклом проекта

Этап вовлечения	Этап разработки проекта
Выявление заинтересованных сторон и формирование проектных команд	Формирование основной команды проекта
Выявление проблематики маршрутной сети и формирование первоначального видения	Предпроектный анализ и формирование видения проекта
Вовлечение в разработку проекта сети общественного транспорта	Подготовка драфта проекта сети общественного транспорта
Вовлечение в обсуждение проекта сети общественного транспорта	Разработка проектной и проектно-сметной документации
Блок теоретических занятий на тему эффективности работы общественного транспорта в городе и нормативно-правового законодательства в сфере транспорта	Доработка и проверка проекта специалистами. Постепенное изменение маршрутной сети в городе согласно проекту
Вовлечение горожан к принятию маршрутной сети в эксплуатацию и получение обратной связи	Принятие маршрутной сети в эксплуатацию
Вовлечение в управление маршрутной сетью	Дальнейшая эксплуатация объекта

Данный механизм выработан по аналогии с алгоритмом разработки проекта благоустройства с вовлечением горожан, представленным в методических рекомендациях по вовлечению горожан в проекты благоустройства, которые были разработаны при поддержке ДОМ.РФ и Кб Стрелка [1].

На всех этапах вовлечения могут использоваться различные методы соучаствующего проектирования. Данные методы – это различные стратегии, инструменты и техники, которые помогают участникам проекта работать вместе в рамках единого процесса проектирования, учитывать разные точки зрения и обеспечивать участие всех в принятии решений. Рассмотрим несколько примеров таких методов:

1. Фокус-группы.

Фокус-группы – это группы людей, которые собираются вместе, чтобы обсудить идеи, мнения и опыт их использования. Часто фокус-группы используются на начальном этапе проекта, для определения потребностей и желаний пользователей.

2. Ко-дизайн.

Ко-дизайн – это совместная работа над проектом, в которой участвуют различные заинтересованные стороны. В процессе ко-дизайна все участники вносят свой вклад в разработку проекта, что обеспечивает более широкое понимание потребностей и мнений.

3. Итеративный дизайн.

Итеративный, или циклический дизайн – это методика, при которой проект разрабатывается поэтапно, с последующим промежуточным сбором обратной связи участников проекта. Такой подход позволяет постоянно уточнять требования и убеждаться, что конечный продукт удовлетворяет потребности пользователей.

4. Прототипирование.

Прототипирование – это создание небольшого и быстрого прототипа, который имитирует работу конечного продукта. Часто прототипирование применяется для включения пользователя в процесс проектирования и получения обратной связи от него.

5. Тестирование.

Участие в тестировании – это метод, при котором конечный продукт тестируется пользователями. Такой подход позволяет убедиться, что продукт соответствует требованиям пользователя и исправить возникшие ошибки.

6. Структурированное интервью.

Структурированные интервью являются методом получения качественной информации от представителей группы потребителей или экспертов. В ходе интервью задаются определенные вопросы, на основе ответов составляется доклад, который учитывается при проектировании.

7. Картирование пользовательского опыта.

Картирование пользовательского опыта – это метод, который помогает понимать, как пользователю необходимо использовать продукт. Этот метод используется для того, чтобы узнать, какие проблемы существуют и как их исправить.

8. Бэтчинг и порционный дизайн.

Бэтчинг – это метод, при котором возможности, которые могут быть представлены в различных временных интервалах, группируются вместе и создаются различные сервисы. Порционный дизайн – это процесс разделения продукта на мелкие функциональные элементы, то есть блокирую их по подмножествам, что позволяет участвующим уточнять каждый блок отдельно.

9. Коучинг-сессии.

Коучинг-сессии нередко используют в соучастии, чтобы не просто создать множество предложений и идей, но и помочь всем участникам разрешить конфликты и достигнуть общего знаменателя.

10. Моделирование.

При моделировании проектируемый объект создается в виде модели на компьютере. Пользователи могут попробовать и протестировать модель, прежде чем изготовить окончательный продукт.

Конечно, стоит помнить, что каждый метод ограничен своими преимуществами и проблемами, так что важно выбирать тот метод, который наиболее эффективен в конкретной ситуации.

Основная часть

Но особую роль при соучаствующем проектировании играют цифровые методы вовлечения и взаимодействия. Ведь мы живём в то время, когда цифровизация является одним из трендов развития почти всех сфер жизни общество и города.

Конечно же, транспортная сфера не является исключением. Согласно «Стратегии цифровой трансформации транспортной отрасли Российской Федерации» можно выделить следующие тренды цифровизации транспортной отрасли в целом:

- беспилотники для пассажиров и грузов;
- зеленый коридор для пассажиров и бесшовная логистика;
- цифровое управление транспортной системой РФ;
- цифровизации транспортной безопасности;
- цифровые двойники объектов транспортной инфраструктуры;
- экономика цифровой взаимопомощи;
- внедрение сервисов для популяризации карпулинга;
- развитие технологий машинного зрения;
- внедрение электронного учёта пассажиропотока;
- цифровое участие пользователей [2].

Все эти тренды можно учитывать при разработке проекта маршрутной сети, а последний из упомянутых напрямую связан с соучаствующем проектированием в планировании сетей общественного транспорта.

В рамках реализации данного проекта будет полезно использование следующих цифровых сервисов и инструментов:

- создание необходимых каналов связи;
- удалённый доступ;
- работа с GIS-инструментарием и 3d-моделями;
- разработки вэб-сайтов проектов;
- цифровые платформы для образования в сфере транспортного планирования;
- нативная реклама в СМИ;
- ведение социальных сетей проекта;
- переговоры и переписка с заинтересованными сторонами и горожанами с помощью электронных сервисов;
- интерактивные карты;
- электронные опросы.

Но важна не только методика и инструменты соучаствующего проектирования, но, как уже было сказано, и этап проекта, на которых те или иные меры применяются.

На этапе предпроектного анализа и формирования видения маршрутной сети будут полезны такие цифровые инструменты, как онлайн-сервисы с голосованием и опросами и цифровые онлайн-доски. Участники проекта могут использовать онлайн-сервисы с голосованием и опросами как при сборе общественного мнения при социологических исследованиях, так и для формирования совместного решения. А цифровые доски позволяют пользователям рисовать, замечать и комментировать другие планы и сборы участников.

На этапе подготовки эскизного проекта маршрутной сети будут актуальны веб-приложения для визуализации и обсуждения, а также GIS-инструментарий и 3D-моделирование. Веб-приложения позволяют пользователям загружать схемы,

планировки, макеты и другие данные проекта, что упрощает работу группы над проектом. 3D-моделирование позволяет пользователям создавать и редактировать виртуальные объекты, что также упрощает их взаимодействие и работу над проектом. А с помощью GIS-инструментария горожане могут просмотреть новые маршруты на интерактивной карте и оставить свои комментарии и пожелания для доработки специалистами.

Следующий этап, который заключается в разработке проектно-сметной документации, не требует внедрения механизма соучаствующего проектирования. Стоит помнить, что транспортная отрасль очень сложная сфера, и некоторые задачи в ней лучше доверять опытным специалистам.

Следующий этап – принятие в эксплуатацию и пользование маршрутной сетью. На нём будут полезны такие цифровые методы, как прототипирование с замером ключевых показателей и участие в тестировании с электронной обратной связью. Первый заключается в создании небольшого и быстрого прототипа, который имитирует работу конечного продукта. Прототипирование часто применяется для включения пользователя в процесс проектирования и получения обратной связи от него, а также для тестирования конечных решений. Втором методе конечный продукт тестируется пользователями. Такой подход позволяет убедиться, что продукт соответствует требованиям пользователя и исправить возникшие ошибки при помощи электронной обратной связи.

Выводы

Цифровые методы взаимодействия при соучаствующем проектировании могут включать в себя использование различных онлайн-платформ и инструментов, которые позволяют участникам процесса совместно работать над проектом, обмениваться мнениями и принимать общие решения. Использование цифровых методов соучаствующего проектирования может значительно ускорить и упростить процесс работы над проектом и сделать его более доступным для всех заинтересованных. Они предлагают новые возможности для улучшения процессов формирования и оптимизации маршрутных сетей общественного транспорта.

Стоит отметить, что развитая транспортная инфраструктура – одна из важнейших составляющих индекса качества городской среды [3]. Также это показатель уровня жизни и комфорта в городе в целом.

Цифровые методы участия, такие как опросы через интернет и мобильные приложения для сбора информации, позволяют собрать данные о потребностях горожан относительно маршрутов и услуг общественного транспорта. Эти данные могут быть обработаны и использованы для определения потребностей в различных точках города, объединения географически близких маршрутов, определения времени движения транспорта между остановками и т.д.

Предоставление участия и обратная связь от пользователей может помочь определить, какие маршруты имеют наибольшее число пассажиров и какие услуги следует улучшить. Учет вклада горожан в процесс формирования маршрутных сетей общественного транспорта может уменьшить количество жалоб на услуги, а также помочь повысить удобство и качество при оказании услуг.

Использование цифровых методов участия в процессе предпроектного анализа маршрутных сетей общественного транспорта и формирования их видения позволяют увеличить прозрачность процесса и участие горожан, что может привести к более точному определению потребностей и более эффективному расходованию средств городского бюджета на эти цели.

Одна из главных проблем маршрутных сетей общественного транспорта — это комплексность их использования. Города становятся все больше и усложняются, активность населения тоже растет, и пользователи перестают понимать, какие маршруты

и остановки оптимальны для их потребностей. Цифровые методы участия могут сделать этот процесс более доступным и прозрачным.

Также важно отметить, что цифровые методы участия позволяют получать быструю и детальную обратную связь от горожан об услугах общественного транспорта. Это позволяет ответственным лицам быстро определять проблемные места и предлагать устранение проблем. Обратная связь от пользователей может быть запрашиваемой или получаться через мониторинг социальных сетей и отзывов в мобильных приложениях.

Таким образом, использование цифровых методов участия при соучаствующем проектировании в планировании систем общественного транспорта имеет большие преимущества: от получения данных о потребностях и предпочтениях горожан до создания маршрутов для различных групп населения. В итоге это может привести к более удобному, доступному и качественному общественному транспорту в городе.

Литература

1. Инструменты вовлечения горожан в проекты благоустройства [Электронный ресурс]. 2018. URL: https://8architects.com/more. (дата обращения: 21.02.2023).
2. ПАСПОРТ Стратегии цифровой трансформации транспортной отрасли Российской Федерации [Электронный ресурс]. 2021. URL: https:// mintrans. gov.ru/ documents/ 8/11374?type=0. (дата обращения: 21.02.2023).
3. Индекс качества городской среды в разрезе городов и субъектов Российской Федерации за 2022 год [Электронный ресурс]. 03.04.2023. URL: https:// minstroyrf. gov.ru/docs/309391/?clear_cache=Y. (дата обращения: 03.04.2023).

УДК 332.01

КОМПЛЕКСНЫЙ ПОДХОД К ПРОЕКТИРОВАНИЮ ОБЩЕСТВЕННЫХ ПРОСТРАНСТВ

Ваврина С.С.[1] (студент)
Научный руководитель – д.филос.н., доцент Горнова Г.В.[1]
1 – Университет ИТМО
email: 370714@edu.itmo.ru

Аннотация
Статья посвящена применению комплексного подхода при проектировании системы общественных пространств крупного города в целях улучшения городской среды и повышения качества жизни в городе. Рассматривается понятие системы общественных пространств как совокупности пространств, отражающих структуру и особенности конкретного города.
Ключевые слова
Общественные пространства, комплексный подход, городское планирование, система общественных пространств, структура города.

Система общественных пространств крупного города, основанная на пространственно-планировочном каркасе территории, состоит из набора взаимосвязанных общественных пространств и имеет свои ключевые особенности в каждом отдельном городе и требует тщательной разработки на основе анализа градостроительной документации, нормативной правовой базы, потребностей местного населения, а также физико-географических особенностей местности. Комплексное планирование сети общественных пространств должно стать основой для построения системы связанных и обоснованных городских пространств, необходимых для повышения уровня комфортности городской среды.

Представленная работа направлена на решение проблемы отсутствия комплексного характера развития общественных пространств Санкт-Петербурга, отсутствия или недостаточного применения методик участия местных сообществ в проектировании пространств, учета низовых практик, а также использования ландшафтно-природного и культурного потенциала системы общественных пространств.

Указанная проблема приводит к городским конфликтам, неиспользованному потенциалу городского пространства.

Основным определением общественного пространства можно выделить определение, которое рассматривает любую общедоступную территорию города в качестве общественного пространства.

Система общественных пространств рассматривается как совокупность общественных пространств города, тесно взаимосвязанная с его ландшафтными и культурными особенностями, формирующую единый облик городского пространства.

В научной литературе общественные пространства рассматриваются с точки зрения различных подходов:

- семиотический подход – общественные пространства отражают культурные особенности территории на основе некоторой знаковой (символьной) системы (В.Л. Глазычев, К. Линч, Н. П. Анциферов, Ю.М. Лотман, Д.Л. Спивак и др.). Городская среда воплощается в мифах, метафорах, художественных произведениях. Например, существуют метафоры «город-машина», «город-сад»;
- антропологический подход рассматривает человека в контексте общественных пространств, в частности, можно выделить концепцию «третьего места» Р. Ольденбурга. В рамках антропологического подхода исследуются общественные пространства и городская среда как место обитания современного человека [1];

− социологический подход рассматривает общественные пространства как среду для социального взаимодействия (Вебер М., Дюркгейм Э., Дэвид Б. Кларк и др.);

− социально-культурный подход – сочетание культурных и коммуникативных потребностей, реализуемых на территории общественных пространств (Капков С.А., Зеленцова Е.В., Шептухина Л.И. и др.). В частности, на территории общественных пространств горожане получают новый опыт, знания, возможность реализовать творческий потенциал и так далее;

− архитектурно-эстетический подход подразумевает совокупность архитектурных и интеллектуальных качеств пространства, которые формируют городскую среду (Фризби Д., Я. Гейл., Масталерж и др.) [2].

В целях проектирования качественных общественных пространств, а также для повышения качества городской среды предлагается использование комплексного подхода к пониманию общественных пространств. Рассматриваемый подход является некоторым синтезом указанных выше методологий.

Особое внимание комплексного подхода уделяется системе общественных пространств, совокупность которых отражает сущность конкретного города, его ключевые особенности.

Например, система общественных пространств курортного города Сочи направлена на туристическую и рекреационную функцию. Основные общественные пространства лежат вокруг прибрежной линии, вдоль моря.

По мнению профессора В.Л. Глазычева, «Общественное пространство является определяющим признаком города, общественное пространство и есть город» [3].

Комплексный подход при проектировании общественных пространств включает в себя учет таких факторов как:

− историко-культурный аспект – исторические и культурные факторы, определяющие особенности территории, ее идентичность как исторической памяти, культурных общественных ожиданий. Историко-культурная идентичность места формирует памятники и ориентиры, которые становятся маркерами места, определяющие особенности и отличия территории (например, литературные произведения, историческое назначение территории, архитектурные особенности городской среды);

− пространственная структура города – разделение городской территории на определенные зоны – ядро города (исторический центр), срединная зона, периферийная зона [5]. Для каждой из этих зон присущи свои функции, особенности использования территории, которые напрямую влияют на формирования системы общественных пространств. Например, ядру города, как правило, характерна высокая деловая активность и туристическая привлекательность, что напрямую влияет на те общественные пространства, которые будут формироваться в данной зоне;

− ландшафтно-климатические особенности – особенности климата, природного ландшафта – наличие прибрежных зон, горной местности, ветрового режима территории и тому подобное. Например, прибрежные города Черноморского побережья – Сочи, Адлер и в силу своих климатических и ландшафтных особенностей стали туристическими направлениями, что влияет на то, какие общественные пространства характерны для таких городов;

− градостроительная документация – юридически закрепленное в документах обоснование и положения развития территории, содержащееся в нормативно-правовых актах [4];

− социальный аспект – особенности социальных коммуникаций на городской территории, существующие местные сообщества и низовые практики использования территории. Например, в Санкт-Петербурге одной из низовых практик горожан является использование крыш жилых домов в качестве общественных пространств, вследствие чего в историческом центре возникает целая система взаимосвязанных маршрутов и

видовых точек. Такая система общественных пространств обусловлена также пространственным и культурным аспектом города, что требует дальнейшего исследования и анализа;

− психологический аспект – влияние факторов городской среды на психологическое состояние, уровень стресса местных жителей и пользователей общественных пространств. При формировании системы общественных пространств должен быть заранее учтен фактор влияния среды на психологический комфорт горожан, снижение конфликтов и стресс-факторов среды.

При проектировании общественных пространств согласно комплексному подходу можно возможно использование таких методологий как:

− оценка психологического благополучия и удовлетворенности граждан той или иной территорией на основе социологических опросов, анализа использования территорий гражданами (натурное наблюдение, цифровые методы исследования перемещений граждан);

− уровень туристической привлекательности территории (статистика туристического потока);

− соответствие проектируемых пространств градостроительной документации;

− соответствие системы общественных пространств историко-культурной идентичности городской территории на основе методологий анализа социальных сетей (сентимент-анализ), а также опросов общественного мнения;

− учет имеющихся низовых практик и сформированных локальных сообществ при проектировании пространств на основе методик предпроектных натурных, антропологических и социологических исследований, а также использования практики соучаствующего проектирования.

Анализируя практику применения принципов проектирования общественных пространств, основанных на системном подходе, можно привести в пример High Line Park в Нью Йорке, в основе которого были использованы принципы сохранения исторической идентичности (сохранение железнодорожной ветки), доступности городской среды, использованы природные элементы благоустройства, а также было учтено мнение местного сообщества горожан, что позволило создать общественное пространство, которое является одним из самых известных парков в мире и привлекает большое число туристов ежегодно.

Среди отечественных примеров можно выделить реконструкцию центральных улиц города Рыбинска в Ярославской Области. В рамках реконструкции была произведена реставрация фасадов зданий, реконструирована улично-дорожная сеть, а также восстановлены вывески на основе архивных данных начала XX века с сохранением исторических материалов и колористических решений.

В рамках проекта были применены принципы комплексного проектирования общественных пространств – учет локальной исторической идентичности (Рыбинск как исторически торговый город), доступности городской среды, улицы были озеленены, а также при создании пространства участвовали не только городские власти, но и бизнес-сообщество, общественные организации.

Данное пространство позволило повысить. Туристическую привлекательность Рыбинска, что было одной из целей реконструкции. Рассматриваемый проект, в частности Крестовая улица Рыбинска занял первое место в конкурсе «Торговля России 2020» как лучшая торговая улица России. Непосредственно проект «Музей живой старинной вывески под открытым небом» был номинирован на приз в номинации «За сохранение лучших традиций российского гостеприимства» в рамках конкурса «Звезда качества России».

Применение комплексного подхода в проектировании общественных пространств позволит повысить качество городской среды и качество жизни населения, повысить туристическую привлекательность города, снизить количество конфликтов и стресс-

факторов при использовании общественных территорий, что положительно скажется на городской экономике.

Применение основных принципов системного проектирования пространства может быть использовано при создании широкого круга городских общественных пространств, поскольку данные принципы имеют общий характер.

Литература

1. Аванесов С.С. Городское пространство как антропологический феномен // Праксема. Проблемы визуальной семиотики. 2018. №2 (16). С. 10-31.
2. Акунина Ю.А., Ванина О.В. Проектирование креативных общественных пространств: социально-культурный подход // Вестник Московского государственного университета культуры и искусств. 2019. №3 (89). С. 167-174.
3. Городская среда. Технология развития: Настольная книга / Глазычев В.Л., Егоров М.М., Ильина Т.В. и др. Москва: Ладья. 1995. 239 с.
4. Рахманова Ю.А., Садыкгов Э.А. Классификация градостроительной документации // Актуальные проблемы землеустройства, кадастра и природообустройства. 2021. С. 275-281.
5. Филанова Т.В., Шуруев Ф.В. Анализ процесса формирования системы общественных и рекреационных пространств в крупнейшем исторически сложившимся городе // Архитектон: известия вузов. 2015. №49. С. 63-72.

УДК 004.942, 719

ОПРЕДЕЛЕНИЕ СТРАТЕГИИ УПРАВЛЕНИЯ ОКН НА ОСНОВЕ ОЦЕНКИ ЕГО СОСТОЯНИЯ И ОКРУЖАЮЩЕГО КОНТЕКСТА

Дрожжин А.И.[1] (аспирант), **Назаренко Е.А.**[1] (студент)
Научный руководитель – к.т.н. Митягин С.А.[1]
1 – Университет ИТМО
e-mail: drozhzhin@itmo.ru

Аннотация
В данной работе изучается вопрос сохранения памятников культурного наследия путем рассмотрения возможных стратегий сохранения таких объектов с учетом экономического обоснования дальнейшей эксплуатации памятника и вывода его из состояния аварийности. Актуальной задачей является определение возможностей унифицированного подхода к определению дальнейшей судьбы объекта культурного наследия.
Ключевые слова
Объекты культурного наследия, Градиентный бустинг, Выбор стратегий.

Статья посвящена проблеме управления объектами культурного наследия. Под объектом культурного наследия мы рассматриваем объекты недвижимого имущества с исторически связанными с ними территориями, возникшие в результате исторических событий, представляющие ценность с точки зрения истории, архитектуры, градостроительства, искусства, науки, эстетики и являющиеся свидетельством цивилизаций, подлинными источниками информации о зарождении и развитии культуры [1, 2]. Управление объектами культурного наследия осуществляется на нескольких уровнях и основано на выборе оптимальной стратегии, включающей варианты функционального наполнения объекта, оценки объемов инвестиций и прогноз их возврата, а также налагаемые ограничения.

Проблема управления такого рода объектами связана с тем, что традиционные методы формирования стратегий управления часто не позволяют получить варианты, одновременно обеспечивающие инвестиционную привлекательность проекта и сохранность его исторического облика. Средняя окупаемость инвестиций в подобного рода объекты – 20-25 лет [3, 4], что крайне редко может считаться положительным или нейтральным фактором при принятии решения относительно вложения инвестиций. Это приводит утрате исторического облика и памяти места, а в ряде случаев и к утрате объектов целиком. Инвесторам проще и дешевле даже с возможными штрафами снести объект, на который уже готовы документы о внесении его в охранный список (но еще не состоящий в нем), чем работать с таким объектом. Яркими примерами могут послужить сносы в Санкт-Петербурге зданий СКК «Петербургский» в 2020 г. и Манежа казарм Финляндского полка в 2022 г.

При этом важнейшим фактором определения перспектив эксплуатации здания является его окружение. Для объектов культурного наследия значимую роль играют вхождение такого объекта в исторический ансамбль, наличие объектов, связанных через исторические события, культурные аспекты или художественные свойства. Более того, на инвестиционную привлекательность влияет наполнение и состояние окружающей среды с точки зрения выбора новой функции и обоснования этого выбора. Одним из подходов, позволяющих учесть взаимосвязь объектов окружающей урбанизированной среды и объектов культурного наследия для формирования альтернативных вариантов управления выступают сетевые модели городской среды. Применение этого класса моделей позволяет учесть дополнительные эффекты окружения объекта культурного наследия, которые для некоторых объектов позволят обеспечить выбор инвестиционно-

привлекательного сценария, обеспечивающего вместе с тем сохранение исторического облика объекта.

В данной работе авторами произведено исследование и предложен подход к системному моделированию сценариев управления объектами культурного наследия на основе сетевой модели городской среды. Экспериментальные исследования метода производились на примере города Санкт-Петербурга, богатого объектами культурного наследия, что позволило выделить классы объектов, для которых предложенные методы наиболее перспективны.

Сохранение культурного наследия предполагает выбор стратегий управления объектами культурного наследия, которые позволяют обеспечить сохранность по меньшей мере внешних, а в лучшем случае и конструкционных особенностей объекта. Как правило, здания, являющиеся объектами культурного наследия, требуют повышенных затрат на поддержание и эксплуатацию, а выполняемые ими функции и не всегда способны обеспечить эти затраты. С другой стороны, ограничения, налагаемые на эксплуатацию и реконструкцию объектов культурного наследия ограничивают варианты возможных функций этих объектов. Часто это приводит к реализации таких стратегий как снос объекта культурного наследия или замена современными копиями [5-7].

В современной урбанистике признается значимость объектов культурного наследия как фундамента гражданского самосознания [8-10], восприятия культурной идентичности и эмоционального отношения граждан к месту. С культурологической и исторической точек зрения оптимальной является полная музеефикация объектов. По этой причине существует практика государственной поддержки сохранения объектов культурного наследия во всем мире. Однако, количество объектов, как сохраняемых, так и требующих сохранения может быть велико и это может требовать существенных финансовых затрат на реставрацию или хотя бы консервацию всех объектов [11, 12]. Вследствие чего финансирование зачастую выделяется только на статусные объекты [13-15]. По некоторым оценкам [16-17] выделяемые государством на реставрацию и сохранение объектов деньги почти в 20 раз меньше минимальной границы необходимой суммы. На проведение комплексной реставрации и поддержание объектов в новом состоянии потребуются более значительные затраты.

В настоящее время применяются архитектурные, законотворческие и экономические методы сохранения объектов культурного наследия. Архитектурные методы включают практики сохранения объекта как архитектурного объекта, вопросы реставрации и консервации объекта или его элементов и пр. [22, 18, 19]. Законотворческие инициативы направлены на внесение изменений в действующие законодательства стран и наднациональных организаций с целью облегчения бюрократического бремени, с которым сталкиваются как государства, так и частные лица или компании, которые работают с объектами культурного наследия [20, 21]. Экономические методы направлены на выработку подходов к повышению привлекательности конкретных объектов для частных инвесторов при обеспечении сохранность объекта культурного наследия [23-25].

В данной работе рассматривается метод формирования рекомендаций к выбору оптимальных стратегий управления объектами культурного наследия в рамках экономической группы методов, в сложившихся законодательных условиях и не рассматривая детализацию архитектурного исполнения.

Наиболее предпочтительными с точки зрения сохранения культурной идентичности места и исторической памяти выступают проекты, связанные с реставрацией объекта.

Затратная часть формируется из комплекса работ по изучению памятника, восстановления его облика виртуально, создания проекта и, собственно, работ по воссозданию утраченных конструктивов, декоративных элементов, отделки и пр. Часто

такие работы требуют подходов и материалов, которые сейчас не используются в массовом строительстве, что сильно удорожает стоимость проводимых работ.

Доходная часть формируется на основе двух вариантов дальнейшего использования объекта и их комбинации: после реставрации объект может продолжить выполнять функции, для которых он был предназначен изначально (например, жилье), либо стать частью туристической инфраструктуры (музеефикация). Смежным вариантом можно назвать использования прежде жилого объекта для размещения там отеля – функция с одной стороны, сохраняется, с другой стороны, такой объект вовлекается уже в сферу туризма.

Таким образом, доходная часть определяется следующими основными факторами:

а) охватом потенциальной аудитории;

б) доступностью объекта для посещения;

в) емкостью объекта;

г) художественной и исторической значимостью объекта;

е) контекстом объекта.

Выбор стратегии осуществляется инвестором персонально для каждого объекта культурного наследия и зависит как от параметров объекта, так и от возможностей и целей инвестора. По этой причине должны быть рассмотрены задачи выбора объекта и выбора сценария управления.

В основе предлагаемого метода лежит оценка самого объекта через его техническое состояние, окружение и анализ текущего использования аналогов.

Для определения лучшей с точки зрения доходности стратегии использования ОКН собрана информация об обеспеченности районов города различными сервисами. Основными источниками данных выступили платформа цифровой урбанистики ИТМО и картографический сервис OpenStreetMap. Собранная информация включает в себя сведения об обеспеченности районов образовательными учреждениями разного уровня, продовольственными и непродовольственными магазинами, медицинскими учреждениями и банками. Для описания обеспеченности использованы перцентили – проценты наблюдений, которые лежат ниже текущего значения. Использование такого формата измерения обеспеченности позволяет оценить не только количество определенных сервисов, но и то, насколько конкретная область города в них нуждается. Оценка обеспеченности производилась с учетом объектов, расположенных в окрестности радиусом 500 метров, что соответствует 15-минутной доступности, учитывая нелинейный доступ от центра описываемой окружности к границе.

Для каждой из предлагаемых стратегий приводится приблизительное значение прибыли и окупаемости. Для этого использованы следующие данные: информация о средней прибыльности бизнеса в конкретной сфере, информация о площади, необходимой для реализации бизнеса в конкретной сфере. Данные получены из объявлений о продаже бизнеса для 3448 объектов. Для каждого варианта стратегии использования прибыль рассчитывается как произведение средней прибыльности стратегии на квадратный метр коммерческой площади и площади здания, пригодной для использования. Для части стратегий не видится возможным рассчитать эти показатели, поскольку они несут социальную функцию или же данные об их прибыли не находятся в открытом доступе. К таким стратегиям относятся музеефикация, приспособление ОКН под государственные школы, детские сады, медицинские учреждения и банки.

Для составления итогового ранжированного списка оценено техническое состояние здания и требующиеся для его восстановления инвестиции. Для этого собрана общая информация о зданиях и данные об износе их конструктивных элементов. Общая информация об объектах представлена на платформе цифровой урбанистики ИТМО. Она включает в себя 1544 записи с информацией о расположении зданий, их постройке, конструктивных особенностях и др. Данные об износе конструктивных элементов

собраны с портала «Наш Санкт-Петербург». Они включают в себя информацию о степени разрушения 988 объектов и мероприятиях, необходимых для восстановления здания. Для предсказания используются более 70 параметров здания, включающих размеры, дату постройки, наличие коммуникаций, информацию о проведенных капитальных ремонтах, состоянии конструктивных элементов здания и пр. Значения целевого признака для обучения, валидации и тестирования моделей извлечены из информации о завершенных тендерах, размещенной в единой информационной системе в сфере закупок, а также рассчитаны с использованием Методики определения сметной стоимости строительства, реконструкции, капитального ремонта, сноса объектов капитального строительства, работ по сохранению объектов культурного наследия (памятников истории и культуры) народов Российской Федерации (РФ) на территории РФ.

Финальная оценка требующихся вложений получена с помощью модели градиентного бустинга библиотеки CatBoost, обученной на данных о состоянии зданий и архивной информации о восстановлении схожих объектов. Оптимальный набор гиперпараметров, максимизирующий прогностическую эффективность моделей, выбран путем перебора по сетке с использованием кросс-валидации. Для проверки качества модели предсказания типа восстановления здания использовалось F1-мера с макроусреднением. Для проверки качества предсказания стоимости восстановления использовалась метрика MAPE (mean absolute percentage error, средняя абсолютная ошибка в процентах):

Результаты проверки моделей на тестовой выборке с использованием данных из тендеров:

- модель предсказания типа восстановления: F1-мера = 0.87;
- модель предсказания стоимости восстановления: MAPE = 18.3%.

В соответствии с полученными результатами, наибольшее влияние на стоимость восстановления объекта культурного наследия оказывают размеры здания, степень износа фасадов и требующийся тип восстановления.

В качестве дополнительного варианта приспособления ОКН предложена возможность сдачи в аренду помещений для коммерческого использования. Для предсказания прибыльности этой стратегии приспособления использована информация о сдаче коммерческих помещений в аренду, полученная из интернет-сервиса для размещения объявлений. Для каждого помещения получены следующие данные: адрес, категория помещения, стоимость аренды, площадь. Итоговое качество модели оценивалось на основе тестовой выборки по метрике MAPE, значение которой составило 13.4%.

Итоговый ранжированный список стратегий формируется следующим образом:

- выбираются топ-5 стратегий по критерию обеспеченности. Предлагается также определение стратегии «Музеефикация», основанием для определения которой служит значимость объекта, а не информация об обеспеченности;
- выбранные стратегии упорядочиваются по сроку окупаемости восстановления здания.

Итоговый сервис представляет собой интерактивную карту с выделенными объектами культурного наследия. Она позволяет получить общую информацию о каждом объекте, а также топ-5 стратегий предлагаемого приспособления. Кроме того, в результате работы алгоритма формируется база данных, которая в дальнейшем может быть использована для обратной процедуры подбора здания под конкретную стратегию, выбранную заказчиком.

Предложенный метод позволяет осуществить предварительную оценку объектов культурного наследия для потенциальных заказчиков на основе открытых данных. Использование метода для уже восстановленных и приспособленных зданий показало высокую степень соответствия реальным данным. Это подтверждает эффективность

применения метода для объектов, находящихся в бесхозном состоянии. Повысить точность метода в дальнейшем позволят данные Комитета по государственному контролю, использованию и охране памятников истории и культуры Санкт-Петербурга. В свете увеличивающегося интереса к сохранению культурного наследия среди волонтеров, общественных организаций и государственных учреждений, планируется дальнейшая работа по улучшению алгоритма с целью его внедрения для спасения объектов культурного наследия.

Литература

1. Logan W. S. Closing Pandora's Box: Human Rights Conundrums in Cultural Heritage // Cultural heritage and human rights. 2007. С. 33-52.
2. Федеральный закон "Об объектах культурного наследия (памятниках истории и культуры) народов Российской Федерации" от 25.06.2002 N 73-ФЗ [Электронный ресурс]. – Режим доступа: http:// www.consultant.ru/document/cons_doc_LAW_37318/ (дата обращения: 16.03.2023).
3. Кудимов И.С. Повышение инвестиционной привлекательности объектов культурного наследия народов Российской Федерации // Имущественные отношения в Российской Федерации. 2020. №10 (229). С. 104-106.
4. Work plan for culture 2019-2022 [Электронный ресурс]. Режим доступа: https:// eenc.eu/uploads/ eenc-eu/ 2022/03/04/7869028db14908a0ecd4d6c4c7385b46.pdf (дата обращения: 18.04.2023).
5. Шапиро Л.Г. Особенности предварительной проверки первичной информации о преступлениях, связанных с посягательствами на объекты культурного наследия народов Российской Федерации // Правовая политика и правовая жизнь. 2021. №4. С. 99-106.
6. Такахо М.А., Братошевская В.В. Оценка рациональности сноса зданий архитектурно-исторической среды // Вестник науки. 2020. т. 3 № 6. (27). С. 181-182.
7. Полякова М.А. «Новодел» в усадьбе: «pro» et «contra» // Наследие и современность 2020. т.3 №2. С. 44-53.
8. Каминская Н.Д., Эртман Е.В.. Формирование исторического самосознания молодёжи средствами туристско-краеведческой деятельности // Вестник Московского государственного университета культуры и искусств. 2021. №5 (103). С. 144-154.
9. Лихачева А. Национальная идентичность и будущее России: доклад. Москва: Валдай. 2014. 71 с.
10. Касаткина С.С. Историко-культурное наследие городов России как элемент культурного универсума // Вестник Омского государственного педагогического университета. Гуманитарные исследования. 2021. №1 (30). С. 21-24.
11. Михайлова Н.Н.. Современные ценности культурного наследия и новые подходы к его сохранению // Традиционное прикладное искусство и образование. 2020. №1(32). С. 130-138.
12. Титаренко И.Н. Сохранение культурного наследия исторических поселений в России: история и современные проблемы // Вестник Томского государственного университета. 2020. № 450. С. 177-184.
13. Брендоусова Е.К. Методы сохранения архитектурных традиций при реновации зданий и сооружений // Баландинские чтения. 2020. т. 15 №1. С. 16-22.
14. Карпова Е.С. Реконструкция исторических сооружений в современном культурном контексте // Architecture and Modern Information Technologies. 2021. №4 (57). С. 191-211.

15. Палий К.Р. К вопросу о государственно-частном партнерстве в сфере охраны объектов культурного наследия Санкт-Петербурга // Управленческое консультирование. 2019. №5 (125). С. 140-150.
16. Heritage at risk. World report 2016-2019 [Электронный ресурс]. Режим доступа: https://openarchive.icomos.org/id/eprint/2430/1/hr20_2016_2019.pdf (дата обращения: 26.04.2023).
17. Лаврова Т.А. Финансирование сохранения объектов исторического и культурного наследия: методический аспект // Петербургский экономический журнал. 2013. № 1. С. 8-14.
18. Menodza M.A.D., Franco E. de la H., Gomez J.E.G. Technologies for the preservation of cultural heritage – A systematic review of the literature // Sustainability. №15 (1059). 2023 [Электронный ресурс]. Режим доступа: https://doi.org/10.3390/su15021059 (дата обращения: 22.04.2023).
19. Ricca M., Alberghina M.F., Randazzo L., Schiavone S., Donato A., Albanese M.P., La Russa M.F. A combined non-destructive and micro-destructive approach to solving the forensic problems in the field of cultural heritage: Two case studies // Applied sciences. т. 11. №15. 2021 [Электронный ресурс]. Режим доступа: https://doi.org/ 10.3390/app 11156951 (дата обращения: 21.04.2023).
20. Reap J.K. Introduction: Heritage legislation and management // Built herigate. №6. 2022. [Электронный ресурс]. Режим доступа: https://doi.org/10.1186/s43238-022-00059-9 (дата обращения: 21.04.2023).
21. National intangible cultural heritage (ICH) legislation & initiatives: a case study of fourteen countries [Электронный ресурс]. Режим доступа: https:// unesdoc.unesco. org/ark:/48223/ pf0000368351.locale=en (дата обращения: 20.04.2023).
22. Goddard-Bowman R. Something old is something new: The role of heritage preservation in economic development // Papers in Canadian economic development. №9. 2014. https://doi.org/10.15353/pced.v9i0.23 (дата обращения: 21.04.2023).
23. Loulanski T. Cultural heritage in socio-eceonomic development: local and global perspectives // Environments. 2006. №34(2). С. 51-69.
24. Marcinkowska M. The recovery and reconstruction of cultural heritage: The socio-economic approach // Heritage reconstruction and people: integrated recovery after trauma. 2022. С. 81-93.
25. Tisma S., Uzelac A., Jelincic D.A., Franic S., Skrtic M.M. Overwiew of social assessment methods for the economic analysus of cultural heritage investments // Journal of risk and financial management. 2022. №15(8): 327 [Электронный ресурс]. Режим доступа: https://doi.org/10.3390/jrfm15080327 (дата обращения: 21.04.2023).

УДК 130.2

ЦИФРОВЫЕ ТЕХНОЛОГИИ В ИССЛЕДОВАНИИ ХУДОЖЕСТВЕННЫХ ЦЕННОСТЕЙ

Ефименко В.П.[1] (аспирант)
Научный руководитель – д.филос.н., доцент Руди А.Ш.[1]
1 – Федеральное государственное бюджетное образовательное учреждение высшего образования «Омский государственный университет путей сообщения»
e-mail: vsegda-est@rambler.ru

Аннотация
В работе рассмотрена сложившаяся ситуация систематизациикаталожных сведений и технологических данных произведений искусства в отечественных музеях. Цель исследования – подчеркнуть необходимость создания в свободном доступе единой электронной базы данных культурных ценностей, включающей как каталожные, так и технико-технологические сведения.
Ключевые слова
Экспертиза художественных ценностей, технико-технологические исследования предметов искусства, электронная база технологического материала произведений искусства, цифровые технологии в сохранении искусства, технологическая экспертиза.

Культурное пространство городов в значительной степени формируется социальными институтами, сохраняющими культурное наследие, в немалой степени с привлечением возможностей умной среды, виртуальной реальности. Язык и форма самого современного искусства тесно связаны с цифровыми технологиями, которые начинают применять в искусстве с середины XX в. Среди искусствоведов принято считать, что первая в мире экспозиция цифрового искусства была представлена немецким математиком, физиком и философом Георгом Нисом (Georg Nees, 1926–2016) на персональной выставке «Компьютерная графика» (Computergraphik) в 1965 году. Работы художника были размещены в стенах «Учебной галереи Штутгарского колледжа» [1]. Активными субъектами деятельности по созиданию культуры в городской цифровой среде являются художественные музеи. К главным задачам музеев относится отбор из множества произведений искусства особо ценных и в дальнейшем – их сбережение и изучение. Эти задачи являются факторами, способствующими укреплению позиции комплексного научного подхода в экспертизе художественных ценностей. Важную роль в комплексной экспертизе имеют искусствоведческий, источниковедческий, технологический анализы и другие виды изучений из областей смежных наук. Данные, полученные в ходе исследований подлинных артефактов, представляют собой эталонный материал для атрибуции работ неизвестных авторов. Поэтому специалисты (искусствоведы, музейные сотрудники, эксперты культурных ценностей и др.) стремятся к пополнению материала на каждое конкретное произведение искусства; также они предпочитают иметь доступ к информации об интересующих их предметах для осуществления своей исследовательской работы на более качественном уровне.

С конца XX века технологическая экспертиза прочно закрепилась в атрибуционной и реставрационной деятельностях, независимо от того, что анализ материалов предмета искусства – дорогостоящая процедура. В особых случаях, как например, при проведении дополнительного исследования, ее запрашивают на таможне, в судебных спорах и в других правовых ситуациях. Музеи же стремятся наладить непрерывное последовательное технологическое изучение своих коллекций. При ведущих музеях страны, таких как, например: Государственный Эрмитаж, Государственная Третьяковская галерея, Государственный Русский музей, Государственный музей изобразительных искусств имени А.С. Пушкина, оборудованы научно-

исследовательские лаборатории, ориентированные на исследование художественных предметов, обогащение банка данных технологического материала и выявление эталонных экземпляров. Специалисты по технологической экспертизе подтверждают выбранный курс, в частности, К.В. Шумихин, эксперт технико-технологической экспертизы Государственной Третьяковской галереи объясняет: «Благодаря возможностям разного вида технологических исследований сегодня есть шанс по-новому взглянуть на традиционные художественные практики различных периодов, школ и отдельных мастеров» [2, с. 169].

Региональные музеи по ряду объективных обстоятельств (географической удаленности, ограниченности финансирования и др.) не имеют возможности проводить полные технологические исследования произведений искусства. Выходом из ситуации становится объединение усилий по привлечению специалистов из центральных научных лабораторий страны. Примером может служить сотрудничество между научно-исследовательской лабораторией Московского музея современного искусства и Омским областным музеем изобразительных искусств имени М.А. Врубеля. Научно-исследовательская лаборатория московского музея дает возможность осуществлять технологические исследования коллекций регионального музея, ранее ему недоступные, и в определенной мере способствует развитию технологических исследований в омском музее. Между коллегами происходит обмен профессиональным опытом, появляются совместные проекты, направленные на изучение и популяризацию произведений живописи из музея имени М.А. Врубеля. Уже на начальной стадии работы возникла потребность в получении технологических данных из лабораторий других музеев страны в качестве сравнительного материала.

Данная ситуацияподтверждает повышение востребованности в технологических сведениях при создании условий их использования. Спрос на технологические сведения вышел за пределы отдельных лабораторий. Назрела потребность в создании единой электронной базы данных технологического материала по всей России, а возможно, и включая Европу. Необходимость создания общей базы данных культурных ценностей констатировалась отдельными специалистами в области искусства еще во второй половине XX в., когда технической возможности не было. Так, советский искусствовед М.Я. Либман (1920–2010) в 1989 г. писал: «Работу по учету и определению художественных объектов в Советском Союзе должна венчать централизованная картотека наподобие “Генерального каталога”, начатого в Германии еще в первые десятилетия XX века. Эта централизованная картотека в конечном счете включит художественные ценности всех музеев любого подчинения» [3, с. 38].

Следует иметь в виду, что музейная деятельность как централизованная система может работать при соблюдении двух основных условий: а) организация и контроль функционированиясистемы на государственном уровне; б) наличие технического оснащения,соответствующего поставленным перед профессиональной системой задачам. Такая согласованность в вопросе создания единой электронной картотеки произведений искусства окончательно оформилась в 2015 г. после началаработ по ведению «Государственного каталога Музейного фонда Российской Федерации». В настоящее время данный каталог активно пополняется благодаря работе музеев всей страны. Бесспорно, каталог представляет огромный исследовательский интерес. Его данные – фамилия автора, название произведения, размер, материал исполнения и другие сведения, например провенанса картины – оказывают существенную помощь в изучении предметов искусства. Благодаря каталогу, а также оцифрованным коллекциям с каталожными сведениями, представленным на официальных сайтах музеев – как нашей страны, так и зарубежных – существенно увеличился эффективный результат атрибуционной работы. Но вместе с тем отмечается несоответствие качества заполнения многих электронных каталогов заявленным требованиям, не стал исключением и рассматриваемый пример Государственного каталога. Зачастую не указываются размер

произведения, материал и техника исполнения, однако эти данные представляют собойважную информационную составляющую для процесса атрибуции. Востребованность данных, полученных в результате технико-технологической экспертизы, демонстрирует зарубежный опыт. Например, среди исследователей из разных стран популярностью пользуются информационные сайты Исследовательского института Гетти (Getty ResearchInstitute), Института сохранения предметов искусства Гетти (Getty ConservationInstitute), Нидерландского института истории искусств (Rijksbureau voor Kunsthistorische Documentatie, сокращенно – RKD). Помимо каталожных сведений, при соблюдении дополнительных условий можно получить техническую документацию: изображения произведений в видимом, ультрафиолетовом, инфракрасном диапазонах излучения, рентгенограммы и описания к ним. Препятствием быстрому процессу атрибуции и других исследований в системе электронных технологий является отсутствие культуры открытости. В сложившейся мировой практике специализированная информация предоставляется за плату.

На условиях свободного доступа технологические данные художественных произведений подаются как материалы отдельных выставочных проектов. Например, веб-сайт «Ближе к Ван Эйку: заново открывая Гентский алтарь» (The Ghent Altarpieceand Further Worksby Janvan Eyck), созданный по проекту Lastihg Support, при поддержке Фонда Гетти (GettyFoundation) предоставил отчеты разного вида исследований, реставрации Гентского алтаря работы Яна ванн Эйка (1432. Д., м., 375×520, Собор Святого Бовона, Гент, Бельгия) и более тысячи электронных изображений в высоком разрешении с возможностью увеличения деталей, в видимом, ультрафиолетовом, инфракрасном диапазонах излучения, рентгенограмм [4]. Современным образцом демонстрации материала, касающегося различных видов изучения и реставрации памятников культурного наследия можно считать Проект «Внутри Брейгеля» (Inside Bruegel), созданный Историко-художественным музеем в Вене (Kunsthistorisches Museum Wien) при поддержке Фонда Гетти, как результат реставрации одиннадцати картин П. Брейгеля Старшего (PieterBruegeldeOude, 1525–1569) [5]. Кроме возможности изучать изображения, полученные в процессе реставрации, и анализа технологии при помощи различных физическихметодов, информационный ресурс позволяет сравнивать между собой снимки, синхронизировать их в удобном для пользователя порядке. Данный сайт интересенширокой публике и предлагает профессионалам ценный исследовательский материал.

В нашей стране практика оказания услуг по предоставлению технологических данных произведений искусства профессионалам через официальные сайты музеев и специализированных лабораторий недостаточно отработана. Ответы на обращения в организации за технологическим материалом зачастую зависят от человеческого фактора и, к сожалению, устойчива традиция утаивания информации. На сайтах свободного доступа можно найти отдельные изображения технологического материала произведений искусства с выставок. Зачастую использовать их профессионалам невозможно из-за низкого качества разрешения.

Таким образом, решением проблемы доступности технологической информации произведений искусства может быть пересмотр порядка профессионального взаимодействия музейных сотрудников, который необходимо выстроить на принципах профессионального доверия, открытости, прозрачности. Популяризация специфических данных (демонстрация их в открытом доступе) расширит представление зрителя об искусстве в целом. Единая база технологических сведений, доступная профессионалам, решит множество вопросов в атрибуционной и другого рода исследовательской работе в сфере изучения предметов культурных ценностей.

Литература

1. Как развивалось цифровое искусство / Проект Гос. музея изобр. ис-в им. А.С. Пушкина; ред. Л. Гамина: [сайт]. URL: https: // The blueprint – независимый сайт о моде, красоте и современной культуре – theblueprint.ru (дата обращения: 02.02.2023).
2. Шумихин К.В. Некоторые особенности в художественной практике русской живописи конца XIX – начала XX века // Научно-практическая конференция имени А.П. Ковалёва. 2019. М.: Гос. Третьяковская галерея. 2021. С. 169–181.
3. Либман М.Я. Актуальные вопросы атрибуции памятников искусства // Искусство. 1974. № 2. С. 36–38.
4. Closertovan Eyck (Ближе к Ван Эйку): [сайт]. URL: https:// Legacy.closertovaneyck.be (дата обращения: 14.01.2023).
5. Inside Bruegel (Внутри Брейгеля): [сайт]. URL: http:// insidebruegel.net/(дата обращения 05.04.2023).

УДК 004.62

ИССЛЕДОВАНИЕ МЕТОДОВ ВЫЯВЛЕНИЯ ТОЧЕК ОБЩЕСТВЕННОГО ИНТЕРЕСА НА ОСНОВЕ ОТКРЫТЫХ ДАННЫХ

Катынсус А.В.[1] (студент)
Научный руководитель – к.т.н. Митягин С.А.[1]
1 – Университет ИТМО
e-mail: alexkat20@rambler.ru

Аннотация
Рассматриваются различные методы выявления точек общественного интереса на основе различного вида данных. Формируется список методов, используемых в представленных решениях. Формулируется подход к решению поставленной задачи по выявлению точек общественного интереса на основе открытых данных.

Ключевые слова
Обработка естественного языка, анализ данных, точки общественного интереса, парсинг, выделение ключевых слов.

Работа над точками общественного интереса, их выявление и анализ начались не так давно. За последнее время появилось очень много способов анализа данных, было введено много новых понятий, связанных с данными [1, 2]. Сейчас решения все чаще принимаются с опорой на результаты анализа больших данных. К таким методам приходят и в урбанистике, проектировании городов и городской инфраструктуры. Понятие «точка общественного интереса» тоже появилось не так давно, и оно напрямую связано с данными. Бизнесы, городские администрации, застройщики собирают данные и стараются принимать все решения на основе имеющихся данных. В России разработка и анализ таких точек пока не сильно популярны. Но даже при таких условиях здесь стараются использовать данные от людей (анкеты, опросы), чтобы понять, насколько привлекательным или комфортным является пространство. Также у точек общественного интереса нет никаких нормативных и правовых актов, которые бы позволили однозначно их определять, выделять и объяснять их назначение. Данная тема только начинает разрабатываться, как в России, так и за рубежом.

Данные, получаемые с помощью выделения точек общественного интереса являются крайне важными и для принятия решений в разных бизнесах. Так, например, предприниматели могут на основе данных определить, где поставить точку продажи, какая целевая ниша будет в данном месте. Телекоммуникации смогут лучше предоставлять свои услуги с помощью точных адресов, лучше определять зону покрытия сетей. Банковская сфера может лучше прогнозировать способность людей брать и возвращать кредиты на основании возраста и рода деятельности людей, которые посещают ближайшие точки общественного интереса. Транспортная система выигрывает в улучшении остановок для транспорта и инфраструктуры в целом. Девелоперы и застройщики могут лучше планировать заселяемость нового здания людьми, могут решать, какой тип здания нужно строить: бизнес-центр, офис, торговый центр или жилой дом. Таким образом, можно спрогнозировать проходимость или окупаемость.

Как было сказано ранее, еще одним игроком, который выигрывает от выявления точек общественного интереса, является правительство или городская администрация. Улучшение пространств, формирование городской среды, исправление недочетов при проектировании, планирование пешеходных улиц, мониторинг состояния общественных пространств, больниц, магазинов – это все возможно благодаря собираемым данным.

Неучет точек общественного интереса или наличие невыявленных точек могут привести к плохим последствиям для всех акторов в городе, начиная с жителей,

заканчивая администрациями. Общественное пространство с негативными отзывами по всем точкам, которые в нем есть, влияет на его привлекательность, расположение отдела банка в неблагополучном районе влияет на состояние всего банка в целом. Как было сказано выше, очень многие проблемы и конфликты могут быть решены и предотвращены с помощью анализа и выявления таких точек.

Существуют относительно недавние исследования, которые затрагивают заинтересованность людей в общественных местах города. В исследовании 2016 года под названием «Public Space, public interest and Challenges of Urban Transformation» [3] (Общественные пространства, общественный интерес и задачи урбанистической трансформации), проведенном в двух городах: Приштина (Косово) и Скопье (Северная Македония), - исследователь общался с людьми в городе и задавал вопросы о том, где им больше нравится проводить время. Он выделил три типа пространств: дороги и тропинки, бульвары и площади и зеленые пространства и места для отдыха. В качестве данных можно указать ответы на вопросы, которые были заданы, это нельзя назвать открытыми данными, так как людям не всегда нравится проходить опросы, иногда они склонны к ответам, которые от них и так ожидают, что может отличаться от действительности. При этом данное исследование очень близко к теме научной исследовательской работы. Результатами исследования стали следующие факты:

- улицы и тропинки чаще используются молодыми людьми и людьми среднего возраста;
- площади и бульвары наиболее популярны у более взрослых людей;
- зеленые пространства и зоны отдыха популярны абсолютно у всех категорий людей.

В данном исследовании заранее были известны типы точек общественного интереса, что облегчало работу.

В России исследования общественных пространств и выявление точек интереса почти не проводятся, но последнее время это становится все более актуальным. В основном разбирались общественные места Москвы: сюда пригласили представителей из бюро «Гейл Архитектс», которые изучали город и разбирали все точки интереса. Остальные города не являются столь популярными и не могут похвастаться такими исследованиями. Все чаще российские университеты проводят исследования пространств, и студенты делают диссертации, дипломы и исследовательские работы на эту тему.

Так, например, в 2015 году группа студентов провела исследование в городе Иркутск. Исследование называется «Методика исследования общественных пространств центра города как основа их совершенствования» [4]. Главными задачами работы были определение мест наиболее активного пребывания людей и проведение социологического наблюдения на основе данных. В качестве данных были взяты анкеты, которые заполнялись каждый час, и количество прохожих в течение некоторого количества минут. В анкету вносились данные о людях, которые стоят на территории, ждут транспорта сидят на скамейках, едят в кафе и ресторанах, прилежащих к территории. Исследование проводилось в будние и выходные дни, а также в летний и зимний периоды. В качестве результатов исследования можно указать следующее:

- наиболее активными пользователями общественных пространств являются дети;
- активность в зимнее время снижается, но это зависит от обеспеченности пространства соответствующим для времени года оборудованием;
- самое большое количество пешеходов приходится на торговые пешеходные улицы;
- выявлена нехватка парковых пространств;
- дворовые пространства используются не полностью.

Таким образом, можно сказать, что описанное выше исследование очень сильно коррелирует с данным исследованием. Единственное отличие заключается в способе сбора данных и их обработке.

В следующей статье под названием «Urban Data Mining with Natural Language Processing: Social Media as Complementary Tool for Urban Decision Making» [2] (Интеллектуальный анализ городских данных с обработкой естественного языка: социальные сети как дополнительный инструмент для принятия городских решений) авторы разбирали, как с помощью социальных медиа и обработки естественного языка можно влиять на принятие решений в городе. Студенты Массачусетского университета использовали NLP и Data Mining, чтобы лучше понять город Андорра-ла-Велья, посредством анализа отзывов об известных общественных местах из приложения TripAdvisor.

В своей работе исследователи выделили следующие этапы:

- сбор данных (скрэппинг отзывов с сайта);
- предобработка данных с использованием Google Translate (для перевода всех отзывов на английский язык), токкенизации (разделение предложений на отдельные слова) и лемматизации (выделение корней слов);
- анализ данных (NLP);
- пользовательский интерфейс;
- визуализация данных;
- анализ полученных графиков и диаграмм;
- интеграция в городских данных;
- получение результатов и применение их в городской среде.

В качестве результатов работы студенты представляют карты с эмоционально отмеченными точками города. Красные точки – негативные эмоции, зеленые – положительные. В ходе исследования, основываясь на негативных и позитивных отзывах о данных локациях, была выявлена проблема с парковками в городе. Также далее были представлены возможные способы использования полученных систем в городе: анализ транспортной системы, анализ открытых пространств, здравоохранения, безопасности и образования.

В представленной далее статье под названием «Обнаружение точек общественного интереса в городе с помощью GPS такси с адаптивным DBSCAN» [5] (Detecting Points of Interest in a City from Taxi GPS with Adaptive DBSCAN) студенты департамента компьютерной инженерии университета Чулалонгкорна в Бангкоке также решают проблемы выявления точек общественного интереса, используя данные с GPS трекеров и метод кластеризации под названием DBSCAN (Density-based spatial clustering of applications with noise). В своей работе студенты дают следующее определение этим точкам: транспортные центры или центры активностей, достопримечательности и популярные места города. Все эти места должны быть указаны на картах городов, так как они являются центрами притяжения и позволяют жителям познакомиться с городом, отдохнуть и найти занятие по душе.

Данная работа построена на возможном использовании цифровых следов людей, таких как координаты GPS, чтобы определить, какие места больше притягивают жителей. Здесь используется алгоритм кластеризации под названием DBSCAN, который позволяет определять кластеры без учета шумов. Данные для данной работы собирались посредством визуального наблюдения высадки и посадки пассажиров такси, а также использовался датасет с данными GPS трекеров машин такси с февраля по март 2016 года. В качестве данных использовались координаты, так как по они однозначно показывают точку в пространстве. Далее ко всем найденным точкам применялся алгоритм кластеризации, описанный выше. На основании расстояний между точками, количества точек по одним и тем же координатам, частоте появления одинаковых точек все пространства города были разбиты на три категории по количеству людей: низкая

популяция, средняя популяция и высокая популяция. В каждом пространстве были выявлены кластеры, к которым относятся точки общественного интереса.

Выводом к данной работе являются полученные параметры, которые можно использовать в описанном выше алгоритме. С их помощью исследователи могут выявлять точки общественного интереса, используя местоположения остановок такси и других средств передвижения. При проверке данного метода на работоспособность были выявлены следующие точки: магазины, университеты, рестораны, торговые центры и многие другие. Как показывают результаты, в будущем предложенный метод можно успешно использовать для развития городской инфраструктуры во множестве стран.

Создание и проектирование инфраструктуры города являются очень сложными и многоуровневыми процессами. Для их корректного ведения требуется большое количество проанализированных данных, на которых можно опираться для принятия решений. Перед созданием нового объекта всегда требуются грамотная обработка местности, анализ всех точек, выявление интересов и потребностей людей в этом объекте. Таким образом, точки общественного интереса становятся ключевым источником необходимых данных, с помощью которых можно предсказывать поведение людей, востребованность нового объекта или улучшение существующего. Учет расположенных вокруг точек, состояния городской среды рядом и привлекательности пространства позволяет проработать недостатки, решить конфликты и создать удобную и комфортную городскую среду.

Выявление же точек общественного интереса продолжает очень актуальную сегодня тему по разработке и улучшению среды. Весь мир стремится сделать города удобнее и комфортнее для жителей. Все больше исследований проводится в этой сфере: студенты, архитекторы и ученые придумывают новые способы идентификации и анализа территорий.

Все методы работы с точками общественного интереса, описанные выше являются очень важными и полезными для городов, но они являются неполными. В них используются разные виды данных и разные техники. В основном описаны возможности использования опросов и анкет с целью анализа полученных результатов, постов из социальных сетей по конкретным, заранее известным местам, и данные gps трекеров в машинах. В данном исследовании заранее неизвестна ни одна точка общественного интереса, что делает его более сложным и уникальным, также данные для исследования будут собраны из множества источников: социальные сети и карты, что позволит сделать исследование более обширным. Не в каждом исследовании выше используются методы NLP для обработки текста, а также нигде не используется тональность отзывов и постов, что позволит определять привлекательность точки гораздо точнее.

С уверенностью можно говорить, что данное исследование будет полезно городам, администрациям и урбанистам, так как открытых данных становится все больше, к ним есть относительно простой доступ, для их анализа не требуется большое количество людей и техник.

На основе анализа представленных выше методов был сформулирован метод для решения задачи выделения точек общественного интереса на основе открытых данных, который состоит из трех этапов: сбор данных, анализ и выделение точек и определение привлекательности каждой такой точки. Верификация результатов будет сделана с помощью соотнесения полученных данных и выделенных сущностей, а также ручного определения привлекательности места и полученной привлекательностью.

Были выделены следующие возможности использования предложенного метода: принятие решения, с опорой на актуальное состояние городской среды, автоматизация сбора и обработки городских данных, качественное выявление точек общественного интереса, что позволит избегать негативных сценариев взаимодействия городских акторов.

Литература

1. Agryzkov T., Martí P., Nolasco-Cirugeda A., Serrano-Estrada L., Tortosa L. & Vicent J.F. Analysing successful public spaces in an urban street network using data from the social networks Foursquare and Twitter // Applied Network Science. 2016. С. 1 – 15.
2. Nai Chun Chen, Yan Zhang, Marrisa Stephens, Takehiko Nagakura, Kent Larson. Urban Data Mining with Natural Language Processing: Social Media as Complementary Tool for Urban Decision Making // CAADFutures. 17. 2017. С. 101 – 109.
3. Binak Beqaj. Public Space, public interest and Challenges of Urban Transformation // ScienceDirect. IFAC-PapersOnLine. 2016. С. 320–324.
4. Козлова Л.В. Методика исследования общественных пространств центра города как основа их совершенствования // Строительство и архитектура. 2015. С. 1–6.
5. Uriwan Angkhawey, Veera Muangsin. Detecting Points of Interest in a City from Taxi GPS with Adaptive DBSCAN // Conference: 2018 Seventh ICT International Student Project Conference (ICT-ISPC). 2018. С. 1 – 6.

УДК 656.183

МЕТОД ПЛАНИРОВАНИЯ ВЕЛОМАРШРУТНОЙ СЕТИ ГОРОДА

Клюева А.В.[1] (студент)
Научный руководитель – к.п.н., доцент Шмелева И.А.[1]
Университет ИТМО
e-mail: klvanastasiia@yandex.ru

Аннотация
В работе рассмотрены существующие рекомендации и документы по организации велотранспортного движения в городе для дальнейшего их использования в качестве основных требований и алгоритма при создании универсальной программы автоматического расчета веломаршрутной сети. Предлагается использование программы для более быстрого и эффективного планирования веломаршрутной сети любого города.

Ключевые слова
Город, велосипед, метод, программа, алгоритм, транспортная сеть, маршрут.

С ростом автомобилизации помимо доступного и комфортного транспортного средства города обрели регулярные длительные дорожные заторы, проблемы с экологией и здоровьем жителей. В связи с чем, решение для предотвращения подобных проблем многие города мира открыли в развитии городского велосипедного движения. В Нидерландах, Дании и Финляндии велосипед уже несколько десятилетий составляет конкуренцию личному автомобилю и общественному транспорту. Для этого ими была создана соответствующая инфраструктура, основой которой является связная сеть безопасных и эффективных маршрутов.

Копенгаген и Амстердам занимают первые места рейтингов по качеству велосипедной инфраструктуры [1]. В 2020 году поездки на велосипеде среди всех перемещений по обоим городам составили 26% и 30% [2], как представлено на рисунке.

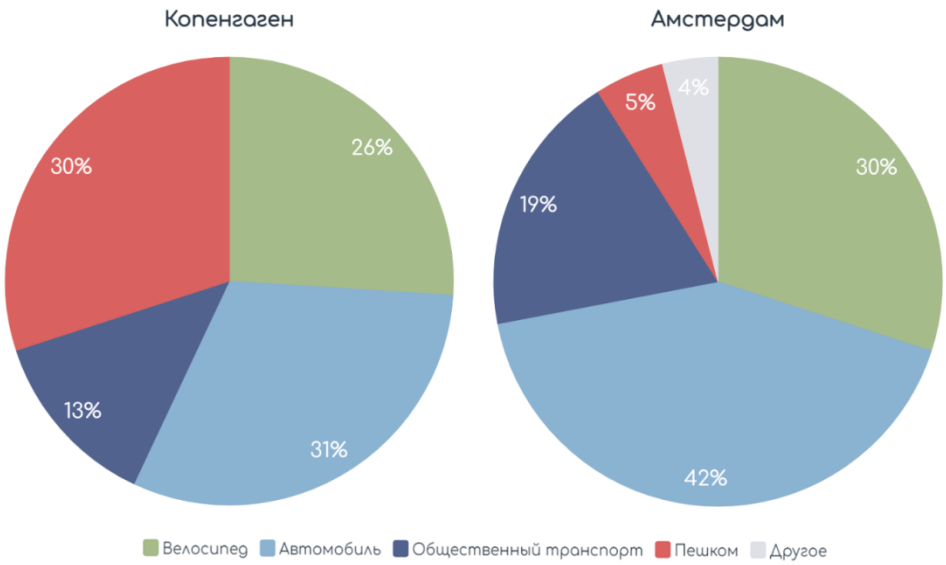


Рисунок. Доля поездок на разных видах транспорта по Копенгагену и Амстердаму в год

Эти города имеют небольшую площадь и довольно теплый климат, однако качественная велосипедная инфраструктура возможна для организации и в таком городе как Оулу. Это финский город, расположенный на широте Архангельска, и средняя температура в январе здесь составляет -9°С, при этом его площадь сравнима с площадью

Санкт-Петербурга и составляет 1410 км^2. Несмотря на это, 20% поездок в год здесь совершается на велосипеде.

К сожалению, на данный момент велосипед в России считается средством для отдыха и прогулок, а не полноценным участником дорожного движения.

При этом по данным опроса агентства исследований ResearchMe 61% россиян заявили, что катались бы на велосипеде чаще, если бы велосипедная инфраструктура в городе была развита лучше [3].

Однако в попытке перенять успешный опыт наиболее велосипедизированных городов, поселения, впервые пробующие планирование велосипедных маршрутов, сталкиваются с многими проблемами. Различный климат, ландшафт, планировка и размер города вынуждает к постоянной адаптации существующих методов планирования под конкретную местность и условия. Чем больше город, тем сложнее создать связную сеть и найти баланс среди множества рекомендаций в конкретной дорожной ситуации. Комплексные схемы веломаршрутной сети или не имеют достаточной связности, что потребует доработки в будущем, или их создание откладывается. Универсальный программный продукт, алгоритмы которого опирались бы на основные принципы внедрения велотранспортной инфраструктуры и базу проектных решений для конкретных ситуаций, мог бы упростить и ускорить процесс планирования и утверждения схем, предлагая развитие маршрутной сети.

Основными принципами для предлагаемого программного продукта могут быть рекомендации «Велосипедная инфраструктура – Общее руководство» от проекта по популяризации велосипедного движения в городах «PRESTO». В нем представлены принципы и практики от специалистов европейских городов по созданию качественной велосипедной инфраструктуры, такие как:

1. Совмещение движения автомобилей и велосипедов там, где возможно; разделение там, где необходимо. Чем выше скорость автомобильного потока, тем выше необходимость разделения движения велосипедов и автомобилей. Более подробно этот принцип представлен в правиле разделения на скоростные зоны:

– если скорость транспортного потока меньше 30 км/ч, можно не выделять велосипеды в отдельный ряд, а организовать совмещенное движение;

– 40 км/ч – отделить велосипедную полосу разметкой;

– 50–60 км/ч – отделить велосипедную дорожку бордюрной разметкой;

– Выше 70 км/ч необходимо полное разделение автомобильного и велосипедного потоков рядом припаркованных автомобилей и другими препятствиями.

2. Развитие веломаршрутной сети совместно со всей транспортной сетью.

3. Пять основных требований к веломаршрутам: безопасность, прямолинейность, связность, привлекательность и удобство. Порядок приоритетности отличается для утилитарных и рекреационных маршрутов, но безопасность всегда остается на первом месте. Для веломаршрутной сети в целом существенное значение имеют три из пяти описанных выше основных требований:

а) безопасность, которую можно достичь через снижение разрешённой скорости движения или уменьшение конфликтных точек путем разделения велосипедистов от автомобилей как во времени, так и в пространстве;

б) прямолинейность значит, что маршруты будут конкурентоспособны и иметь наибольшую эффективность при возможности пользователя доехать до точки назначения по самому прямому и короткому пути;

в) связность означает, что велодорожки не прерываются в неожиданных местах, а предоставляют пользователю разнообразие вариантов построения маршрута до максимального количества пунктов назначения и наличие пересадочных узлов с общественным транспортом [4].

Помимо вышеуказанного руководства Департаментом Московского транспорта в 2014 году также был опубликован «Альбом конструктивных элементов обустройства

велотранспортной инфраструктуры». В нем собраны типовые решения для создания безопасных и комфортных велосипедных маршрутов в городе. Выбор типа велотранспортной инфраструктуры и её конструктивных элементов зависит он существующей и планируемой интенсивности, состава велосипедного и пешеходного движения, а также интенсивности и скорости автомобильного транспорта, и происходит согласно алгоритму:

1. Выбор пункта отправления (точки генерации велосипедного потока) и пункт назначения (объекты сферы занятости, торговли и досуга).
2. Сбор исходных данных. Необходимо собрать следующие данные, в том числе на сопредельных территориях:

– геометрические параметры улиц и дорог;

– интенсивность и состав транспортных потоков и пешеходов;

– существующая организация дорожного движения;

– выявление мест концентрации ДТП и проблемных участков.

3. Разработка предварительного маршрута и его технико-экономическая оценка.
4. Определение параметров велотранспортной инфраструктуры и выбор комплекса элементов обустройства для окончательного маршрута [5].

Исходя из интенсивности транспортных потоков и условий проектирования (например, наличие парковки) одна из схем организации на перегоне и на пересечениях с другими транспортными потоками при их наличии.

5. Определение места и параметров размещения сопутствующих элементов велотранспортной инфраструктуры (велопарковки, ремонтные станции и т.д.).Данный алгоритм возможно использовать для автоматизации выбора и предложения соответствующего типа велосипедного маршрута и необходимых конструктивных элементов.

В документе предлагаются три уровня качества велоинфраструктуры: минимальный стандартный и высокого качества. Для каждого уровня сформированы планировочные решения в зависимости от габаритов улицы на дорогах и интенсивности велосипедного движения.

А именно в создании цифрового продукта, в котором будут возможны загрузка и обработка схем существующей дорожной сети города и связанной с ней базы данных. Данные составят такие характеристики сети, как габариты дорог, скоростной режим, расположение станций метро, функциональные зоны и др. Учитывая принципы создания велосипедной инфраструктуры из руководства проекта PRESTO, согласно алгоритму Альбома конструктивных элементов программа предложит схему развития веломаршрутной сети с описанием необходимых решений и элементов велосипедных маршрутов.

Создание и применение подобного продукта нацелено на ускорение непосредственно самого процесса планирования, и, впоследствии, процесса перехода к комплексному внедрению велосипедной инфраструктуры в городах.

Литература

1. The most bicycle-friendly cities of 2019 // Copenhagenize index 2019: сайт. URL: https://copenhagenizeindex.eu/. (дата обращения: 13.02.2023).
2. Deloitte City Mobility Index 2020. Amsterdam // Deloitte Insights: сайт. URL: https://www2. deloitte.com/nl/nl/ pages/publieke-sector/ articles/deloitte-city-mobility-index.html. (дата обращения: 13.02.2023).
3. Опрос: более 60% россиян катались бы на велосипеде чаще, если бы было больше велодорожек // Агентство социальной информации: сайт. URL: https:// www.asi.org.ru/ news/2021/ 08/17/ issledovanie-61-rossiyan- katalis-na-velosipede-chashhe-esli-by-bylo-bolshe-velodorozhek/. (дата обращения: 13.02.2023).

4. Dufour D. Cycling Policy Guide Infrastructure / Dufour D. Нидерланды: Ligtermoet & Partners. 2010. 61 с.
5. Альбом конструктивных элементов обустройства велотранспортной инфраструктуры. Москва: Департамент транспорта и развития дорожно-транспортной инфраструктуры города Москвы. 2014. 104 с.

УДК 911.9

АНАЛИЗ ДОРОЖНО-ТРАНСПОРТНЫХ ПРОИСШЕСТВИЙ НА ПРИМЕРЕ Г. САНКТ-ПЕТЕРБУРГА

Печенкина Е.В.[1] (студент)
Научный руководитель – к.т.н. Паниди Е.А.[1]
1 – Санкт-Петербургский Государственный Университет
e-mail: Ivanisheva-lv@mail.ru

Аннотация
В данной статье рассмотрен алгоритм выявления горячих точек дорожно-транспортных происшествий на примере города Санкт-Петербурга. Изучены статистические методы для решения данной задачи и реализованы на практике. Предложен подход для решения данной задачи, реализованный с помощью программного кода. Выявлены очаги возникновения дорожно-транспортных происшествий и представлены на итоговой карте.

Ключевые слова
ГИС, статистический анализ, матрица пространственных весов, дорожно-транспортные происшествия, автокорреляция.

В связи с увеличением числа транспортных средств, количество дорожно-транспортных происшествий (ДТП) во всем мире увеличивается с каждым годом, поэтому безопасность дорожного движения в настоящее время является важным вопросом для всех стран. По статистике аварий за 2018 год Российская Федерация занимает 72 место из 175, что заметно ниже многих развитых стран [1]. Также стоит отметить, что каждое четвертое из пяти происшествий происходит в населенных пунктах. Данные показатели свидетельствуют о необходимости детального анализа дорожно-транспортных происшествий, выявлений причин и сопутствующих факторов. Надлежащий анализ данных, связанных с авариями, необходим не только для выявления мест, где велика концентрация происшествий, но и для предоставления необходимой информации транспортным службам. Полученные результаты могут быть эффективно использованы различными ведомствами для принятия более эффективных стратегий планирования и управления в целях улучшения условий дорожного движения, а также снижения аварийности.

В целом можно выделить три главных аспекта в этой области:

1. Выделение очагов возникновения аварий.
2. Их классификация на основе определенных факторов.
3. Выявление методов предотвращения аварий.

В рамках данной работы целью было выбрано выделение горячих точек дорожно-транспортных происшествий на примере города Санкт-Петербурга.

Ученые из разных стран, используя различные методы пространственной статистики, занимаются поиском более точного и эффективного подхода. Поскольку аварии неразрывно связаны с географическими координатами, для выявления и моделирования горячих точек аварий используются методы пространственной статистики на основе ГИС. Чтобы выявить закономерности распределения аварий, а также проанализировать горячие точки, используется метод пространственной автокорреляции (Общий индекс I Морана). Пространственная автокорреляция оценивает, имеется ли кластеризация объектов, или они разбросаны и не находятся рядом. Общий индекс Морана I оценивает общую структуру и тенденцию данных. Он наиболее эффективен, когда пространственные закономерности устойчивы в пределах области интереса. Локальные статистические показатели (такие как Анализ горячих точек (Getis-Ord гетис орд Gi*)) оценивают каждый объект в контексте соседних объектов и сравнивают локальные ситуации с глобальной ситуацией, поэтому в работе был использован именно этот метод.

Для данного проекта выбран определенный диапазон времени (с 2016 года по 2019 год), что включает в себя временную составляющую исследования, построение матрицы пространственных весов с отрисовкой графа дорог и расчет глобальной и локальной автокорелляций.

Для работы использованы статистические показатели ДТП в городе Санкт-Петербурге. Все данные собираются и хранятся на официальном сайте [2]. Для проводимого исследования данные собирались с помощью программного интерфейса. Для этого был написан программный код на языке Python с использованием библиотеки requests для реализации самих запросов [3]. В результате выполнения данного программного кода, были получены файлы за последние 5 лет.

Следующим шагом стало проектирование базы данных. Для этого была выбрана объектно-реляционная система управления базами данных - PostgreSQL. Данная СУБД была выбрана по ряду причин: умеет создавать и хранить сложные структуры данных, имеет открытый исходный код, в сочетании с postgis позволяет оперативно работать с данными, которые хранят геометрию, представлена простым и понятным интерфейсом [4]. Основными сущностями базы были выбраны: ДТП, его участники, их транспортные средства, факторы, объекты УДС рядом с местом происшествия, объекты УДС на месте происшествия, также была включена информация о дорожной системе. На рисунке 1 представлена наглядная схема.

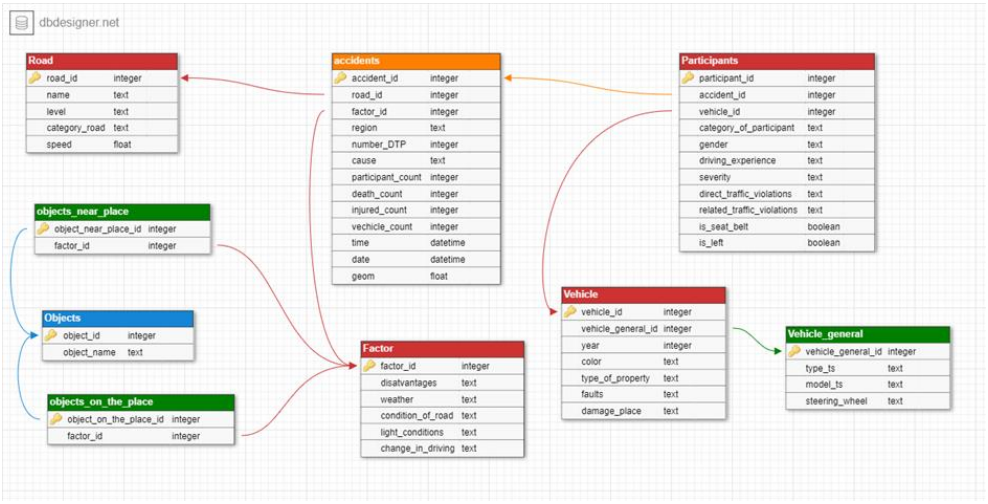


Рис. 1. Структура базы данных

Анализ ДТП невозможен без учета дорожной системы города, ведь эти два понятия неразрывно связаны между собой. Для дорожной сети использовались данные из открытого источника openstreemap, а загрузка производилась с помощью ресурса Geofabrik, позволяющего получать данные для целых федеральных округов. Так как изначально эта сеть включает в себя дороги всего северо-западного округа, то возникла необходимость обрезки и удаления лишнего с помощью утилиты osmconvert. В качестве параметров нужно указать на основе каких данных необходимо строить сеть. Существует некоторые шаблонные файлы, в которых отображены теги, относящиеся к пешеходам, машинам и так далее. Одним из таких шаблонов является файл mapconfig_for_cars.xml, который был скачан с сайта github. В нем содержатся типы дорог, их уровень и максимально допустимая скорость. В рамках данного проекта полный пакет данных не нужен, поэтому возникла необходимость редактирования данного файла. Были удалены лишние проезды, которые не были значимыми для нашей работы. Последним шагом на этом этапе был импорт конечных значений в базу данных.

Для выделения горячих точек дорожно-транспортных происшествий была предложена следующая методика:

1. Агрегация данных.
2. Определение матричных весов на основании расстояний между участками дорог.
3. Расчет глобальной автокорреляции: Moran’s I и Getis-Ord General G.*.
4. Расчет локальной автокорреляции (Getis-Ord General G.*).
5. Классификация результатов по z-score и p-value.

Алгоритм предусматривает сочетания картографической техники (методы ГИС, визуализация результатов на карте) и статистических методов (различные виды анализа). ГИС используется для локализации аварий на цифровой карте и визуализации их распространения. Статистический анализ используется для группировки однородных данных вместе [5]. Надлежащий анализ данных, связанных с авариями, необходим для выявления мест, где чаще всего происходят происшествия. С помощью ГИС пользователь может объединить данные, связанные с авариями и другими факторами, и представить результаты в удобном для работы виде.

В рамках данного исследования за нулевую гипотезу принимается то, что анализируемые атрибуты распределены случайно между объектами в выбранной области изучения; или, говоря иначе, пространственные процессы, создавшие наблюдаемую структуру значений, носят случайный характер. Здесь задача заключается в поиске пика z оценки в зависимости от порогового значения на основании расчета локальной автокорреляции. Если z оценка положительна, то можно отклонить нулевую гипотезу. Пространственное распределение высоких значений в наборе данных пространственно более кластеризовано, чем это ожидалось бы, если бы обозначенные пространственные процессы были действительно случайными. Если оценка z отрицательная, то также можно отклонить нулевую гипотезу. Но пространственное распределение низких значений в наборе данных более кластеризовано пространственно, чем если бы обозначенные пространственные процессы были действительно случайными. В результате, был проведен анализ, где выявлены горячие точки в Петроградском районе города Санкт-Петербурга, а также обозначены опасные сегменты дорог. Итогом стала карта, которая представлена на рисунке 2.

Рис. 2. Карта опасных сегментов дорог

Литература

1. Global status report on road safety 2018. Geneva: World Health Organization; 2018 [Электронный ресурс]. Режим доступа: https://www.who.int/publications-detail-redirect/9789241565684 (дата обращения: 03.10.2020).
2. Сайт статистики ГИБДД [Электронный ресурс]. Режим доступа: http://stat.gibdd.ru/ (дата обращения: 10.01.2021).
3. Документация Python [Электронный ресурс]. Режим доступа: http://docs-python.ru/ (дата обращения: 13.09.2020).
4. Официальный сайт postgis [Электронный ресурс]. Режим доступа: https://postgis.net/(дата обращения: 13.09.2020).
5. Павлов Ю.В., Королева Е.Н. Пространственные взаимодействия: оценка на основе глобального и локального индексов Морана // Пространственная Экономика. 2014. №3. С. 95–110.

УДК 130.2

ПРАКТИКИ НОСТАЛЬГИРОВАНИЯ В ЦИФРОВОМ ПРОСТРАНСТВЕ СОВРЕМЕННОГО ГОРОДА

Пластун Ю.А.[1] (соискатель)
Научный руководитель – д.филос.н., доцент Руди А.Ш.[1]
1 – Федеральное государственное бюджетное образовательное учреждение высшего образования «Омский государственный университет путей сообщения»
e-mail: julia257@bk.ru

Аннотация
В работе рассмотрены ностальгические практики в пространстве города, охватывающие два исторических периода нашей страны: дореволюционный и советский. Выдвигается инициатива новой практики ностальгирования по советскому как наиболее «востребованному» периоду ностальгических переживаний среди наших соотечественников, которая объединит жителей города всех возрастов одновременно как в цифровом, так и в городском пространстве.
Ключевые слова
Ностальгия, город, цифровой архив, QR-код, реставрация, ностальгия по советскому, ностальгия без памяти, чужая ностальгия, ностальгия по непрожитому, практики ностальгирования.

На данном этапе мировой истории человечество сталкивается с небывало высокой интенсивностью общественных изменений, и город является таким пространством, в рамках которого динамика и характер этих изменений представлены нагляднее всего. В то же время, как реакция на процессы постоянной трансформации, зачастую приносящей все новые и новые проблемы, могут быть отмечены противоположные социальные тенденции: к замедлению, сохранению статуса-кво, возвращению к истокам. Люди все чаще стали обращаться к опыту прошлого, к воспоминаниям о коллективной истории. Все это свидетельствует о том, что в обществе сегодня существует некая усталость от настоящего и страх относительно будущего. Прошлое же в восприятии человека эпохи современности становится самым безопасным и понятным местом.

Данное умонастроение способно вызывать ностальгические переживания, которые могут выражаться в практиках ностальгирования, посредством которых человек обращается к существовавшему, а также воображаемому прошлому.

Особенностью ностальгии является преобладание положительных эмоций, связанных с прошлым. Для наших соотечественников прошлое связано с существованием СССР. Недавний опрос ВЦИОМ показал, что «каждый второй россиянин убежден, что в СССР было больше хорошего (51%), и только 5% говорят об обратном, еще 38% полагают, что хорошего и плохого было поровну. Чем старше человек, тем чаще он видит в советском прошлом больше хорошего: среди граждан старше 60 лет так ответили 63%, это в два раза больше в сравнении с 18-24-летними (31%). Молодые граждане чаще других говорят, что в те времена было и хорошее, и плохое (48% и 43% среди 18-24-летних и 25-34-летних соответственно)» [1]. Почти половина опрошенных «признались, что хотели бы восстановления Советского Союза». По возрастному составу среди таковых преобладают россияне старше 60 лет (58%), однако, 18-24-летние тоже испытывают подобные чувства: таким образом высказались 27% опрошенных этой возрастной группы, при этом не более 19% из них испытывают чувство сожаления относительно распада СССР [1].

Из результатов опроса можно заключить, что на сегодняшний день немало молодых людей, которым не довелось жить во времена СССР, относятся к прошлому нашей страны достаточно позитивно, а четверть опрошенных вовсе желали бы восстановления СССР. Эти данные свидетельствуют о наличии ностальгических настроений среди молодежи, которые основываются не на индивидуальной памяти, а на

конструировании той эпохи посредством медиа и в продукции массовой культуры, а также трансляции ностальгических переживаний старшего поколения родственников. На это указывает и опрос: «У россиян старше 35 лет в числе доминирующих ассоциаций появляется детство/молодость... <…> У молодежи 18-24 лет ассоциативный ряд наполнен историческими маркерами» [1]. Данный феномен получил название ностальгии без памяти, или чужой ностальгии. Его исследователи среди причин такого явления называют неудовлетворенность настоящим и отсутствие позитивных образов будущего, то есть, такая ностальгия представляет собой стратегию совладания, форму замещения [2, с. 14; 3]. По большому счету, эти причины общие для тех, кто ностальгирует по прожитому, и для тех, кто к этой ностальгии присоединяется.

Подобные общественные тренды находят отражение в облике города и взаимодействии с ним: «Городская реконструкция, происходящая в настоящее время, теперь уже отнюдь не футуристическая, а скорее ностальгическая; город воображает свое будущее посредством импровизации со своим прошлым. Время прогресса и модернистской эффективности, материализованное в многочисленных башнях с часами и телебашнях, больше не определяет хронометрию современного города. Напротив, происходит устойчивое продолжение видимых и невидимых городов прошлого, городов мечтаний и воспоминаний, которые оказывают влияние одновременно как на новые проекты градостроительной реконструкции, так и на неформальные городские ритуалы самого низшего уровня, которые помогают нам вообразить более гуманизированную общественную среду» [4, с. 165-166].

Наблюдения С. Бойм, сделанные в самом начале XXI в., сегодня невероятно точны и актуальны. В этой связи интересен пример города Рыбинск Ярославской области, где в 2017 г. был реализован проект жителя города Д. С. Кузнецова, фолк-музыканта, под названием «Музей Живой Старинной Вывески под открытым небом». Его инициатива заключалась в том, чтобы воссоздать дореволюционные вывески на зданиях в историческом центре города. Она была поддержана мэрией и Русским географическим обществом. Кузнецовым также были определены цветовая гамма зданий во время реставрации центральных улиц города. Все реставрационные решения Кузнецова тщательно выверены на соответствие оригинальному городскому стилю и технологиям.

Таким образом, уникальный пример реставрации Рыбинска тоже можно назвать некоей ностальгией по непрожитому, по инициативе «снизу» реализованной в пространстве городской среды. Референтной эпохой в этом случае становится не Советский Союз, а царская Россия XIX в., культура которой, вероятно, выступает источником вдохновения в творчестве музыканта Д. С. Кузнецова. Как бы ни парадоксально это прозвучало, но пример реставрации Рыбинска, на наш взгляд, является скорее рефлексирующей, чем реставрационной ностальгией (по типологии С. Бойм). Несмотря на то, что автором проекта Музея была взята установка на историческую достоверность, надписи на вывесках при их аутентичном дизайне задают не ожидаемо серьезный, а ироничный контекст восприятия городского пространства.

Исторический облик центра Рыбинска выгодно отличается среди остальной части города и выделяется в ряду реставрационных проектов в других городах, что демонстрирует привлекательность российского культурного наследия в современных условиях и создает положительный исторический образ дореволюционной России, а также преемственность истории.

Среди цифровых урбанистических практик ностальгирования, формирующихся «снизу», можно назвать создание интернет-пользователями виртуального архива визуальной истории города. На сегодняшний день пользователями-энтузиастами оцифровано из официальных и личных архивов и добавлено в общий сетевой доступ множество фотографий городских и сельских пейзажей. Создаются сайты (например, pastvu.com и retromap.ru), где ретро-снимки размещают на виртуальных картах населенных пунктов. Эти фотографии составляют ценность, в первую очередь, для

специалистов в области краеведения, исторической антропологии, урбанистики, однако, могут также представлять интерес для обычных жителей города: как для тех, кто относится к старшим поколениям, так и для молодежи. Тем, кто мог быть свидетелем прежнего облика города, важно сохранить память о нем для себя и современников, чтобы не утратить чувство принадлежности к тому, что было знакомо и дорого, иными словами, составляло личную и групповую идентичность. Не менее важным представляется передача свидетельств, воспоминаний младшим поколениям, приобщение их к истории семьи и родного края, отдельных значимых мест в городе: микрорайона, где человек живет или жил, общественных пространств, мест памяти.

С целью популяризации истории города среди молодых людей – активных пользователей интернета – представляется перспективным оборудовать жилые здания и объекты городской инфраструктуры небольшими настенными табличками, на которых будет изображен индивидуальный QR-код, а также создать приложение, которое бы считывало эти коды и содержало общую базу данных о городских объектах, которую было бы резонно сформировать на основе уже имеющихся в сети изображений и информации, расположенных на сайтах, примеры которых упоминались выше. Базу данных было бы уместно сделать доступной для пополнения фотографиями и личными историями всеми желающими через приложение либо сайт, их публикация будет осуществляться с применением модерации. Модераторами могли бы быть краеведы или знатоки города.

Итак, любой пользователь смартфона смог бы считать QR-код, который направит к старым изображениям данного места с указанием даты снимка, информации о нем, а также к историям, связанным с данным местом у горожан. Под каждым QR-кодом было бы указано приложение, которое необходимо скачать или открыть, если оно уже установлено. Преимущества данной инициативы перед вышеупомянутыми сайтами заключаются в характеристиках удобства и оперативности: нет необходимости искать место на карте, достаточно открыть приложение и навести экран смартфона на QR-код. Кроме этого, про упомянутые сайты знают немногие, тогда как наличие табличек с QR-кодом на зданиях города вызвало бы любопытство у неограниченного числа горожан, особенно у наиболее активных пользователей информационных сетей.

Участие в создании виртуального городского архива и обращение к нему способно разнообразить досуг жителей города всех возрастов: горожане одновременно могут вносить личный вклад в общезначимое дело и исследовать город, знакомясь с его историей. Посредством удовлетворения ностальгического запроса, существующего в российском обществе, данная инициатива укрепит идентичность локальных сообществ, поможет сформировать участливое отношение к своему городу и району, а также активную гражданскую позицию и интерес к истории своей страны в целом.

Литература

1. 100 лет СССР: забыть нельзя вернуться? [Электронный ресурс]. Режим доступа: https:// wciom.ru/ analytical- reviews/ analiticheskii- obzor/100-let- sssr-zabyt-nelzja-vernutsja (дата обращения: 26.04.2023).
2. Романова А.П., Федорова М.М. «Советская ностальгия» несоветского цифрового поколения // Южно-российский журнал социальных наук. 2021. Т. 22. № 1. С. 6–18.
3. Сапогова Е.Е. Экзистенциально-психологические аспекты переживания «ностальгии без памяти // Проблемы современного образования. 2019. №2. С. 28–45.
4. Бойм С. Будущее ностальгии. М.: Новое литературное обозрение. 2019. 680 с.

УДК 711.123

ПОТЕНЦИАЛ СИМВОЛИЧЕСКОГО РЕСУРСА ГОРОДА В РЕВИТАЛИЗАЦИИ ЛОКАЛЬНОЙ ГОРОДСКОЙ ИДЕНТИЧНОСТИ

Судакова В.В.[1] (студент).
Научный руководитель – д.ф.н., доцент Горнова Г.В.[1]
1 – Университет ИТМО
e-mail: vladislavasudakova@yandex.ru

Аннотация
В работе рассматривается вопрос ревитализации городской идентичности и ее роли в накоплении символических ресурсов города. Цель исследования: систематизировать существующие знания о символическом ресурсе (капитале) города и установить взаимосвязь между ним и процессом «переосмысления»/«возрождения» локальной идентичности.
Ключевые слова
Символический капитал, символический капитал города, трансформационный процесс, ре-терминология, ревитализация, городская идентичность.

Постиндустриальная эпоха, кроме волны глобальной информатизации, стала временем роста значимости «символического обмена» в коммуникативных процессах, кроме того, преумножив важность глобального и локального в мире. По этой причине конкуренция между территориями зачастую идет на символическом уровне – за значимость и узнаваемость места среди других, за развитие его устойчивой идентификации во внешней среде. Одним из важнейших условием успешного развития места оказывается наличие, так называемых, «мягких ресурсов», имеющих символическое основание. Символический капитал места, его использование и увеличение, становится средством преобразования территориальной реальности, которое способно аккумулировать символическую прибыль, конвертируемую в иные формы капиталов – экономический, социальный, культурный – для территорий. Поэтому и становятся актуальными исследования, посвященные символическому потенциалу территорий, ее локальной идентичности, возможностям и проблемам бренда места, и иным другим явлениям, в которых происходит процесс территориальной символизации и непосредственно «символическое» маркирование пространства.

Первоначально в исследовании был изучен термин «символический капитал», как совокупность символических ресурсов – не только в области урбанистики, но и в пересекающихся с ней сферах науки. Термин сформулирован французским социологом и философом Пьером Бурдье в середине XX века, который выделял четыре формы капитала – экономический, культурный, социальный и символический, но подчеркивал, что «символический является особой формой – как капитал в любой его форме, представляемой символически в связи с узнаванием или неузнаванием» [1]. Данное определение показывает, что символическую форму может приобретать любой другой вид капитала вследствие влияния габитуса как социально сконструированной когнитивной способности.

В качестве носителя символического капитала может выступать и город. В таком случае он становится и бенефициаром той выгоды и прибыли, которую дает его использование.

В широком смысле любой капитал подразумевает совокупность ресурсов – тогда символический капитал города можно определить, как совокупность его символических ресурсов таких как:

1. Природные и географические особенности территории.
2. Производство уникальной продукции.
3. Известные люди (как прошлого, так и современности).

4. Визуальные компоненты территории (материальные объекты – памятники, символика, архитектурное наследие).
5. Статусы и бренды.
6. Уникальные и знаковые события, явления истории и культуры (фестивали, исторические даты, мифы, легенды) (рис. 1).

Рис. 1. Схема классификации символических ресурсов, составляющих символический капитал. Источник: создано Судаковой Владиславой (автор)

Но сами по себе ресурсы не несут никакой значимости, ценности и прибыли. Исходя из коммуникативной парадигмы и постструктуралистской логики «данные символические (преимущественно культурные) ресурсы территорий могут приобрести значимость и узнаваемость у местных жителей или у потенциальных жителей, бизнесменов, будущих туристов в ходе процессов коммуникации» [2].

Можно сказать, что символический капитал места работает как его репутационная идентичность. Но все же знания о механизмах работы символического капитала как в городских, так и в иных сферах исследований, достаточно фрагментарны. До сих пор не сформулированы методики, которые позволили бы измерить «престиж» и «репутацию», накопление которых подразумевает данный вид капитала. И именно поэтому актуальным является изучить с точки зрения формирования, накопления и преумножения символического капитала места трансформационные процессы города.

Оно имеет множество значений в разных сферах науки – проанализировав их, можно сказать, что в общем понимании трансформация – это процесс изменений, который подразумевает преобразование фундаментальных атрибутов какой-либо системы.

Трансформационные процессы в городе могут носить как непреднамеренный, так и намеренный характер. Большая часть литературы фокусируется на непреднамеренных трансформациях, которые обычно следуют за деструктивными событиями. В данной работе рассматриваются намеренные трансформационные процессы. Современные исследования формируются вокруг их трех характеристик: возможностей, механизмов и результатов. Но возможности исследователи делят еще на два типа: способность реагировать на конкретные изменения и общие способности повышать уровень развития

и благополучия. Например, потенциал адаптации к конкретным изменениям может включать разработку мер реагирования на климатические изменения. А, потенциал общих - повышать благосостояние людей посредством вовлеченности и социальной сплоченности. Именно общие адаптивные способности преднамеренных трансформаций интересуют нас.

В мировой практике существует множество понятий, созданных для определения городских трансформаций. В англоязычных источниках зачастую используются термины, начинающиеся с приставки “re”, что формирует новую “ре-терминологию” [3]. Слова, образованные с ее помощью, обозначают новую интерпретацию, процесс воссоздания, ссылающийся на что-то уже существовавшее. В рамках данной работы был взят термин ревитализации, который лучших других трансформаций описывает процесс изменения среды, затрагивающий изменение и ее идентичности.

В урбанистике термин «ревитализация» обозначает процесс воссоздания и оживления городского пространства. Классическим примером является переосмысление старых промышленных территорий и создание в них креативных кластеров. Но ревитализация городской среды может опираться не только на ее материальные объекты. Так история возрождения Коломны, началась с рецепта коломенской пастилы. Цепная реакция запустила интерес к наследию города и его капитализацию – в городских преобразованиях.

Для того, чтобы понять роль городской идентичности в трансформационных процессах города, необходимо деконструировать сам термин. Понятие городская идентичность означает «устойчивое представление о себе как о жителе именно этого города» [4]. Ее структура включает четыре компонента: когнитивный (отвечающий за отношение человека к объективным характеристикам среды), аффективный (являющийся отражением эмоционально-чувственного восприятия факторов первого), ценностно-смысловой (как отражение мнения своих жителей во времени в объективированных формах городской культуры) и практический (отвечающий за миграционные установки горожан).

Однако сопоставив классификацию символических ресурсов территории и структуру городской идентичности, мы видим, что каждый из его видов лишь при условии «переживания и принятия» горожанами, может назваться символическим. Это доказывает гипотезу, что городская идентичность инициирует создание символических ресурсов.

Процесс ревитализации может происходить по отношению к городской идентичности насильственно, «руками» приглашенных экспертов. В таком случае успешность зависит от принятия предложенных трансформаций местным сообществом. Примером насильственной практики может послужить искусственно созданный бренд «Ярославль – столица Золотого кольца». С туристической точки зрения, возможно, статус увеличил турпоток, но точно не повлиял на увеличение символического капитала. Местные жители высмеивают новый статус, не понимая, как у кольца может быть «столица». Историки задаются вопросом, почему именно Ярославль.

Успешным примером ревитализации территориального бренда, выполненной внешним экспертом, можно назвать кейс Урюпинска – как «Столицы российской провинции». Но важно отметить, что непосредственный разработчик бренда, провел множество полевых исследований города и его горожан, аккумулируя субъектность.

Второй вариант – когда ревитализация городской идентичности идет первоочередно. За осознанием человека себя жителем некоторого города, идет осознание его причастным к ранее существовавшим традициям и обычаям, а как следствие возрождение как материальной, так и нематериальной городской среды.

Кейс проекта Текстиль в Ярославле – пример этой практики. За ревитализацией локальной идентичности местными жителями, желающими изучить историю старой ткацкой мануфактуры, началась трансформация городской среды – создание

культурного центра, музея, новых практик и традиций. Инвестиционный интерес вырос. Появилась новая символическая связь района с горожанами.

Примером кейса, где ревитализация городской идентичности началась с одного жителя, является социальный эксперимент в Данилове. К созданию креативной среды города, местным экспертом были привлечены школьники. На первых этапах они были пассивными исполнителями, но спустя время сами стали инициаторами новых смыслов и проектов.

С целью сокращения временных, экономических и иных издержек, на основе существующих кейсов, мы можем разработать алгоритм оценки потенциала конвертируемости символического капитала в проектные инициативы по территориальному развитию (рис. 2).

Рис. 2. Схема алгоритма оценки потенциала конвертируемости символического капитала в проектных инициативах по территориальному развитию.
Источник: создано Судаковой Владиславой (автор)

Для этого необходимо выявить аккумулируемые средой ресурсы. За этим, инициировать процесс ревитализации, то есть оценить степень «принятия» и «отождествления» местным сообществом новых символических маркеров. И в итоге, отсеять те ресурсы, которые не будут им «приняты». Полученный символический капитал может быть использован в разнообразных проектах по территориальному развитию как малых единиц – поселков, деревень, так и целых регионов: в мастер-планах, стратегиях территориального развития, брендинге территории и в во многом другом. Это поможет значительно повысить конкурентоспособность городов в борьбе за человеческие и иные ресурсы, создав новые «символические» маркеры связи между личностью и городом.

В дальнейшем сформулированные тезисы лягут в основу разработки методики исследования механизмов формирования символического капитала территорий и оценки его конвертируемости в проектные инициативы по территориальному развитию.

Литература

1. Бурдье П. Практический смысл. СПб: Алетейя. 2001. 562 с.
2. Федотова Н.Г. Символический капитал места: понятие, особенности накопления, методики исследования. // Вестн. Том. гос. ун-та. Культурология и искусствоведение. 2018. №29.
3. Гунько М.С., Пивовар Г.А., Аверкиева К.В. Ревитализация в малых городах европейской России (на примере Боровичей, Выксы, Ростова) // Известия РАН. Серия географическая. 2019. №5. С. 18-31.
4. Горнова Г.В. Визуально-семиотические аспекты городской идентичности // Праксема. Проблемы визуальной семиотики. 2019. № 6 (21). С. 62–74.

СОДЕРЖАНИЕ

БИОТЕХНОЛОГИИ……... 5

Беспалая Д.С.
ОБОСНОВАНИЕ ВЫБОРА ИНГРЕДИЕНТОВ ДЛЯ ПРОИЗВОДСТВА ФУНКЦИОНАЛЬНОЙ СНЕКОВОЙ ПРОДУКЦИИ ДЛЯ ПРОФИЛАКТИКИ ОНКОЛОГИЧЕСКИХ ЗАБОЛЕВАНИЙ……. 6

Гаврецкая А.В., Шестель А.А.
ОСОБЕННОСТИ РАЗРАБОТКИ ЭКОЛОГИЧЕСКИХ ДЕКЛАРАЦИЙ ДЛЯ РЕСУРСОСБЕРЕГАЮЩИХ КОМПЛЕКСНЫХ ТЕХНОЛОГИЙ ПЕРЕРАБОТКИ БИОСЫРЬЯ……... 11

Данилюк М.А.
ТЕХНОЛОГИЯ ПОЛУЧЕНИЯ РЫБНЫХ ПОЛУФАБРИКАТОВ СО СБАЛАНСИРОВАННЫМ СОСТАВОМ ДЛЯ ГЕРОНТОЛОГИЧЕСКОГО ПИТАНИЯ……. 15

Дерягина А.С.
ТЕХНОЛОГИЯ ФУНКЦИОНАЛЬНОГО КОФЕЙНОГО ПРОДУКТА С ИСПОЛЬЗОВАНИЕМ ИНУЛИНА…… 18

Жернякова А.В.
СРАВНЕНИЕ РАСТИТЕЛЬНЫХ И БАКТЕРИАЛЬНЫХ ИСТОЧНИКОВ ВИТАМИНА В12……... 22

Иващенко Л.В., Амбарцумов Т.Г., Захарова У.Е.
АНАЛИЗ ВОЗМОЖНОСТИ ПРОИЗВОДСТВА ГОРОХОВОГО ИЗОЛЯТА В РОССИИ…… 26

Йимер Г.А.
РАЗРАБОТКА РЕЦЕПТУР БЕЗГЛЮТЕНОВЫХ МАКАРОННЫХ ИЗДЕЛИЙ…….. 29

Коновалов Д.С.
ЗЕЛЁНАЯ ХИМИЯ КАК ОСНОВА ДЛЯ ПОЛУЧЕНИЯ ПОЛИУРЕТАНО-КОМПОЗИТНОГО МАТЕРИАЛА ИЗ БИО-СЫРЬЯ…….. 32

Наумова А.В., Утсаль В.А., Яккола А.Н., Наумов И.А., Островидова Е.В.
ПРИМЕНЕНИЕ МЕТОДА ХРОМАТО МАСС- СПЕКТРОМЕТРИИ ДЛЯ ОЦЕНКИ БЕЗОПАСНОСТИ РЫБЫ И РЫБНЫХ ПОЛУФАБРИКАТОВ…… 37

Сергуненкова О.Б.
ОБОГАЩЕНИЕ ДЕСЕРТА НА ОСНОВЕ ВТОРИЧНОГО МОЛОЧНОГО СЫРЬЯ ПРОБИОТИЧЕСКИМ ШТАММОМ…… 39

Ситникова А.О., Капралова В.М.
ИМПЕДАНСНЫЕ ХАРАКТЕРИСТИКИ ПОЛИМЕРНОГО ВОЛОКНИСТОГО КОМПОЗИТНОГО МАТЕРИАЛА……. 44

Фофанов В.С., Мишин К.К., Скворцов Н.В.
β-ЦИКЛОДЕКСТРИН – СИСТЕМА ДОСТАВКИ
ПОЛИФЕНОЛОВ НОВОГО ПОКОЛЕНИЯ…… 49

Фролова К.О.
АНТОЦИАНЫ И БЕТАЛАИНЫ
КАК ИНДИКАТОРЫ ПОРЧИ МЯСНЫХ ПРОДУКТОВ…….. 55

Шестель А.А., Гаврецкая А.В.
РАЗРАБОТКА ЭКОЛОГИЧЕСКОЙ ДЕКЛАРАЦИИ НА ПРИМЕРЕ
РЫБНЫХ ПОЛУФАБРИКАТОВ СО СЪЕДОБНЫМ ПОКРЫТИЕМ…… 61

ИНФОРМАЦИОННАЯ БЕЗОПАСНОСТЬ……. 65
Amanuel Estifanos Asfha, Zakoldaev D.A., Abhishek Vaish
INFORMATION SECURITY RISKS AT AUTOMATED
INDUSTRIAL INFORMATION SYSTEM FIRMS…… 66

Асьминкин Ф.А.
ОБЗОР МЕТОДОВ МОНИТОРИНГА ВСТРАИВАЕМЫХ
СИСТЕМ НА ПРИМЕРЕ КОНТРОЛЛЕРА ЛИНИЙ
СИСТЕМЫ АВАРИЙНОГО ОПОВЕЩЕНИЯ…… 73

Блинов А.В.
ПРОБЛЕМЫ БЕЗОПАСНОСТИ MESH-СЕТЕЙ
БЕСПИЛОТНЫХ ЛЕТАТЕЛЬНЫХ АППАРАТОВ…… 82

Huiyao Dong
MULTI-TASK DEEP LEARNING-BASED INTRUSION
DETECTION MODEL FOR IOT NETWORK…… 86

Ишутина Е.
ОСОБЕННОСТИ ОБЕСПЕЧЕНИЯ БЕЗОПАСНОСТИ
ПОЛЬЗОВАТЕЛЬСКИХ ДАННЫХ WEB-СЕРВИСОВ…….. 91

Кирьянова А.П., Давыдов В.В.
О ПОСТРОЕНИИ И ПРИМЕНЕНИИ КРИПТОГРАФИЧЕСКИХ
ХЭШ-ФУНКЦИЙ, ОСНОВАННЫХ НА ГРАФАХ…… 95

Кича И.В., Тимкин А.К.
ПРИМЕНЕНИЕ ДРЕВОВИДНЫХ СТРУКТУР
ДЛЯ РАСПРЕДЕЛЕННОГО ХРАНЕНИЯ ДАННЫХ…… 98

Коновалов П.И., Андрушкевич Д.В.
ОСОБЕННОСТИ БИНАРНОЙ ЭКСПЛУАТАЦИИ МЕТОДА
SIGRETURN ORIENTED PROGRAMMING (SROP)…….. 103

Круглов Д.Д., Морозова Е.М., Пирожникова О.И.
АНАЛИЗ ПАТЕНТНЫХ ИССЛЕДОВАНИЙ
И НАУЧНО-ТЕХНИЧЕСКОЙ ЛИТЕРАТУРЫ
ПО ВОПРОСАМ ЭЛЕКТРОМАГНИТНЫХ ЗАМКОВ…….. 107

Ксенофонтов В.С.
РАЗРАБОТКА ПРОГРАММНОГО СРЕДСТВА ДЛЯ ОБНАРУЖЕНИЯ АНОМАЛИЙ В ФУНКЦИОНИРОВАНИИ ПРОЦЕССОВ НА ОСНОВЕ ТРАССИРОВКИ СИСТЕМНЫХ ВЫЗОВОВ В ОПЕРАЦИОННОЙ СИСТЕМЕ LINUX…… 111

Лаврентьев С.А., Андрушкевич С.С., Дудкин А.С.
АЛГОРИТМЫ ОБФУСКАЦИИ ПРОГРАММ, РАЗРАБАТЫВАЕМЫХ НА ЯЗЫКАХ C/C++……. 118

Ласкус Е.О.
ОЦЕНКА ЗАЩИЩЕННОСТИ ЦИФРОВЫХ ДВОЙНИКОВ ПАЦИЕНТОВ…… 122

Линкер М.А.
МЕТОДИКА АУТЕНТИФИКАЦИИ ПОЛЬЗОВАТЕЛЕЙ МОБИЛЬНЫХ УСТРОЙСТВ НА ОСНОВЕ НЕПРЕРЫВНОГО АНАЛИЗА ПОВЕДЕНЧЕСКИХ ХАРАКТЕРИСТИК…… 127

Меренков Д.Н.
РАЗРАБОТКА МОДЕЛИ СИСТЕМЫ АВТОМАТИЗИРОВАННОЙ ПРОВЕРКИ ТЕКСТОВ НА ПРЕДМЕТ СОДЕРЖАНИЯ ИНФОРМАЦИИ ОГРАНИЧЕННОГО ДОСТУПА…… 132

Ниткин И.С., Кирьянова А.П.
ПОСТ-КВАНТОВЫЕ ПРОТОКОЛЫ ВЫРАБОТКИ ОБЩЕГО СЕКРЕТНОГО КЛЮЧА…….. 140

Осипов И.Н.
СИСТЕМЫ ЗАЩИТЫ ВЕБ-ПРИЛОЖЕНИЙ……. 145

Прыгунов М.И.
БЕЗОПАСНАЯ СИСТЕМА ОБМЕНА МГНОВЕННЫМИ СООБЩЕНИЯМИ С ПОВЫШЕННОЙ КОНФИДЕНЦИАЛЬНОСТЬЮ ДАННЫХ…… 150

Stepanian A.A.
PORT SCANNING DETECTION USING MACHINE LEARNING MODELS……. 153

Тимкин А.К., Кича И.В.
ФАКТОРЫ ВЛИЯЮЩИЕ НА СИСТЕМЫ БЕЗОПАСНОСТИ В МУЛЬТИАГЕНТНОЙ СИСТЕМЕ АВТОНОМНЫХ НЕОБИТАЕМЫХ ПОДВОДНЫХ АППАРАТОВ…….. 159

Токаев А.Д.
КОНТРОЛЬ КАЧЕСТВА ДАННЫХ С ПОМОЩЬЮ БЛОКЧЕЙН ТЕХНОЛОГИЙ……. 163

Хазагаров А.А., Воробьева А.А.
ИСПОЛЬЗОВАНИЕ ТРАНСФОРМЕРОВ ДЛЯ ЛИНГВИСТИЧЕСКОЙ ИДЕНТИФИКАЦИИ АВТОРА……... 168

Чернов Р.И., Горошков В.А., Островский Д.П., Еремук В.В.
РАЗРАБОТКА ЛАБОРАТОРНОГО СТЕНДА ДЛЯ
ИССЛЕДОВАНИЯ ОПТИЧЕСКОЙ СИСТЕМЫ СВЯЗИ…….. 171

Якимов Я.Д.
РАЗРАБОТКА АРХИТЕКТУРЫ СИСТЕМЫ МОНИТОРИНГА
БЕСПИЛОТНОГО МАНЕВРОВОГО ЛОКОМОТИВА……. 174

ИНФОКОММУНИКАЦИОННЫЕ ТЕХНОЛОГИИ…… 180
Dayoub A., Filyanin I.V.
THE DEVELOPMENT OF IOT APPLICATION
ON EDGE COMPUTING PLATFORM…….. 181

Двойникова А.А.
АНАЛИЗ ВОВЛЕЧЕННОСТИ И ЭМОЦИЙ
СОБЕСЕДНИКОВ ВИРТУАЛЬНОЙ КОММУНИКАЦИИ…… 185

Инкин М.Е., Столбов М.Б.
ИССЛЕДОВАНИЕ ВЛИЯНИЯ УСЛОВИЙ
ПРОВЕДЕНИЯ ЗВУКОЗАПИСЕЙ НА РЕЗУЛЬТАТ
РАСПОЗНАВАНИЯ ДИКТОРОВ ПРИ ИСПОЛЬЗОВАНИИ
АВТОМАТИЧЕСКОГО МЕТОДА СРАВНЕНИЯ…….. 191

Кузнецова Е.Г.
РАЗРАБОТКА АВТОМАТИЗИРОВАННОЙ СИСТЕМЫ АУДИТА
ПОКАЗАТЕЛЕЙ ЭФФЕКТИВНОСТИ БИЗНЕС-ПРОЦЕССОВ……... 196

Кумашев Э., Аминов Н.С.
РАСПОЗНАВАНИЕ ЖЕСТОВ В УМНОМ ДОМЕ……. 201

Петухов С.А.
РАЗРАБОТКА ИНФОРМАЦИОННОЙ СИСТЕМЫ ЭЛЕКТРОННОГО
ФОТОАРХИВА КРАЕВЕДЧЕСКОЙ ИНФОРМАЦИИ……. 212

Попов И.А.
УЯЗВИМОСТИ СМАРТ-КОНТРАКТОВ SOLIDITY……. 215

Рыбкин Я.С.
ПРОБЛЕМЫ АНАЛИЗА КРИПТОМАЙНИНГА И ИХ РЕШЕНИЕ…… 224

Самородова М.Э.
ОБНАРУЖЕНИЕ СЛЕДОВ АУДИОМОНТАЖА ПРИ
СПУФИНГ-АТАКЕ НА СИСТЕМУ ГОЛОСОВОЙ БИОМЕТРИИ……. 231

Холькина А.А.
ИСПОЛЬЗОВАНИЕ НЕЙРОСЕТЕВОГО МОДЕЛИРОВАНИЯ
В СИСТЕМАХ КОНТРОЛЯ И УПРАВЛЕНИЯ ДОСТУПОМ
НА ОСНОВЕ ЛИЦЕВОЙ БИОМЕТРИИ……. 236

Чубур С.В., Сабиров Т.Р.
СИСТЕМА ГЕОПРОСТРАНСТВЕННОЙ ПРИВЯЗКИ
И ВИЗУАЛИЗАЦИИ ДАННЫХ ИЗ
ИНФОРМАЦИОННО-НОВОСТНЫХ ИСТОЧНИКОВ……. 246

Шеметов Ф.А.
УСКОРЕНИЕ ИНФЕРЕНСА НЕЙРОННЫХ
СЕТЕЙ ПРЯМОГО РАСПРОСТРАНЕНИЯ…… 252

ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ…… 262
Абрегова З.Х.
СПЕЦИФИКА СБОРА ДАТАСЕТА ПРИ СОЗДАНИИ ЦИФРОВОГО
ПОМОЩНИКА ДЛЯ СЛАБОВИДЯЩИХ ЛЮДЕЙ……... 263

Авдюшина А.Е., Крисанов Р.В.
СБОР ГЕОЗАВИСИМЫХ ДАННЫХ ИЗ НОВОСТНЫХ РЕСУРСОВ…….. 267

Алексеев А.С.
RCPSP RESOURCE-CONSTRAINED PROJECT
SCHEDULING PROBLEM (ЗАДАЧА ПЛАНИРОВАНИЯ
ПРОЕКТОВ В УСЛОВИЯХ ОГРАНИЧЕННОСТИ РЕСУРСОВ)……. 271

Бетеня Д.С.
УСТРОЙСТВО ДЛЯ ПОМОЩИ СЛАБОВИДЯЩИМ
ЛЮДЯМ НА БАЗЕ ТЕХНОЛОГИИ EDGE AI…… 275

Вандышев К.А., Горшков Р.Д., Коныхов В.С.
ПОХОД К СОЗДАНИЮ ИНТЕЛЛЕКТУАЛЬНОЙ
СИСТЕМЫ ПОСТРОЕНИЯ 3Д МОДЕЛИ ПОМЕЩЕНИЙ
ПРИ ПОМОЩИ БЕСПИЛОТНЫХ ЛЕТАТЕЛЬНЫХ АППАРАТОВ…….. 280

Вологина Е.А.
ГЕНЕРАЦИЯ ИЗОБРАЖЕНИЙ НА ОСНОВЕ СЕМАНТИЧЕСКИ
НАГРУЖЕННЫХ БОЛЬШИХ ХУДОЖЕСТВЕННЫХ ТЕКСТОВ…….. 283

Газизулин А.Ф.
ПРОГНОЗИРОВАНИЕ ЦЕН АКЦИЙ НА БИРЖЕ
С ИСПОЛЬЗОВАНИЕМ РАЗРОЗНЕННЫХ ИСТОЧНИКОВ……... 288

Дощенко А.И.
МЕТОДЫ ВЫДЕЛЕНИЯ РЕКЛАМЫ И УТОЧНЕНИЯ ЕЁ
КАТЕГОРИЙ НА ОСНОВЕ АНАЛИЗА НЕСТРУКТУРИРОВАННЫХ
ДАННЫХ СОЦИАЛЬНЫХ СЕТЕЙ……. 293

Дубовской А.А.
ОБЗОР МЕТОДОВ ИЗВЛЕЧЕНИЯ
ГРАФОВ ЗНАНИЙ ИЗ ИЗОБРАЖЕНИЙ……. 300

Егорова Е.В., Чеканова Е.Д.
РАЗРАБОТКА ЧАТ-БОТА ДЛЯ ФОРМИРОВАНИЯ
РЕКОМЕНДАЦИЙ МЕСТ ОБЩЕСТВЕННОГО ПИТАНИЯ
С УЧЕТОМ ПРЕДПОЧТЕНИЙ ПОЛЬЗОВАТЕЛЯ..... 304

Зайчикова А.И.
КЛАСТЕРИЗАЦИЯ ВАКАНСИЙ ПО КЛЮЧЕВЫМ
НАВЫКАМ КАНДИДАТА В СФЕРЕ
ИНФОРМАЦИННЫХ ТЕХНОЛОГИЙ…… 309

Здоровеннов М.Р.
ПРОЕКТИРОВАНИЕ КЛИЕНТ-СЕРВЕРНОГО ПРИЛОЖЕНИЯ
ДЛЯ РАБОТЫ С ФИНАНСОВЫМИ ЗАДОЛЖЕННОСТЯМИ…….. 315

Зиманова Е.А.
АНАЛИЗ ПРЕДСТАВЛЕННЫХ НА РЫНКЕ CRM-СИСТЕМ
ДЛЯ УЧЕБНЫХ ЯЗЫКОВЫХ ЦЕНТРОВ…….. 319

Коровин А.И.
ПРОЕКТИРОВАНИЕ ФУНКЦИОНАЛЬНОСТИ
И ДИЗАЙНА МОБИЛЬНОГО ПРИЛОЖЕНИЯ
ДЛЯ КОЛЛАБОРАЦИИ МУЗЫКАНТОВ…… 324

Краснов Т.А., Компаниец Р.И., Дудкин А.С.
РАЗРАБОТКА МОДУЛЯ БЕЗОПАСНОСТИ ИСКУССТВЕННОГО
ИНТЕЛЛЕКТА НА ОСНОВЕ ОГРАНИЧЕНИЙ ПОВЕДЕНИЯ
СИСТЕМЫ ПОДДЕРЖКИ И ПРИНЯТИЯ РЕШЕНИЙ……. 327

Кривоносов Е.Д., Нечкасова О.А., Перл И.А.
СРАВНЕНИЕ ЭФФЕКТИВНОСТИ ПРОГНОЗИРОВАНИЯ
ВРЕМЕННЫХ РЯДОВ С ПОМОЩЬЮ МАШИННОГО ОБУЧЕНИЯ
И РАЗЛИЧНЫХ ВИДОВ НЕЙРОННЫХ СЕТЕЙ……... 330

Лакиза А.Н.
КЛАССИФИКАЦИЯ ЖАНРА МУЗЫКАЛЬНОГО
АЛЬБОМА ПО ОБЛОЖКЕ С ПОМОЩЬЮ
СВЕРТОЧНЫХ НЕЙРОННЫЙХ СЕТЕЙ…… 335

Бодунов Г.А., Лобань И.В., Шишкарев З.А., Дудкин А.С.
ПРИМЕНЕНИЕ ТЕХНОЛОГИЙ ИСКУССТВЕННОГО
ИНТЕЛЛЕКТА В ЗАДАЧАХ МОНИТОРИНГА И АГРЕГИРОВАНИЯ
СОБЫТИЙ ИЗ ОТКРЫТЫХ ИСТОЧНИКОВ С ПРИВЯЗКОЙ
К ГЕОГРАФИЧЕСКИМ КООРДИНАТАМ……. 339

Макаров К.Д.
ПРОЕКТИРОВАНИЕ МОБИЛЬНОГО ПРИЛОЖЕНИЯ ДЛЯ КОНТРОЛЯ
И ПЛАНИРОВАНИЯ ТРЕНИРОВОЧНОГО ПРОЦЕССА…… 345

Менисов А.Б., Захаров О.О.
СТАТИЧЕСКИЙ АНАЛИЗ ФРЕЙМВОРКОВ МАШИННОГО ОБУЧЕНИЯ…….. 349

Меньщиков М.А.
АНАЛИЗ ДООБУЧЕННЫХ ЯЗЫКОВЫХ МОДЕЛЕЙ
ДЛЯ ЗАДАЧИ ГЕНЕРАЦИИ ДИСТРАКТОРОВ…… 353

Мурзина А.А.
ПРОГНОЗИРОВАНИЕ ГЛИКЕМИИ В УСЛОВИЯХ МАЛЫХ ВЫБОРОК…… 360

Никифорова А.Д., Роговой В., Суров И.А.
КЛАССИФИКАТОР ПРЕДЛОЖЕНИЙ АНГЛИЙСКОГО ЯЗЫКА
ДЛЯ ПРОЦЕССНО-СМЫСЛОВОГО АНАЛИЗА ТЕКСТОВ…… 364

Раздьяконов А.А.
ЭВРИСТИКА АЛГОРИТМОВ КЛАССИФИКАЦИИ В ЗАДАЧЕ
ОПРЕДЕЛЕНИЯ УБЫТОЧНЫХ СЕГМЕНТОВ ПОРТФЕЛЯ ОСАГО…… 368

Савко С.В., Виноградова А.П., Попкова Я.В.
БОТ-НАВИГАТОР ПО ПРОГРАММАМ
ОБУЧЕНИЯ РЭУ ИМ. Г.В. ПЛЕХАНОВА……. 373

Сердюков А.Н., Замятин А.А.
РАЗРАБОТКА ИНСТРУМЕНТА ДЛЯ ИНТЕРАКТИВНОЙ
ВИЗУАЛИЗАЦИИ И СКАФФОЛДИНГА ГЕНОМНЫХ
СБОРОК ПО ДАННЫМ HI-C…… 380

Сташишин А.И.
ИСПОЛЬЗОВАНИЕ МОБИЛЬНЫХ ПРИЛОЖЕНИЙ
ДЛЯ ОТКАЗА ОТ ВРЕДНЫХ ПРИВЫЧЕК…… 386

Тирский Б.Д., Рунушкина О.Р., Верко Р.А.
РАЗРАБОТКА АЛГОРИТМА ДЛЯ РАСШИФРОВКИ
И ОПРЕДЕЛЕНИЯ АББРЕВИАТУР В ТЕКСТЕ……. 391

Тихонов Л.А.
РАЗРАБОТКА ПРОЕКТА АВТОМАТИЗАЦИИ
РАЗВЕРТЫВАНИЯ ИНФРАСТРУКТУРЫ ДЛЯ РЕШЕНИЯ
ЗАДАЧ МАШИННОГО ОБУЧЕНИЯ…….. 396

Улизько М.В., Добренко Д.А., Голубев А.А.
ПРОГНОЗИРОВАНИЕ ТРЕНДОВ В РАЗРЕЖЕННЫХ
ВРЕМЕННЫХ РЯДАХ С ПРИМЕНЕНИЕМ НЕЙРОННЫХ СЕТЕЙ…….. 400

Фам Т.А., Субботин А.Н.
НАБЛЮДЕНИЕ ЗА ПАЦИЕНТОМ
С ПРИМЕНЕНИЕМ ТЕХНОЛОГИЙ ОКРУЖАЮЩЕГО
ИНТЕЛЛЕКТА И УСТРОЙСТВ ИНТЕРНЕТА ВЕЩЕЙ…… 403

Фролов М.А.
АНАЛИЗ ВЗАИМОДЕЙСТВИЯ РАСПРЕДЕЛЕННЫХ
СЕРВИСОВ С ИСПОЛЬЗОВАНИЕМ ТЕЛЕМЕТРИИ…… 408

Черников К.М.
ОПТИМИЗАЦИЯ НЕЙРОСЕТЕВЫХ АЛГОРИТМОВ
ПУТЕМ АВТОМАТИЗАЦИИ ВЫЧИСЛЕНИЙ ПОКАЗАТЕЛЕЙ
ГРАДИЕНТНОГО СПУСКА И ФУНКЦИИ АКТИВАЦИИ…….. 412

Шутов Д.Э.
ПРОЕКТИРОВАНИЕ МОДУЛЯ ПОЛЬЗОВАТЕЛЕЙ
В СИСТЕМЕ УПРАВЛЕНИЯ УЧЕБНЫМИ ПРОЕКТАМИ…….. 417

Юрьев М.Р.
ТЕХНОЛОГИИ СИМУЛЯЦИИ ПОВЕРХНОСТИ ОКЕАНА
В РЕЖИМЕ РЕАЛЬНОГО ВРЕМЕНИ В ИГРОВЫХ ДВИЖКАХ…… 420

ЦИФРОВАЯ УРБАНИСТИКА……. 424
Алексеевская А.А.
ОЦЕНКА ВЛИЯНИЯ ВЕТРОВОГО РЕЖИМА
НА ОПЫТ ИСПОЛЬЗОВАНИЯ ТЕРРИТОРИИ
ЖИЛЫХ ДВОРОВ САНКТ-ПЕТЕРБУРГА……. 425

Апухтина Ю.В.
СОЗДАНИЕ ЕДИНОЙ УНИВЕРСАЛЬНОЙ БАЗЫ
ДАННЫХ ТЕРРИТОРИЙ ПРИКРЕПЛЕНИЯ ШКОЛ
К ЖИЛЫМ ДОМАМ В ГОРОДАХ РОССИИ……. 431

Бумагин Д.Г.
ЦИФРОВЫЕ МЕТОДЫ ВЗАИМОДЕЙСТВИЯ ПРИ СОУЧАСТВУЮЩЕМ
ПРОЕКТИРОВАНИИ В ПЛАНИРОВАНИИ СИСТЕМ
ОБЩЕСТВЕННОГО ТРАНСПОРТА…… 438

Ваврина С.С.
КОМПЛЕКСНЫЙ ПОДХОД К ПРОЕКТИРОВАНИЮ
ОБЩЕСТВЕННЫХ ПРОСТРАНСТВ…….. 443

Дрожжин А.И., Назаренко Е.А.
ОПРЕДЕЛЕНИЕ СТРАТЕГИИ УПРАВЛЕНИЯ ОКН НА ОСНОВЕ
ОЦЕНКИ ЕГО СОСТОЯНИЯ И ОКРУЖАЮЩЕГО КОНТЕКСТА……. 447

Ефименко В.П.
ЦИФРОВЫЕ ТЕХНОЛОГИИ В ИССЛЕДОВАНИИ
ХУДОЖЕСТВЕННЫХ ЦЕННОСТЕЙ……. 453

Катынсус А.В.
ИССЛЕДОВАНИЕ МЕТОДОВ ВЫЯВЛЕНИЯ
ТОЧЕК ОБЩЕСТВЕННОГО ИНТЕРЕСА
НА ОСНОВЕ ОТКРЫТЫХ ДАННЫХ……. 457

Клюева А.В.
МЕТОД ПЛАНИРОВАНИЯ ВЕЛОМАРШРУТНОЙ СЕТИ ГОРОДА……. 462

Печенкина Е.В.
АНАЛИЗ ДОРОЖНО-ТРАНСПОРТНЫХ ПРОИСШЕСТВИЙ
НА ПРИМЕРЕ Г. САНКТ-ПЕТЕРБУРГА…… 466

Пластун Ю.А.
ПРАКТИКИ НОСТАЛЬГИРОВАНИЯ В ЦИФРОВОМ
ПРОСТРАНСТВЕ СОВРЕМЕННОГО ГОРОДА…… 470

Судакова В.В.
ПОТЕНЦИАЛ СИМВОЛИЧЕСКОГО РЕСУРСА
ГОРОДА В РЕВИТАЛИЗАЦИИ ЛОКАЛЬНОЙ
ГОРОДСКОЙ ИДЕНТИЧНОСТИ…… 473

XII Конгресс молодых ученых
Сборник научных трудов

Том 2

Редакционно-издательский отдел Университета ИТМО
Зав. РИО Н.Ф. Гусарова
Дизайн обложки Отдел развития бренда Университета ИТМО

Вёрстка Я.Я. Платунова
Подписано к печати 09.12.2023
Заказ № 4762 от 09.12.2023
Тираж 100 экз.

Печатается в авторской редакции

Отпечатано: Учреждение «Университетские коммуникации»
199034, Санкт-Петербург, В.О., Биржевая линия, 16