УФИМСКИЙ УНИВЕРСИТЕТ

приоритет2030^
Лидерами становятся

ITIDS'2025

XI Международная научная конференция

ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ ИНТЕЛЛЕКТУАЛЬНОЙ ПОДДЕРЖКИ ПРИНЯТИЯ РЕШЕНИЙ

ТРУДЫ - Том 1



The 11th International Scientific Conference on

Information Technologies for Intelligent Decision Making Support

Proceedings v. 1

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РФ
УФИМСКИЙ УНИВЕРСИТЕТ НАУКИ И ТЕХНОЛОГИЙ

ITIDS'2025

Труды XI Международной научной конференции

Информационные технологии интеллектуальной поддержки принятия решений

Уфа, Россия
11–13 ноября, 2025

Том 1

XI Международная научная конференция "Информационные технологии интеллектуальной поддержки принятия решений"

MINISTRY OF SCIENCE AND HIGHER EDUCATION OF THE RUSSIAN FEDERATION
UFA UNIVERSITY OF SCIENCE AND TECHNOLOGY

ITIDS'2025

Proceedings of the 11th International Scientific Conference on

Information Technologies for Intelligent Decision Making Support

Ufa, Russia
November 11–13, 2025

Volume 1

The 11th International Scientific Conference on Information Technologies for Intelligent Decision Making Support

УДК 004.9
ББК 16.2
И74

Организаторы конференции
Уфимский университет науки и технологий (УУНиТ), Уфа, Россия;

Государственный комитет Республики Башкортостан по науке и высшему образованию;
Чандигархский университет (ЧУ, Индия);
Ваальский технологический университет (ВУТ, ЮАР);
Российская ассоциация искусственного интеллекта (РАИИ);
Башкирское отделение Научного Совета РАН по методологии
искусственного интеллекта и когнитивным исследованиям (БО НСМИИ и КИ РАН)

Публикуется по решению кафедры вычислительной математики
и кибернетики УУНиТ.
Протокол № 6 от 03.02.2026 г.

Редакционная коллегия:
д-р техн. наук, профессор **Г. Р. Шахмаметова** (*отв. редактор*) (УУНиТ, Уфа, Россия);
д-р физ.-мат. наук, профессор **О. А. Кривошеева** (УУНиТ, Уфа, Россия);
д-р техн. наук, профессор **А. Г. Кравец** (ВолгГТУ, Волгоград, Россия);
д-р техн. наук, профессор **А. В. Мельников** (ЮНИИ ИТ, Ханты-Мансийск, Россия);
ст. преподаватель **А. В. Ковтуненко** (*техн. редактор*) (УУНиТ, Уфа, Россия)

Информационные технологии интеллектуальной поддержки принятия решений:
И74 труды XI Международной научной конференции (г. Уфа, 11–13 ноября 2025 г.) / отв. ред. Г. Р. Шахмаметова / В 2-х т. Т. 1 [Электронный ресурс] / Уфимск. ун-т науки и технологий. – Уфа: Уфимский университет, 2026. – 386 с. – URL: https://uust.ru/media/documents/ digital-publications/2026/053.pdf – Загл. с титула экрана.
ISBN 978-5-7477-6354-8
Том 1: ISBN 978-5-7477-6355-5

Конференция является международной площадкой для обмена научными достижениями между исследовательскими сообществами Российской Федерации, Республики Индия, Южно-Африканской Республики в области информационных технологий интеллектуальной поддержки принятия решений.

Статьи, включенные в Труды, представлены в авторской редакции.

УДК 004.9
ББК 16.2

ISBN 978-5-7477-6355-5 (том 1)
ISBN 978-5-7477-6354-8

© Уфимский университет, 2026

UDC 004.9
BBK 16.2
I-74

Organized by
Ufa University of Science and Technology (UUST), Ufa, Russia

State Committee of the Republic of Bashkortostan for Science and Higher Education
Chandigarh University (CHU, India)
Vaal University of Technology (VUT, RSA);
Russian Association of Artificial Intelligence (RAAI)
Bashkir Branch of the Scientific Council of the Russian Academy of Sciences on Methodology of Artificial Intelligence and Cognitive Research

Published by decision of the Department of Computational Mathematics and Cybernetics UUST Protocol No. 6 dated 03/02/2026.

Editorial team:

Giuzel Shakhmametova	UUST, Ufa, Russia
Olesya Krivosheeva	UUST, Ufa, Russia
Alla Kravets	VolgGTU, Volgograd, Russia
Andrey Melnikov	URIIT, Khanty-Mansiysk, Russia
Alexandra Kovtunenko	UUST, Ufa, Russia

Information Technologies for Intelligent Decision Making Support:
I-74 Proceedings of the 11th International Scientific Conference "Information Technologies for Intelligent Decision Making Support" (November 11–13, Ufa, Russia) / ed. G. R. Shakhmametova / In 2 volumes. Vol. 1 [Electronic resource] / Ufa University of Science and Technology. – Ufa: Ufa University, 2026. – 386 p. – URL: https://uust.ru/ media/documents/digital-publications/2026/053.pdf – Title from the screen title.
ISBN 978-5-7477-6354-8
Volume 1: ISBN 978-5-7477-6355-5

The Conference is high quality international meeting providing a forum for exchange of scientific achievements, an international platform for the exchange of scientific achievements between the research communities of the Russian Federation, the Republic of India, the Republic of South Africa in the field of information technologies for intelligent decision making support.

Articles included in the Proceedings are presented in the author's edition.

UDC 004.9
BBK 16

ISBN 978-5-7477-6355-5 (volume 1)
ISBN 978-5-7477-6354-8

© Ufa University, 2026

ПРЕДСЕДАТЕЛЬ ОРГАНИЗАЦИОННОГО КОМИТЕТА

Захаров В. П.	д-р хим. наук, профессор (УУНиТ, Уфа, Россия)

МЕЖДУНАРОДНЫЙ ОРГКОМИТЕТ

Кривошеева О. А.	д-р физ.-мат. наук, профессор (УУНиТ, Уфа, Россия)
Мустафина С. А.	д-р физ.-мат. наук, профессор (Госкомитет РБ по НиО, Россия)
Шахмаметова Г. Р.	д-р техн. наук, профессор (УУНиТ, Уфа, Россия)
Богданова Д. Р.	канд. техн. наук, доцент (УУНиТ, Уфа, Россия)
Сунит Кумар Гупта	директор AIT (ЧУ, Чандиграх, Индия)
Суреш Касван	д-р, профессор (ЧУ, Чандигарх, Индия)
Транос Зува	д-р, профессор (ВТУ, ЮАР)

ПРЕДСЕДАТЕЛЬ ПРОГРАММНОГО КОМИТЕТА

Шахмаметова Г. Р.	д-р техн. наук, профессор (УУНиТ, Уфа, Россия)

ПРОГРАММНЫЙ КОМИТЕТ

Антонов В. В.	д-р техн. наук, профессор (УУНиТ, Уфа, Россия)
Алиев Р. А.	член-корреспондент НАНА (НАНА, Баку, Азербайджан)
Бежаева О. Я.	канд. техн. наук, доцент (УУНиТ, Уфа, Россия)
Бельтюков А. П.	д-р физ.-мат. наук, профессор (УдГУ, Ижевск, Россия)
Богданова Д. Р.	канд. техн. наук, доцент (УУНиТ, Уфа, Россия)
Болодурина И. П.	д-р техн. наук, профессор (ОГУ, Оренбург, Россия)
Васильев В. И.	д-р техн. наук, профессор (УУНиТ, Уфа, Россия)
Вольфенгаген В. Э.	д-р техн. наук, профессор (ЮрИнфоР-МГУ, Россия)
Гардашова Л. А.	д-р техн. наук, профессор (АГУНП, Баку, Азербайджан)
Гвоздев В. Е.	д-р техн. наук, профессор (УУНиТ, Уфа, Россия)
Грибова В. В.	д-р техн. наук, профессор (ИАПУ ДВО РАН, Владивосток, Россия)
Джасприт Сингх Баттх	д-р, профессор (ЧУ, Чандигарх, Индия)
Елхова О. И.	д-р филос. наук, профессор (УУНиТ, Уфа, Россия)
Исмагилова А. С.	д-р физ.-мат. наук, профессор (УУНиТ, Уфа,Россия)
Картак В. М.	д-р физ.-мат. наук, профессор (УУНиТ, Уфа, Россия)
Кравец А. Г.	д-р техн. наук, профессор (ВолгГТУ, Волгоград, Россия)
Кривошеева О. А.	д-р физ.-мат. наук, профессор (УУНиТ, Уфа, Россия)
Масизана Одри	д-р, профессор (Университет Ботсваны, Ботсвана)
Мельников А. В.	д-р техн. наук, профессор (ЮНИИ ИТ, Ханты-Мансийск, Россия)
Миронов В. В.	д-р техн. наук, профессор (УУНиТ, Уфа, Россия)
Миронов К. В.	PhD, Dr. (УУНиТ, Уфа, Россия)
Муравьева Е. А.	д-р техн. наук, профессор (УГНТУ, Стерлитамак, Россия)
Мустафина С. А.	д-р физ.-мат. наук, профессор (УУНиТ, Уфа, Россия)
Петренко В. И.	канд. техн. наук, доцент (СКФУ, Ставрополь, Россия)
Петунин А. А.	д-р техн. наук, профессор (УрФУ, Екатеринбург, Россия)
Расулов С. Р.	д-р техн. наук, профессор (АГУНП, Баку, Азербайджан)
Родионова Л. Н.	д-р экон. наук, профессор (УУНиТ, Уфа, Россия)
Сирота А. А.	д-р техн. наук, профессор (ВГУ, Воронеж, Россия)
Сметанина О. Н.	д-р техн. наук, профессор (УУНиТ, Уфа, Россия)
Сунь Сяою	д-р, профессор (Шэньянский инженерный институт, Китай)
Суреш Касван	д-р, профессор (ЧУ, Чандигарх, Индия)
Тебуева Ф. Б.	д-р физ.-мат. наук, профессор (СКФУ, Ставрополь, Россия)
Транос Зува	д-р, профессор (Технологический университет Ваала, ЮАР)
Тюрин А. В.	д-р мед. наук, доцент (БГМУ, Уфа, Россия)
Христодуло О. И.	д-р техн. наук, профессор (УУНиТ, Уфа, Россия)
Шерыхалина Н. М.	д-р техн. наук, профессор (УУНиТ, Уфа, Россия)

GENERAL ORGANIZING CHAIR

Vadim Zakharov	UUST, Ufa, Russia

INTERNATIONAL ORGANIZING COMMITTEE

Olesya Krivosheeva	UUST, Ufa, Russia
Svetlana Mustafina	UUST, Ufa, Russia
Gouzel Shakhmametova	UUST, Ufa, Russia
Diana Bogdanova	UUST, Ufa, Russia
Sunit Kumar Gupta	CHU, Chandigarh, India
Suresh Kaswan	CHU, Chandigarh, India
Tranos Zuva	Vaal University of Technology, RSA

GENERAL PROGRAM CHAIR

Giuzel Shakhmametova	UUST, Ufa, Russia

PROGRAM COMMITTEE

Vyacheslav Antonov	UUST, Ufa, Russia
Rafiq Aliyev	Baku, Azerbaijan
Oksana Bezhaeva	UUST, Ufa, Russia
Anatoly Beltyukov	UdSU, Izhevsk, Russia
Diana Bogdanova	UUST, Ufa, Russia
Irina Bolodurina	OSU, Orenburg, Russia
Vladimir Vasiliev	UUST, Ufa, Russia
Vyacheslav Wolfengagen	JurInfoR-MSU, Moscow, Russia
Latafat Gardashova	ASUNP, Baku, Azerbaijan
Vladimir Gvozdev	UUST, Ufa, Russia
Valeriya Gribova	IACP FEB RAS, Vladivostok, Russia
Jaspreet Singh Batth	CHU, Chandigarh, India
Oksana Elkhova	UUST, Ufa, Russia
Albina Ismagilova	UUST, Ufa, Russia
Vadim Kartak	UUST, Ufa, Russia
Alla Kravets	VolgGTU, Volgograd, Russia
Olesya Krivosheeva	UUST, Ufa, Russia
Audrey Masizana	University of Botswana, Botswana
Andrey Melnikov	URIIT, Khanty-Mansiysk, Russia
Valery Mironov	UUST, Ufa, Russia
Konstantin Mironov	UUST, Ufa, Russia
Elena Muravyeva	USNTU, Sterlitamak, Russia
Svetlana Mustafina	UUST, Ufa, Russia
Vyacheslav Petrenko	NCFU, Stavropol, Russia
Alexander Petunin	UrFU, Yekaterinburg, Russ
Sakit Rasulov	ASUNP, Baku, Azerbaijan
Lyudmila Rodionova	UUST, Ufa, Russia
Alexander Sirota	VSU, Voronezh, Russia
Olga Smetanina	UUST, Ufa, Russia
Sun Xiaoyu	Shenyang Engineering Institute, China
Suresh Kaswan	CHU, Chandigarh, India
Fariza Tebueva	NCFU, Stavropol, Russia
Tranos Zuva	Vaal University of Technology, RSA
Anton Tyurin	BSMU, Ufa, Russia
Olga Christodulo	UUST, Ufa, Russia
Natalia Sheryakhalina	UUST, Ufa, Russia

Предисловие

Указом Президента Российской Федерации от 10 октября 2019 г. № 490 «О развитии искусственного интеллекта в Российской Федерации» утверждена Национальная стратегия развития искусственного интеллекта на период до 2030 года.

XI Международная научная конференция “Information Technologies for Intelligent Decision Making Support” ITIDS’2025 посвящена интенсивно развиваемым направлениям информационных технологий интеллектуальной поддержки принятия решений в области интеллектуальных моделей и методов, машинного обучения, нейросетевых технологий нечетких моделей и методов, системного анализа и математических методов принятия решений в сложных системах, а также вопросам их практического применения. В конференции приняли участие представители университетов, научно-исследовательских центров, научных школ Российской Федерации (Уфа, Москва, Санкт-Петербург, Химки, Екатеринбург, Таганрог, Ижевск, Владивосток, Челябинск, Оренбург, Ханты-Мансийск, Старый Оскол, Пермь, Самара, Иркутск, Стерлитамак, Кумертау, Пятигорск, Хабаровск, Ставрополь, Волгоград, Калининград, Казань, Иваново, Чита, Майкоп), Республики Индия, Республики Ирак, Южно-Африканской Республики, Республики Казахстан, Республики Таджикистан, Королевства Нидерландов, Республики Польша.

Тематика конференции направлена на обсуждение широкого круга вопросов, связанных с использованием последних достижений науки в области информационных технологий интеллектуальной поддержки принятия решений. Основные направления работы конференции: искусственный интеллект и машинное обучение; информационные технологии и информационные системы; интеллектуальные методы принятия решений в социально-экономических системах; методы защиты информации и обеспечения информационной безопасности; математическое моделирование в задачах принятия решений; приложения технологий искусственного интеллекта в системах различного назначения; философия искусственного интеллекта; интеллектуальные методы обработки биомедицинских данных.

От имени программного комитета выражаю благодарность всем участникам XI Международной конференции «Information Technologies for Intelligent Decision Making Support» ITIDS’2025.

Председатель программного комитета ITIDS’2025 Г. Р. Шахмаметова

Foreword

By the Decree of the President of the Russian Federation of October 10, 2019 No. 490 "On the Development of Artificial Intelligence in the Russian Federation," the National Strategy for the Development of Artificial Intelligence for the Period up to 2030 was approved.

XI International Scientific Conference "Information Technologies for Intelligent Decision Making Support" ITIDS'2025 is devoted to intensively developing areas of information technologies of intelligent decision support in the field of intelligent models and methods, machine learning, neural network technologies of fuzzy models and methods, system analysis and mathematical decision-making methods in complex systems, as well as issues of their practical application. The conference was attended by representatives of universities, research centers, scientific schools of the Russian Federation (Ufa, Moscow, St. Petersburg, Khimki, Yekaterinburg, Taganrog, Izhevsk, Vladivostok, Chelyabinsk, Orenburg, Khanty-Mansiysk, Stary Oskol, Perm, Samara, Irkutsk, Sterlitamak, Kumertau, Pyatigorsk, Khabarovsk, Stavropol, Volgograd, Kaliningrad, Kazan, Ivanovo, Chita, Maykop), the Republic of India, the Republic of Iraq, the Republic of South Africa, the Republic of Kazakhstan, the Republic of Tajikistan, the Kingdom of the Netherlands, the Republic of Poland.

The theme of the conference is aimed at discussing a wide range of issues related to the use of the latest achievements of science in the field of information technology and intellectual decision support. The main areas of work of the conference: artificial intelligence and machine learning; information technology and information systems; intelligent decision-making methods in socio-economic systems; information protection and information security methods; mathematical modeling in decision-making problems; applications of artificial intelligence technologies in systems for various purposes; artificial intelligence philosophy.

On behalf of the program committee, I would like to thank all the participants of the XI International Conference "Information Technologies for Intelligent Decision Making Support" ITIDS'2025.

General Program Chair of ITIDS'2025 Giuzel Shakhmametova

Оптимизация распределения оборотного капитала на электронных маркетплейсах с использованием двухуровневой модели на основе глубоких *Q*-сетей

Бастанов А. Э.
ООО «Инвестиционное агентство «Градиент»
г. Уфа, Россия
gradinv@gmail.com

Абраров А. М.
ООО «Цифра»
г. Уфа, Россия
azat.abrarov@soykasoft.ru

Аннотация[1]

Растущая конкуренция на электронных маркетплейсах требует применения методов искусственного интеллекта для оптимизации финансов в условиях неопределенности. В статье представлена двухуровневая модель оптимального распределения оборотного капитала, которая интегрирует стратегический анализ товарного портфеля на основе PTM–матрицы и оперативное управление закупками при помощи алгоритма глубокого *Q*-обучения. Эксперименты на синтетических данных для 50 товаров показали статистически значимое превосходство разработанного метода над традиционными алгоритмами: рентабельность капитала (ROCE) составила 23,7 % при сокращении потерь от дефицита (SLC) до 8,4 %, что на 5,3 и 3,7 п.п., соответственно, превышает показатели лучшего из базовых методов. Разработанный формализованный бизнес-процесс обеспечивает интеграцию стратегических ориентиров и операционных решений, повышая общую эффективность управления.

Ключевые слова: глубокое обучение с подкреплением, *Q*-обучение, Dueling DQN, Prioritized Experience Replay, оборотный капитал, управление ассортиментом, электронная коммерция, PTM–матрица.

1. Введение

Эволюция электронной коммерции и обострение конкуренции на маркетплейсах порождают комплекс вызовов для продавцов, ключевым из которых является оптимизация распределения ограниченного оборотного капитала. Высокая волатильность спроса, агрессивная ценовая конкуренция и выраженная сезонность формируют среду, в которой традиционные модели управления ресурсами демонстрируют низкую эффективность. Данное противоречие актуализирует необходимость разработки новых, адаптивных и интегрированных подходов к управлению.

Проблема управления товарным ассортиментом и оборотным капиталом традиционно рассматривается в контексте теории управления запасами. Фундаментальные работы [1–3] заложили основы статических и стохастических моделей, таких как модель экономичного размера заказа (EOQ) и ее модификации. Эти аналитически строгие модели обеспечивают надежные решения для условий стабильного спроса, однако оказываются недостаточно эффективными в быстро меняющейся конкурентной среде маркетплейсов, где факторы спроса носят нестационарный и многомерный характер.

В качестве альтернативного подхода рассматривается применение оптимизационных моделей, разработанных в теории исследования операций [4]. Данное направление включает линейное и нелинейное программирование для задач с соответствующими ограничениями, а также динамическое программирование для многоэтапных процессов принятия решений [5]. Однако для сложных, нелинейных и высокодинамичных сред электронной коммерции эти методы часто оказываются недостаточно гибкими и требуют существенных упрощающих допущений.

Более гибким подходом для учета неопределенности представляется стохастическое моделирование [6], предлагающее методический аппарат для работы со случайными параметрами, такими как спрос. Методы Монте-Карло и байесовские подходы позволяют моделировать вероятностные сценарии, однако их применение преимущественно ограничивается задачами анализа и симуляции, а не оперативного управления в реальном времени.

[1] Труды XI Международной научной конференции "Информационные технологии интеллектуальной поддержки принятия решений", 11-13 ноября, Уфа-Чандигарх-Фандербейлпарк, 2025

С развитием методов искусственного интеллекта появились новые возможности для адаптивного управления в условиях неопределенности. Особый интерес представляют методы глубокого обучения с подкреплением (Reinforcement Learning, RL). Алгоритмы глубокого *Q*-обучения и их модификации демонстрируют значительный потенциал для управления ресурсами в сложных динамических средах [7–10]. Особого внимания заслуживают такие архитектурные усовершенствования, как дуэльные сети (Dueling DQN) [11] и механизмы внимания [12], которые повышают способность алгоритмов к обобщению в условиях многомерного пространства состояний.

Для стратегического анализа ассортимента традиционно применяются портфельные модели, такие как матрица BCG [13]. Однако ее классические критерии (доля рынка и темпы его роста) считаются упрощенными для современных условий электронной коммерции, где критическую важность приобретают операционные метрики, такие как рентабельность и оборачиваемость.

Таким образом, в литературе выявлен пробел, связанный с отсутствием формализованных и верифицированных подходов, интегрирующих стратегический портфельный анализ товарного ассортимента и оперативное управление оборотным капиталом на основе RL в единый адаптивный контур.

2. Цель и задачи исследования

Целью настоящего исследования является разработка двухуровневой модели управления оборотным капиталом для продавцов на маркетплейсах, интегрирующей стратегический портфельный анализ и оперативное распределение ресурсов с помощью алгоритма глубокого *Q*-обучения.

Для достижения поставленной цели были решены следующие задачи:

1) разработать методику стратегической классификации и анализа товарного ассортимента, адаптированную для специфики маркетплейсов;

2) формализовать комплексный бизнес-процесс, обеспечивающий интеграцию стратегических и операционных решений;

3) формализовать задачу оперативного распределения оборотного капитала как марковского процесса принятия решений;

4) предложить архитектуру агента глубокого *Q*-обучения, способного учитывать комплексное влияние факторов рентабельности, оборачиваемости и сезонности;

5) провести вычислительный эксперимент на синтетических данных для сравнительной оценки эффективности предложенного подхода с традиционными методами.

3. Научная новизна и практическая значимость

Научная новизна работы заключается в следующем:

1) разработана и предложена оригинальная двухуровневая модель, интегрирующая стратегический портфельный анализ и оперативное управление оборотным капиталом на основе RL в единый контур;

2) разработано и предложено методическое обеспечение портфельного анализа,, адаптированное для маркетплейсов, расширяющее классические подходы;

3) произведены постановка и формализация задачи распределения оборотного капитала как марковского процесса принятия решений с учетом стратегического контекста товарных позиций.

Практическая значимость определяется:

1) разработкой готовой архитектуры RL-агента для поддержки оперативных решений в сфере управления оборотным капиталом;

2) доказательством эффективности предложенного подхода по сравнению с традиционными методами, выраженным в ключевых финансовых метриках;

3) формализацией бизнес-процесса, который может быть положен в основу систем поддержки принятия решений для продавцов.

4. Двухуровневая модель управления оборотным капиталом продавца на маркетплейсах

Высокая волатильность спроса и динамичная конкурентная среда требуют разработки современных систем управления, обеспечивающих как оперативную реакцию на изменения рынка, так и достижение устойчивых конкурентных преимуществ. В электронной коммерции деятельность продавцов может быть описана в рамках базовых стратегий, адаптированных по М. Трейси и Ф. Вирсема [14]:

1) операционное превосходство – достижение лидерства за счет минимизации издержек и эффективности операционных процессов;

2) лидерство в продукте – создание уникального товарного предложения, защищенного от прямой ценовой конкуренции за счет инноваций, качества, дизайна и премиального позиционирования;

3) глубокое понимание ниши (адаптация стратегии «близости к клиенту») – фокусировка на узком сегменте с предложением кастомизированных решений.

Реализация любой из этих стратегий требует системного подхода к управлению ассортиментным портфелем и оптимального распределения ресурсов. Для решения этой задачи в настоящем исследовании

предлагается двухуровневая модель, обеспечивающая интеграцию стратегического выбора приоритетов и операционного контура принятия решений.

4.1.Стратегический контур: классификация ассортимента на основе PTM–матрицы

Для операционализации конкурентной стратегии продавцов на электронных маркетплейсах целесообразно применение инструментов портфельного анализа. Однако классические модели, в частности матрица BCG [13], демонстрируют ограниченную применимость в условиях цифровой экономики, поскольку опираются на параметры, сложные для измерения (относительная доля рынка) и слабо коррелирующие с операционной эффективностью отдельных товарных позиций.

В качестве решения предлагается модифицированная матрица «Прибыльности и Оборачиваемости» (Profit-Turnover Matrix, PTM), основанная на следующих критериях:

– ось ординат отражает маржинальную рентабельность с учетом полных переменных издержек;

– ось абсцисс представляет коэффициент оборачиваемости запасов;

– площадь круга соответствует объему продаж в денежном выражении.

Данная модификация позволяет перейти от анализа рыночных позиций к оценке финансово-операционной эффективности товарных позиций, используя доступные и релевантные для принятия управленческих решений метрики.

Пересечение двух дихотомических переменных формирует четыре квадранта (рис. 1), каждому из которых соответствует определенная товарная категория и типовая портфельная стратегия:

– «Хиты» (высокая рентабельность, высокая оборачиваемость): ключевые драйверы прибыли, требующие приоритетного управления запасами, оптимизации рекламных кампаний и активной работы с отзывами;

– «Премиум» (высокая рентабельность, низкая оборачиваемость): товары с уникальным ценностным предложением для узкой аудитории, нуждающиеся в премиальном позиционировании и точечном таргетинге;

– «Генераторы» (низкая рентабельность, высокая оборачиваемость): обеспечивают объем продаж и приток клиентов, требуют жесткого контроля издержек и использования в кросс-селлинговых схемах;

– «Балласты» (низкая рентабельность, низкая оборачиваемость): ресурсоемкие неперспективные товары, подлежащие выводу из ассортимента.

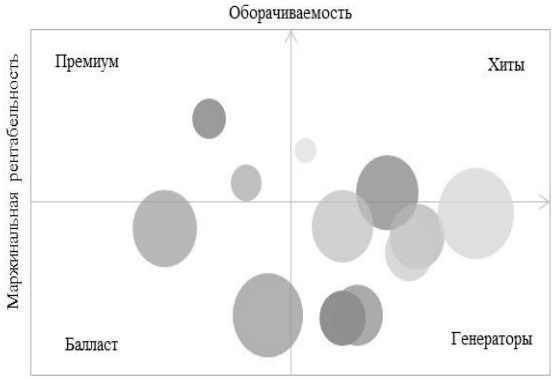


Рис. 1. Пример заполнения матрицы PTM

Предложенная PTM–матрица формирует системную основу для классификации товарного портфеля и разработки дифференцированных стратегий управления различными товарными категориями. Важнейшим преимуществом модели является ее способность служить инструментом операционализации стратегии, обеспечивая переход от концептуального уровня к практическому управлению.

Как демонстрирует табл. 1, матрица устанавливает четкое соответствие между квадрантами и базовыми конкурентными стратегиями, определяя конкретные стратегические задачи управления товарным портфелем для каждого направления.

Таким образом, PTM–матрица выполняет двойную функцию: служит как диагностическим инструментом для оценки текущего состояния портфеля, так и инструментом стратегического управления, позволяющим целенаправленно трансформировать структуру ассортимента в соответствии с выбранной конкурентной стратегией.

Таблица 1. – Связь категорий PTM–матрицы с базовыми конкурентными стратегиями.

Конкурентная стратегия	**Товарный портфель продавца**	**Стратегические задачи управления товарным портфелем**
Операционное превосходство	«Хиты», «Генераторы»	1) оптимизация ключевых операционных процессов: минимизация сроков исполнения заказов и предотвращение дисбалансов запасов через автоматизацию управления поставками и логистикой; 2) системное снижение себестоимости: непрерывный контроль и минимизация всех переменных издержек, включая закупочные цены, логистику, хранение и рекламные расходы; 3) трансформация структуры портфеля в сторону увеличения доли высокомаржинальных товаров с высокой оборачиваемостью путем планомерного перевод «Генераторов» в категорию «Хитов» путем оптимизации их маржинальности.
Лидерство в продукте	«Премиум»	1) создание и поддержание уникального ценностного предложения за счет систематического выявления неудовлетворенных потребностей рынка и разработки продуктов с превосходящими характеристиками, оправдывающими премиальную цену; 2) формирование сильного премиального бренда. Предполагает развитие коммуникации, акцентирующей инновационные преимущества продукции, и обеспечение правовой защиты ключевых разработок; 3) селективное расширение портфеля для на увеличения оборачиваемости премиальных продуктов без снижения цен, например, за счет выхода на смежные сегменты, с целью их перевода в категорию высокомаржинальных «Хитов».
Глубокое понимание ниши	«Премиум, «Хиты»	1) глубокая сегментация рынка и кастомизация предложения. Формирование портфеля, в значительной степени состоящего из продуктов, разработанных для специфических потребностей сегмента, включая возможности индивидуальной настройки; 2) формирование лояльного сообщества через активное привлечение клиентов к разработке продуктов (краудсорсинг) и внедрение персонализированного сервиса; 3) управление портфелем как интегрированным решением. Активное использование кросс-селлинга и комплектов для повышения потребительской ценности, наряду с регулярным аудитом на соответствие динамичным запросам ниши.

4.2. Оперативный контур: зоны управления и процесс принятия решений

На оперативном уровне реализуется тактическое управление товарным портфелем, структурно разделенное на две зоны:

1) зону текущей деятельности: связана с продажей товаров категорий «Хиты», «Генераторы» и «Премиум», а также реализацией операционных планов продавца. Задачи направлены на максимизацию текущего денежного потока и прибыли через оптимизацию запасов, цен и маркетинга;

2) зону развития: связана с инкубацией перспективных товаров в категории «Премиум» и реабилитацией товаров в категории «Балласты», а также реализацией новых проектов и инициатив. Деятельность носит проектный характер и связана с тестированием гипотез для обеспечения будущего роста.

Соотношение бюджетов на текущую деятельность и развитие зависит от жизненного цикла бизнеса и составляет от 40/60 на стадии стартапа, до 85/15 на стадии зрелости.

Бюджет закупок товаров в зоне развития распределяется равномерно, поскольку целью является тестирование гипотез относительно рыночного спроса, а также показателей рентабельности и оборачиваемости товаров.

Предложенная схема комплексного бизнес-процесса стратегически ориентированного управления товарным ассортиментом и финансами продавца (рис. 2)

обеспечивает стратегическую согласованность решений, адаптивность на основе обратной связи и сбалансированность краткосрочных и долгосрочных целей. Механизм позволяет гибко перераспределять ресурсы в соответствии с рыночной динамикой и стратегическими приоритетами компании.

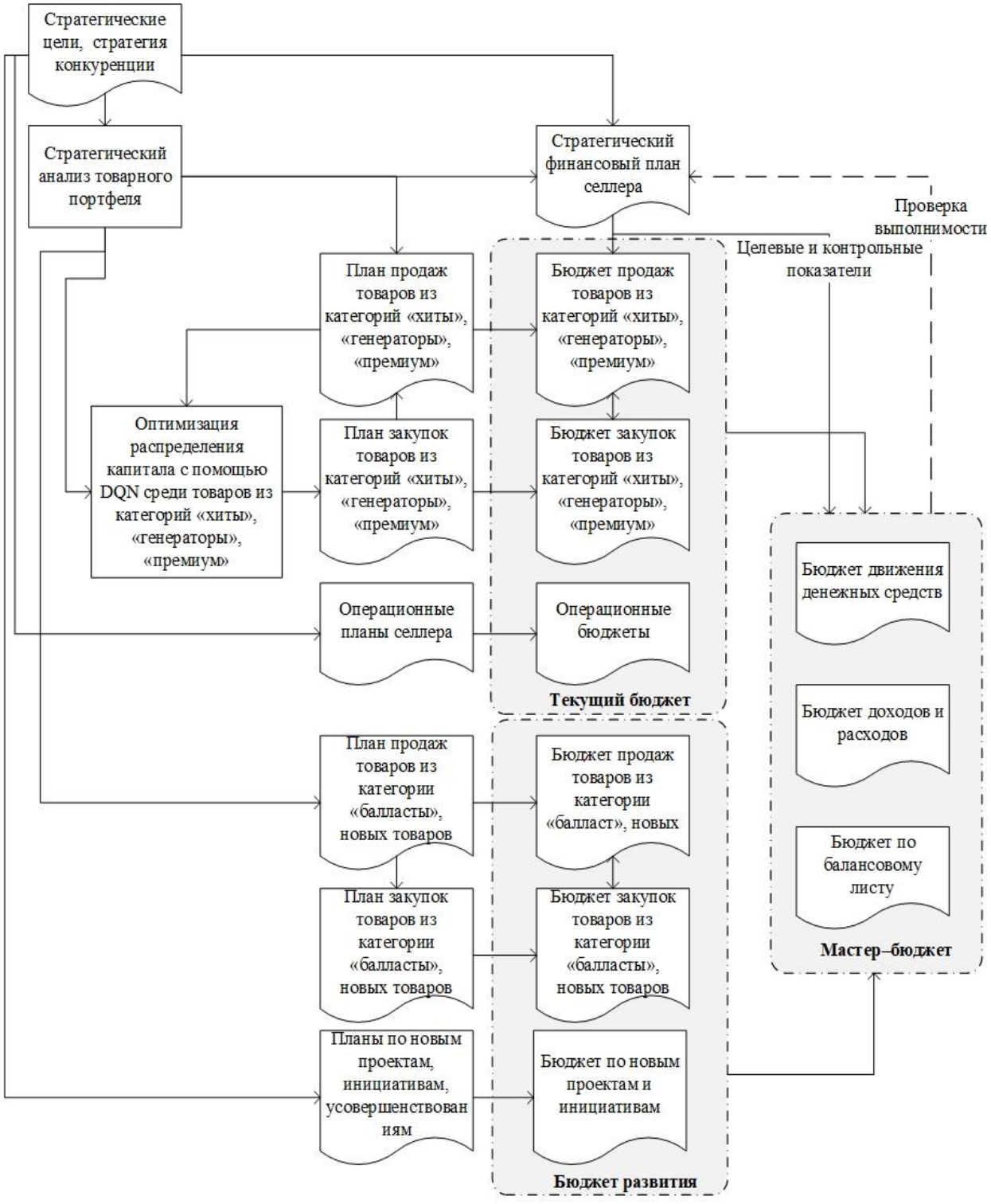


Рис. 2. Схема комплексного бизнес-процесса стратегически ориентированного управления товарным ассортиментом и финансами продавца

5. Постановка задачи оптимизации распределения капитала и метод на основе глубокого обучения с подкреплением

В зоне текущей деятельности актуальной является задача оптимального распределения финансовых ресурсов при формировании бюджета закупки.

Особенностью данной предметной области является высокая волатильность спроса, выраженная сезонность, интенсивная конкурентная среда и необходимость учета множества факторов при принятии решений о распределении ограниченных финансовых ресурсов между товарными позициями.

Для решения указанной задачи предлагается использовать алгоритм глубокого *Q*-обучения, который учитывает особенности предметной области.

5.1. Формализация задачи как марковского процесса принятия решений

Работа продавца на электронном маркетплейсе рассматривается как дискретный многошаговый процесс управления в условиях неопределенности. Целью является максимизация совокупной прибыли на конечном горизонте планирования T путем оптимального распределения ограниченного объема оборотного капитала C_{total} между N товарными позициями:

$$\max_{\pi} E_{\pi}(\sum_{t=1}^{T} \gamma^{t-1} \cdot R_t) \quad (1)$$

где π – стратегия распределения капитала, $\gamma \epsilon [0,1]$ – коэффициент дисконтирования, R_t – прибыль, полученная в периоде *t*.

Для формализации модели приняты следующие допущения:

1) марковское свойство. Предполагается, что текущее состояние системы полностью характеризует всю релевантную историю для принятия решения в момент *t*. Будущее состояние и вознаграждение зависят только от текущего состояния и выбранного действия, а не от всей предыстории;

2) бинарные управляющие воздействия. Агент принимает решение о включении (a_i=1) или исключении (a_i=0) товара *i* в список приоритетных для финансирования в текущем периоде. Объем финансирования определяется автоматически на основе весовых коэффициентов;

3) факторная модель спроса. Стохастичность спроса обусловлена внешними, наблюдаемыми факторами – ожидаемым спросом $\overline{D_i}(t)$ и сезонным коэффициентом $W_i(t)$, которые являются экзогенными параметрами модели;

4) мгновенное пополнение запасов. Заказанный товар поступает на склад мгновенно, в рамках того же временного периода *t*. Временные лаги поставок не учитываются;

5) динамический бюджет. Средства, вложенные в закупку, возвращаются в бюджет по мере реализации товара. Бюджет следующего периода C_{t+1} формируется из текущего бюджета за вычетом затрат на заказы плюс выручка от продаж. В модели не учитываются внешние вливания капитала, изъятия прибыли, а также дополнительные операционные или финансовые расходы;

6) стабильность ценовых параметров. Закупочные цены c_i, продажные цены p_i, удельные затраты на хранение h_i, а также коэффициент штрафа за упущенную выручку из-за дефицита товара θ постоянны и известны на всем горизонте планирования.

Задача моделируется как марковский процесс принятия решений (MDP), заданный кортежем $\langle S, \mathcal{A}, \mathcal{P}, R, \gamma \rangle$:

1) пространство состояний S. Состояние системы S_t в момент *t* описывается вектором:

$$S_t = \left(t, \mathrm{C}_t, \vec{I}_t, \vec{\overline{D}}_t, \vec{W}_t\right) \quad (2)$$

где:

– $t \in \{1,2, \dots, T\}$ – текущий временной период;

– $C_t \in \mathbb{R}^+$ – доступный объем оборотного капитала в начале периода *t*;

– $\vec{I_t} = \left(I_1(t), \dots, I_N(t)\right), I_i(t) \in \mathbb{R}^+$ – уровень запасов товара *i* на складе в период *t*;

– $\vec{\overline{D}}_t = \left(\overline{D}_1(t), \dots, \overline{D}_N(t)\right), \overline{D}_i(t) \in \mathbb{R}^+$ – спрос (ожидаемый) на товар *i* в периоде *t*;

– $\overrightarrow{W_t} = \left(W_1(t), \dots, W_N(t)\right), W_i(t) \in \mathbb{R}^+$ – сезонный коэффициент спроса на товар *i* в период *t*;

2) пространство действий агента $\mathcal{A}$ определяется бинарным вектором:

$$a_t = [a_1, \dots, a_N], a_i \epsilon \{0,1\} \quad (3)$$

где $a_i = 1$ означает принятие решения о закупке *i*-го товара в периоде *t*.

3) функция перехода $\mathcal{P}(S_{t+1}|s_t, a_t)$:

– распределение оборотных средств на каждом шаге *t* реализуется через расчет нормализованных весовых коэффициентов $k_i(t)$ на основе функции *SoftMax*, взвешивающей модифицированный индекс доходности, скорректированный на текущую рыночную конъюнктуру:

$$k_i(t) = \frac{e^{z_i(t)}}{\sum_{j=1}^{N} a_j \cdot e^{z_j(t)}} \quad (4)$$

где $z_i(t)$ – логит для товара *i*:

$$z_i(t) = M_i \cdot b_i \cdot W_i(t) \cdot \overline{D}_i(t) \quad (5)$$

где:

– $M_i = \frac{(p_i - c_i)}{p_i}$ – маржинальная рентабельность *i-го* товара (здесь p_i, c_i – продажная и закупочная цены единицы товара i, соответственно);

– b_i –оборачиваемость *i-го* товара (дней).

Объем закупок для каждого товара i вычисляется по формуле:

$$invest_i(t) = a_i \cdot k_i(t) \cdot C_t \quad (6)$$

– объем поставки товара i в период t:

$$order_i(t) = \frac{invest_i(t)}{c_i} \quad (7)$$

– реализованный спрос (генерируется средой) на основе логнормального распределения [15]:

$$D_i^{actual}(t) = \overline{D}_i(t) \cdot W_i(t) \cdot \varepsilon_i, \varepsilon_i \sim LogNormal(0, \sigma_i) \quad (8)$$

– фактические продажи:

$$sales_i(t) = \min\left(I_i(t) + order_i(t),\ D_i^{actual}(t)\right) \quad (9)$$

– обновление уровня запасов:

$$I_i(t+1) = max(0, I_i(t) + order_i(t) - sales_i(t)) \quad (10)$$

– обновление бюджета:

$$C(t+1) = C_t - \sum_{i=1}^{N} invest_i(t) + \sum_{i=1}^{N} p_i \cdot sales_i(t) \quad (11)$$

– генерация новых параметров спроса: $\overline{D_i}(t)$, $W_i(t)$ стохастически генерируются средой, завершая формирование нового состояния S_{t+1};

3) функция вознаграждения $R(s_t, a_t)$ оценивает финансовый результат от действий агента в периоде t с учетом штрафов за неоптимальность:

$$R(s_t, a_t) = \sum_{i=1}^{N} \Big[(p_i - c_i) \cdot sales_i(t) - h_i \cdot I_i(t+1) - \theta \cdot \max(0, D_i^{actual}(t) - sales_i(t))\Big] \quad (12)$$

где:

– h_i – стоимость хранения единицы товара i;

– θ – коэффициент штрафа за упущенную выручку из-за дефицита товара.

Представленная постановка задачи определяет основу для применения алгоритмов глубоких Q-сетей, для выявления сложных зависимостей между стратегическими приоритетами, рыночной конъюнктурой и операционными решениями.

5.2. Архитектура модифицированной дуэльной DQN-сети

Для аппроксимации Q-функции в задаче оптимального распределения ресурсов была предложена дуэлинговая архитектура сети (Dueling DQN) [11] с механизмом самовнимания (self-attention) [12] и приоритизированной памятью прецедентов (Prioritized Experience Replay, PER) [16]. Выбор данной архитектуры был обусловлен спецификой задачи оптимизации распределения оборотного капитала в многомерном и стохастическом окружении.

Важным этапом предобработки данных является нормализация входного вектора состояния, что обеспечивает стабильность и ускоряет сходимость процесса обучения [17].

Входной вектор состояния S_t включает разнородные компоненты, для которых была применена стандартная нормализация (Z-score) для непрерывных признаков (за исключением сезонных коэффициентов $W_i(t)$, которые уже находятся в сопоставимом диапазоне).

Архитектура сети состоит из трех основных компонентов:

1) энкодер признаков, включающий в себя три полносвязных слоя:

$$h_1 = ReLu(W_1 \cdot s_t + b_1), \dim(h_1) = 256 \quad (13)$$

$$h_2 = ReLu(W_2 \cdot h_1 + b_2), \dim(h_2) = 128 \quad (14)$$

$$h_3 = ReLu(W_3 \cdot h_2 + b_3), \dim(h_3) = 64 \quad (15)$$

2) механизм самовнимания, который преобразует выход энкодера в Query, Key и Value:

$$Q = h_3 \cdot W^Q, K = h_3 \cdot W^K, V = h_3 \cdot W^V \quad (16)$$

Затем вычисляются веса внимания:

$$Attention(Q, K, V) = softmax(\frac{Q \cdot K^T}{\sqrt{d_k}}) \cdot V \quad (17)$$

Выход механизма внимания – вектор h_{att}.

3) дуэлинговые «головы» (Dueling Heads):

$$V(s) = W_V \cdot h_{att} + b_V \quad (18)$$

$$A(s, a) = W_A \cdot h_{att} + b_A \quad (19)$$

$$Q(s, a) = V(s) + A(s, a) - \frac{1}{|\mathcal{A}|}\sum_{a'} A(s, a') \quad (20)$$

где:

– $V(s)$ – функция оценки полезности состояния (вне зависимости от действия);

– $A(s, a)$ – функция оценки преимущества действия;

– $Q(s, a)$ – Q-функция.

Для повышения эффективности обучения используется приоритизированный буфер повторений [12]. Каждый переход $i = (s, a, r, s')$ хранится с приоритетом p_i. Вероятность выборки перехода j определяется как:

$$P(j) = \frac{p_j^\alpha}{\sum_i p_j^\alpha} \quad (21)$$

где приоритет p_j вычисляется на основе TD–ошибки:

$$p_j = |\delta_j| + \varepsilon \quad (22)$$

Для коррекции смещения используется вес важности:

$$w_j = \left(\frac{1}{N} \cdot \frac{1}{P_j}\right)^{\beta} \quad (23)$$

Функция потерь минимизируется с помощью алгоритма Adam:

$$\mathcal{L}(\theta) = \frac{1}{K}\sum_{j=1}^{K} w_j \cdot \left[y_j - Q(s_j, a_j; \theta)\right]^2 \quad (24)$$

где целевое значение:

$$y_j = r_j + \gamma \cdot \max_{a'} Q(s_{j+1}, a'; \theta^-) \quad (25)$$

Веса целевой сети обновляются на каждом шаге обучения:

$$\theta^- \leftarrow \tau \cdot \theta + (1 - \tau) \cdot \theta^- \quad (26)$$

где θ– параметры основной Q–сети, θ^- – параметры целевой сети.

Гиперпараметры модели приведены в табл. 2.

Таблица 2. – Гиперпараметры модели

Параметр	**Значение**	**Описание**
Размер буфера	100 000	Емкость памяти прецедентов
Размер батча, K	64	Количество примеров на шаг обучения
α	0.001	Скорость обучения
γ	0.99	Коэффициент дисконтирования
ϵ_{start}	1.0	Начальная вероятность исследования
ϵ_{end}	0.1	Конечная вероятность исследования
α	0.6	Степень приоритизации
β	0.4 → 1.0	Параметр коррекции смещения
τ	0.01	Частота обновления целевой сети

5.3. Методика генерации синтетических данных

Для обучения и валидации модели были сгенерированы синтетические данные для 50 товаров на горизонте 365 дней. Параметры генерировались следующим образом:

– функция спроса:

$$D_i(t) = D_i^{base} \cdot w_i(t) + \epsilon_t \quad (27)$$

где D_i^{base}– базовый спрос, $w_i(t)$ – мультипликативная сезонность, ϵ_t – случайная ошибка;

– базовый спрос смоделирован нормальным распределением:

$$D_i^{base} \sim \mathcal{N}(\mu_i, \sigma_i^2) \quad (28)$$

где μ_i – среднее значение спроса, σ_i^2– вариация;

– сезонность смоделирована через синусоидальную функцию со смещением:

$$w_i(t) = 1 + k_i \cdot \sin\left(\frac{2\pi t}{T_{season}}\right) \quad (29)$$

где k_i – амплитуда, T_{season} – период сезонности.

Выбранные функции представляют собой компромисс между реалистичностью, простотой и контролируемостью. Они отражают ключевые особенности реальной среды:

– стохастичность спроса через нормальное распределение;

– предсказуемые периодические колебания через синусоидальную сезонность.

Предложенный подход позволяет генерировать репрезентативные синтетические данные для валидации эффективности алгоритма, обеспечивая контроль над параметрами для анализа чувствительности и гарантируя воспроизводимость результатов.

5.4. Анализ эффективности предложенного алгоритма

Для верификации эффективности предложенного DQN-алгоритма была разработана система сравнительной оценки с тремя базовыми методами управления оборотным капиталом:

1) равномерное распределение (M1) – капитал распределяется поровну между всеми товарными позициями. Метод служит базовым уровнем для оценки минимальной эффективности;

2) пропорциональное распределение по рентабельности (M2) – доля капитала соответствует маржинальной рентабельности товара. Представляет собой интуитивный бизнес-подход;

3) эвристический алгоритм на основе оборачиваемости (M3) – приоритет товарам с высокой оборачиваемостью. Представляет продвинутую эвристику, балансирующую между ликвидностью и доходностью.

Для количественной оценки эффективности использовалась система показателей:

1) рентабельность оборотного капитала (ROCE, %) – отношение прибыли к объему инвестированного капитала;

2) коэффициент использования капитала (UCR, %) – доля инвестированного в запасы капитала;

3) коэффициент потерь от дефицита (SLC, %) – доля упущенной выручки из-за отсутствия товара.

Результаты 100 независимых экспериментов (табл. 3) демонстрируют статистически значимое превосходство (p-value <0.05) DQN-алгоритма по всем метрикам.

Полученные результаты свидетельствуют о способности DQN–алгоритма выявлять сложные нелинейные зависимости между маржинальностью, оборачиваемостью и сезонностью спроса. Превышение показателя ROCE на 5.3 п.п. относительно лучшего базового метода (M3) подтверждает эффективность предложенного подхода для задач оптимизации оборотного капитала.

Таблица 3. – Сравнение эффективности алгоритмов

Метод	**ROCE, %**	**UCR, %**	**SLC, %**
M1	12.3 ± 1.5	89.5 ± 3.2	18.7 ± 2.8
M2	15.8 ± 1.1	91.2 ± 2.1	15.3 ± 2.1
M3	18.4 ± 0.9	93.6 ± 1.8	12.1 ± 1.7
Метод DQN	23.7 ± 0.7	96.8 ± 1.1	8.4 ± 1.2

5.5. Ограничения исследования и перспективы дальнейших работ

Несмотря на полученные положительные результаты, настоящее исследование имеет определенные ограничения, которые открывают направления для будущих изысканий:

1) проблема высокой размерности и дискретности решений. При расширении ассортимента (N> 100 товарных позиций) пространство бинарных действий растет экспоненциально, что создает вычислительные трудности и замедляет сходимость алгоритма. Кроме того, предложенная модель предполагает возможность заказа дробных объемов товара, тогда как на практике заказы размещаются целыми партиями. Это упрощение может приводить к получению нереализуемых на практике решений;

Для преодоления этих ограничений перспективным является применение алгоритмов глубокого RL с непрерывным пространством действий (например, DDPG, TD3, SAC) с последующим округлением предсказанного объема заказа до ближайшего целого числа. Альтернативным подходом является разработка иерархической RL-архитектуры, где на верхнем уровне определяется стратегический приоритет товара, а на нижнем уровне решается задача целочисленной оптимизации для расчета точного размера заказа;

2) требования к качеству данных. Эффективность модели существенно зависит от репрезентативности исторических данных. В данной работе не учитывались такие факторы, как промо-акции, динамика рейтингов, маркетинговые активности конкурентов. В этой связи перспективной видится разработка гибридных моделей, интегрирующих RL с передовыми методами прогнозирования временных рядов, а также расширение пространства состояний за счет включения векторных представлений (эмбеддингов) для текстовых описаний товаров и данных веб-аналитики;

3) верификация на реальных данных. Валидация подхода проводилась на синтетических данных, которые не могут в полной мере отразить всю сложность реальной рыночной среды.

Мы планируем провести A/B-тестирование на исторических данных компаний-партнеров и интеграция алгоритма в тестовый контур корпоративной информационной системы;

4) упрощения модели склада. Модель предполагает мгновенное пополнение запасов и отсутствие затрат на оформление заказа. Запланировано введение в модель временных лагов поставки и фиксированных затрат на заказ, что потребует модификации пространства состояний и функции вознаграждения;

5) интерпретируемость решений. Модели на основе глубоких нейронных сетей часто остаются «черными ящиками», что затрудняет их принятие менеджерами.

Для решения этой проблемы рассматривается применение методов explainable AI (XAI), таких как SHAP или LIME, для анализа важности признаков и обеспечения прозрачности принимаемых решений.

Представленные ограничения не снижают ценность полученных результатов, но определяют стратегию для дальнейшего совершенствования методологии и расширения области её применения.

6. Заключение

В работе предложен и верифицирован на синтетических данных инновационный подход к оптимизации распределения оборотного капитала для продавцов на электронных маркетплейсах. Ключевой особенностью разработки является двухуровневая архитектура, которая последовательно интегрирует стратегическую классификацию товарного портфеля с использованием PTM–матрицы и оперативное принятие решений на основе усовершенствованного алгоритма глубокого *Q*-обучения (дуэльная DDQN с механизмами самовнимания и приоритизированной памяти прецедентов).

Экспериментальные результаты подтвердили эффективность предложенного методического обеспечения, продемонстрировав его статистически значимое превосходство над традиционными эвристиками по ключевым метрикам: рентабельности оборотного капитала (ROCE) и коэффициенту потерь от дефицита (SLC). Это свидетельствует о способности алгоритма выявлять сложные нелинейные зависимости между маржинальностью, оборачиваемостью и сезонностью спроса.

Перспективы дальнейших исследований связаны с преодолением выявленных ограничений, а именно: применением алгоритмов с непрерывным пространством действий (например, DDPG, SAC) для работы с крупными ассортиментами, интеграцией дополнительных факторов (промо-акции, рейтинги, данные веб-аналитики) в пространство состояний, верификацией на реальных данных и внедрением методов explainable AI (XAI) для повышения интерпретируемости и доверия к рекомендациям модели со стороны менеджеров.

Представленное решение вносит вклад в развитие методологии интеллектуальной поддержки принятия решений в сфере электронной коммерции и имеет значительный потенциал для практического внедрения в корпоративные информационные системы.

Список используемых источников

1. Foundations of inventory management / P.H. Zipkin. – Boston; Toronto; London: McGraw-Hill, 2000.
2. Porteus E.L. Foundations of Stochastic Inventory Theory. – Stanford University Press, 2002.
3. Silver, E.A., Pyke, D.F. and Peterson, R. (1998) Inventory Management and Production Planning and Scheduling. 3rd Edition, John Wiley, New York.
4. Таха, Хемди А. Введение в исследование операций, 7–е издание.: Пер. с англ. – М.: Издательский дом «Вильямс», 2005. –912 с.
5. Р. Беллман, С. Дрейфус. Прикладные задачи динамического программирования. – М.: Издательство «Наука», 1965.
6. Shapiro A., Dentcheva D., Ruszczynski A. Lectures on Stochastic Programming: modeling and theory. SIAM, 2014.
7. Обучение с подкреплением / Р.С. Саттон, Э.Г. Барто; пер. с англ. – М.: БИНОМ. Лаборатория знаний, 2017. – 399 с.
8. Deng Y., Bao F., Kong Y., Ren Z., Dai Q. Deep Direct Reinforcement Learning for Financial Signal Representation and Trading // IEEE Transactions on Neural Networks and Learning Systems. 2017. Vol. 28. № 3. P. 653–664.
9. Levine S., Finn C., Darrell T., Abbeel P. End-to-End Training of Deep Visuomotor Policies // Journal of Machine Learning Research. 2016. Vol. 17. № 1. P. 1334–1373.
10. Mao H., Alizadeh M., Menache I., Kandula S. Resource Management with Deep Reinforcement Learning // Proceedings of the 15th ACM Workshop on Hot Topics in Networks. 2016. P. 50–56.
11. Wang Z. et al. Dueling Network Architectures for Deep Reinforcement Learning // International Conference on Machine Learning. 2016. P. 1995-2003.
12. Vaswani A. et al. Attention is All You Need // Advances in Neural Information Processing Systems. 2017. P. 5998-6008.
13. Стратегии, которые работают: Подход BCG: сб. ст.: пер. с англ. / Сост. Карл Штерн и Джордж Сток-мл.; – 2-е изд.– М.: Манн, Иванов и Фербер, 2007. – 496 с.
14. Трейси М., Вирсема Ф. Маркетинг ведущих компаний: выбери потребителя, определи фокус, доминируй на рынке.: Пер. с англ.– М.: ООО «И.Д. Вильямс», 2007. – 304 с.
15. Cobb, Barry & Rumí, Rafael & Salmerón, Antonio. (2013). Inventory management with Log-normal demand per unit time. Computers & Operations Research. 40.
16. Schaul T. et al. Prioritized Experience Replay // International Conference on Learning Representations. 2016.
17. Ioffe S., Szegedy C. Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift // International Conference on Machine Learning. 2015. P. 448-456.

Цифровизация управленческого учета: роль и возможности искусственного интеллекта

Кобзев А. М.
ООО «Инфостарт»
Уфа, Россия
e-mail: andrewkobzev97@gmail.com

Аннотация[1]

В статье рассматриваются особенности цифровизации управленческого учета с применением технологий искусственного интеллекта. Анализируются возможности автоматизации рутинных процессов, преимущества и ограничения внедрения искусственного интеллекта в финансово-экономические службы предприятия. Подчеркивается, что искусственный интеллект пока не способен полностью заменить человека, особенно в стратегическом управлении и интерпретации данных. Рекомендуется осторожный и поэтапный подход к внедрению, при котором искусственный интеллект используется как инструмент поддержки, а ключевая роль в принятии управленческих решений сохраняется за специалистами. В статье также отмечаются риски и ошибки применения искусственного интеллекта и предлагаются направления его оптимального использования.

Ключевые слова: цифровизация; управленческий учет; искусственный интеллект; автоматизация; прогнозирование; финансово-экономические службы; риски внедрения.

1. Искусственный интеллект в системе управленческого учета компании

Управленческий учет занимает ключевое место в системе менеджмента современных предприятий. Он обеспечивает руководство компаний достоверной, актуальной и аналитически обработанной информацией, которая используется для стратегического и тактического планирования, анализа затрат, оценки эффективности проектов, прогнозирования и принятия решений. С развитием технологий, особенно в эпоху цифровизации и массового внедрения искусственного интеллекта (далее – ИИ), возникает вопрос о том, насколько управленческий учет может быть автоматизирован/роботизирован и какие функции человека в этой системе сохраняются. ИИ уже активно внедряется в корпоративные процессы, однако в международном сообществе сохраняется скепсис относительно его универсальности и надежности. Ключевой тезис данной работы заключается в том, что искусственный интеллект на текущий момент не способен заменить человека полностью, особенно в процессах принятия решений (в том числе экономических), и его применение требует осторожности.

В последние годы наблюдается активный рост внедрения технологий ИИ в финансово-экономическую сферу. Компании стремятся использовать машинное обучение и интеллектуальные алгоритмы для обработки больших массивов информации, выявления закономерностей и прогнозирования. Согласно отчету «KPMG Global AI in Finance», 71% компаний по всему миру уже применяют ИИ в финансовых процессах, а 41 % используют его в значительной степени [1]. По данным издания «Mordor Intelligence», объем мирового рынка ИИ в бухгалтерии составит 7,52 млрд долл к 2025 г. и вырастет до 50,29 млрд к 2030 г. [2]. Схожие выводы приводят издания «Market Research Future» и «Accounting Today», которые фиксируют значительное увеличение использования автоматизированных решений для отчетности, анализа и аудита [3, 4]. Однако статистика показывает и ограниченность: малые и средние предприятия внедряют ИИ гораздо медленнее из-за дефицита ресурсов и компетенций [5].

Следует подчеркнуть, что внедрение искусственного интеллекта в управленческий учет не ограничивается только автоматизацией базовых процессов. ИИ начинает использоваться для формирования прогнозных моделей, оценки сценариев развития бизнеса и даже поддержки инвестиционных решений. Однако его роль в этих областях ограничена необходимостью экспертной интерпретации: без участия человека результаты моделей могут быть восприняты некорректно. Это связано как с ограничениями самих алгоритмов (зависимость от качества обучающих данных, невозможность учета нефинансовых факторов), так и с высокой стоимостью последствий ошибок в

[1] Труды XI Международной научной конференции "Информационные технологии интеллектуальной поддержки принятия решений", 11-13 ноября, Уфа-Чандигарх-Фандербейлпарк, 2025

управленческих решениях. Таким образом, ИИ следует рассматривать не как замену специалиста, а как инструмент, усиливающий его аналитические способности и расширяющий горизонты анализа.

По результатам анализа вышеприведенных источников можно сделать твердый вывод о том, что внедрение ИИ в области финансово-экономического учета на предприятиях – это уже совсем не будущее, а настоящее время. Наблюдается устойчивый тренд к росту (рис. 1).

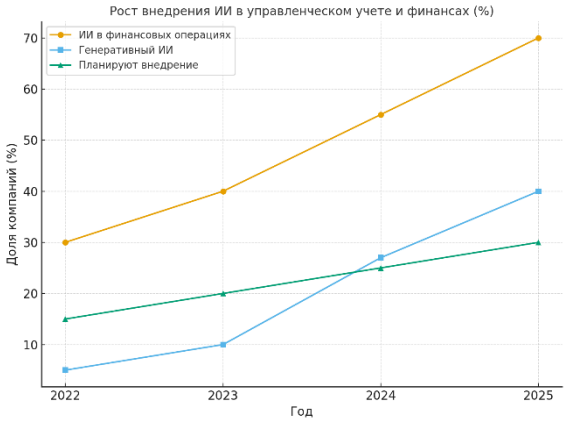


Рис. 1. Рост внедрения ИИ в управленческом учете

Внедрение технологий искусственного интеллекта в управленческий учет носит неоднородный характер в зависимости от региона и уровня развития экономики.

В развитых странах – США, странах Евросоюза и Великобритании – ИИ уже активно используется в аудиторской практике, для автоматизации сверки и обработки документов, а также для анализа больших массивов транзакций с целью выявления аномалий и мошеннических операций. Согласно исследованию консалтинговой фирмы «Deloitte», более 53 % предприятий либо уже интегрировали, либо планируют интеграцию технологий ИИ и продвинутой аналитики в системы управления затратами и рентабельностью [6]. В отчете «PwC» подчеркивается, что ИИ становится частью стандартных инструментов финансовых департаментов: от прогнозирования денежных потоков до риск-менеджмента [7].

В России внедрение идет медленнее, что связано с ограничениями в финансировании, дефицитом специалистов и сложностью адаптации западных решений к отечественным учетным системам (например, продукты фирмы «1С»). По данным Ассоциации профессионалов в области ИИ (РАИИ), к 2024 г. удельный вес российских компаний, применяющих ИИ в финансово-экономических задачах, не превышала 15–18 %, причем наибольшая активность наблюдается в крупных промышленных холдингах и финансовых компаниях [8].

Таким образом, можно отметить, что зарубежный опыт демонстрирует более высокую степень зрелости в использовании ИИ, тогда как российские компании пока находятся на этапе точечных пилотных проектов, преимущественно направленных на автоматизацию рутинных операций.

Использование искусственного интеллекта в управленческом учете открывает значительные возможности. К основным преимуществам можно отнести:

- автоматизацию рутинных задач. Ввод данных, сверка документов, формирование отчетов могут выполняться быстрее и с меньшим количеством ошибок;
- снижение риска ошибок и повышение точности расчетов. Машины менее подвержены утомлению и механическим ошибкам;
- возможность анализа больших объемов информации. ИИ способен находить скрытые закономерности и строить более сложные модели;
- прогнозирование и сценарный анализ. С помощью алгоритмов машинного обучения можно моделировать разные сценарии развития событий;
- выявление мошенничества и аномалий. Алгоритмы позволяют обнаруживать нетипичные операции, которые могут быть признаками злоупотреблений и генерировать риски экономической безопасности предприятия;
- оптимизацию использования ресурсов; автоматизация позволяет высвободить сотрудников для более аналитических и стратегических задач.

Несмотря на преимущества, ИИ имеет ряд ограничений.

Во-первых, его эффективность зависит от качества исходных данных. Ошибки или пробелы в информации могут привести к искаженным результатам.

Во-вторых, ИИ не всегда способен учитывать контекст, человеческий фактор, изменения внешней среды или нормативные требования.

Третья проблема заключается в непрозрачности многих алгоритмов – так называемом «эффекте черного ящика». Кроме того, внедрение ИИ требует значительных финансовых и организационных ресурсов, а также сопряжено с рисками экономической безопасности (в т.ч. кибербезопасности). Наконец, необходимо помнить, что ИИ пока не способен заменить человеческое суждение, ответственность и стратегическое мышление.

Оптимальная стратегия внедрения ИИ в управленческий учет предполагает осторожность и постепенность. Можно выделить следующие этапы:

1) анализ существующих процессов и выявление узких мест;

2) оценка качества данных, доступных для автоматизации;

3) выбор приоритетных участков, где ИИ принесет максимальный эффект (например, отчетность или обработка транзакций);

4) запуск пилотных проектов и оценка их эффективности;

5) масштабирование успешных решений;

6) подготовка сотрудников и формирование новой корпоративной культуры, в которой ИИ рассматривается как помощник, а не замена.

Важно избегать радикальной перестройки нефинансовых процессов исключительно ради автоматизации, так как в условиях ограниченных бюджетов это достаточно дорого и не всегда оправдано. Таким образом, глобальные автоматизированные решения, как правило, предполагают серьезные изменения в основном автоматизируемом бизнес-процессе и его администрировании. Развертывание ИИ-агентов – это решение, которое может преодолеть проблемы с обработкой корпоративной «рутины», заменяя человеческий труд и значительно ускоряя реализацию задач. В дальнейшей работе будут рассмотрены подобные примеры из реальной практики.

Перспективы развития управленческого учета в ближайшие 5–10 лет связаны с постепенным переходом к гибридной модели взаимодействия человека и искусственного интеллекта. Основная тенденция заключается не в замещении специалистов, а в трансформации их роли – от исполнителей рутинных операций к аналитикам и стратегическим советникам.

Согласно прогнозу «Gartner», к 2030 г. более 80 % финансовых процессов будут автоматизированы, однако критически важные управленческие решения останутся в зоне ответственности человека [9]. В консалтинговом отчете «EY» также отмечается, что финансовые службы будут превращаться в центры стратегической аналитики, где ИИ играет роль инструмента, а не субъекта принятия решений [10].

В перспективе развитие может пойти по следующим направлениям:

- создание объяснимых моделей ИИ («Explainable AI») для повышения доверия и прозрачности;
- интеграция ИИ в более сложные сценарии прогнозирования и стресс-тестирования;
- расширение использования интеллектуальных агентов в контуре ERP и BI-систем;
- рост внимания к этическим и правовым аспектам применения ИИ.

Думается, будущее управленческого учета неотъемлемо связано с синергией человеческого интеллекта и искусственных технологий, где главная ценность – усиление аналитического потенциала и повышение качества управленческих решений.

Стоит также отметить и область нормативно-правового (в том числе корпоративного) регулирования деятельности ИИ. Использование прогрессивных технологий, создающих возможность наделения ИИ ролью полноценного участника бизнес-процессов в управленческом учете, неизбежно связано с этическими и правовыми вызовами. Среди наиболее значимых аспектов можно выделить:

- ответственность за ошибки. Если ИИ допускает ошибку при классификации или прогнозировании, то встает вопрос о распределении ответственности: несет ли ее разработчик алгоритма, поставщик системы или конечный пользователь [13];
- прозрачность и объяснимость алгоритмов. Многие модели ИИ, как отмечалось ранее, как «черный ящик», что снижает уровень доверия со стороны руководства и аудиторов. Внедрение принципов «Explainable AI» (XAI) становится важнейшей задачей для обеспечения прозрачности принятия решений [14];
- конфиденциальность и безопасность данных. Управленческий учет оперирует чувствительной финансово-экономической информацией. Необходимы строгие меры защиты, соответствующие требованиям законодательства о персональных данных и корпоративной безопасности [12];
- нормативное регулирование. В ЕС уже принят меморандум «AI Act» [11], направленный на регулирование использования искусственного интеллекта в критически важных областях. Российское законодательство пока находится на стадии формирования, однако в перспективе ожидается разработка аналогичных норм для защиты интересов бизнеса и государства [15];
- этическая составляющая процесса замещения рабочих мест. Автоматизация рутинных функций может привести к сокращению численности персонала. Это требует выработки этически взвешенной кадровой политики, ориентированной на переквалификацию специалистов, а не только на оптимизацию затрат [13].

Таким образом, успешное внедрение ИИ в управленческий учет требует не только технологической готовности, но и решения правовых и этических вопросов, обеспечивающих

доверие к системе со стороны пользователей и общества.

2. Практические примеры внедрения ИИ в систему управленческого учета

2.1. Использование ИИ-агента для повышения прозрачности процесса выбора поставщиков

В ходе аудита расходов предприятия была выявлена системная проблема, связанная с процедурой выбора поставщиков при осуществлении мелких, но регулярных закупок для административно-хозяйственных нужд. На практике сотрудники отдела закупок ориентировались преимущественно на критерий скорости приобретения, что приводило к формированию формальных альтернативных предложений от поставщиков с заведомо завышенной ценой или неудовлетворительными условиями. Такая практика позволяла соблюсти формальные требования регламента (предоставление не менее трех коммерческих предложений), но не обеспечивала должного уровня анализа рыночной ситуации.

Для решения задачи был интегрирован интеллектуальный агент, выполняющий функции второй линии контроля качества выбора поставщика. Агент встроен в процесс согласования закупок в корпоративной CRM-системе, функционирующей в облачной инфраструктуре. На основе анализа открытых источников ИИ формирует собственную рекомендацию по конкретной закупке, сопоставимую по структуре с предложениями, подобранными менеджерами. Благодаря использованию открытого кода CRM-системы удалось минимизировать затраты на интеграцию ИИ-агента в процесс. Таким образом, был обеспечен баланс между необходимостью аналитической проверки рыночных условий и требованием сохранения операционной эффективности процесса закупок.

2.2. Интеграция ИИ-агента в систему управленческого учета для контроля и автоматизации сверки расходов и платежей

Одной из традиционных проблем управленческого учета является несоответствие между отражением платежей и последующим распределением расходов в учетных контурах бюджетов (БДР и БДДС). На рассматриваемом предприятии за проведение платежей отвечает один бухгалтер, за отражение расходов – другой, что зачастую приводило к несогласованности данных (например, выбор разных центров финансовой ответственности (ЦФО) для одной и той же операции). Ошибки носили систематический характер и были обусловлены, главным образом, человеческим фактором.

Руководством была поставлена задача «мягкой» оптимизации процесса, без масштабной перестройки учетной архитектуры. В результате был реализован проект интеграции ИИ-агента в информационную систему «1С: Комплексная автоматизация». Агент выполняет две ключевые функции:

a) контроллер ошибок – автоматически выявляет несоответствия при ручном вводе операций и предлагает корректировки (опираясь на историю взаиморасчетов в контрагентом);

b) ИИ-бухгалтер – самостоятельно относит операции на соответствующие статьи бюджета, ЦФО и проекты, опираясь на исторические записи по контрагентам.

Алгоритм ежедневно формирует реестр рекомендации к корректировкам и может действовать как в режиме консультации (с последующей проверкой), так и в полностью автономном режиме. В настоящее время, после успешного тестирования, сверка операций переносится в контур ежемесячного закрытия управленческой отчетности, что позволило существенно снизить риск ошибок и трудозатраты бухгалтерии.

2.3. Применение ИИ-ассистента для оптимизации SQL-запросов и ускорения работы с управленческой отчетностью

Функция формирования управленческой отчетности на предприятии реализована с использованием среды «Google BigQuery», где аналитика строится на сложных SQL-запросах, включающих объединение нескольких баз данных, создание временных таблиц и рекурсивные конструкции. До внедрения ИИ устранение ошибок или оптимизация запросов требовали участия ИТ-отдела, а сроки доработки могли растягиваться на недели.

Для повышения оперативности был внедрен ИИ-консультант, к которому специалисты финансово-экономической службы направляют SQL-код с уточняющим промтом, выступающим в роли технического задания. Среднее время получения корректного результата сократилось до нескольких минут. Хотя полученный код не всегда является оптимальным (а иногда даже ошибочным) и требует экспертной проверки, временные затраты на код-ревью существенно меньше, чем на традиционную доработку силами ИТ-подразделения. В итоге удалось снизить нагрузку на ИТ-отдел и повысить гибкость работы финансовой службы при подготовке управленческой отчетности.

2.4. Применение ИИ-бота для навигации по регламентам и локальным нормативным актам

Одной из наиболее острых проблем в корпоративной среде является необходимость

работы с большим объемом регламентов и локальных нормативных актов. Чтение и интерпретация этих документов часто становятся затруднительными для сотрудников, особенно при необходимости быстро принять решение в рамках конкретного бизнес-процесса.

Для решения данной задачи на предприятии был развернут специализированный ИИ-бот, интегрированный в корпоративную CRM-систему. Благодаря размещению регламентов в облачной среде, бот всегда располагает актуальной информацией. Сотрудник может задать вопрос в естественной форме («как поступить при согласовании расходов?», «какие сроки сдачи бюджета?») и получить точный ответ, основанный на положениях внутренних документов.

Особенно востребованным оказалось подключение бота к регламенту бюджетирования. Теперь пользователи оперативно получают информацию о сроках подготовки бюджета, перечне необходимых статей, а также порядке согласования, что позволяет существенно снизить количество уточняющих вопросов к финансовой службе.

Таким образом, внедрение ИИ-бота повысило эффективность коммуникаций внутри компании и снизило трудозатраты сотрудников, связанных с поиском и интерпретацией нормативной информации.

3. Заключение

По результатам проведенного исследования можно сформулировать следующие выводы:

- искусственный интеллект становится неотъемлемой частью цифровизации управленческого учета, обеспечивая автоматизацию рутинных задач, повышение точности расчетов и расширение аналитических возможностей предприятия;
- ИИ не способен полностью заменить человека в финансово-экономических службах, поскольку не учитывает контекст, нефинансовые факторы и стратегические приоритеты компании. Ключевыми остаются человеческий контроль и интерпретация данных;
- масштабная перестройка бизнес-процессов ради автоматизации экономически не оправданна для большинства компаний в условиях ограниченных ресурсов. Оптимальным является выборочная автоматизация узких мест;
- практика показывает, что использование ИИ позволяет снизить количество ошибок, ускорить процессы сверки и анализа, а также повысить гибкость финансовых служб за счет разгрузки от рутинных операций;
- Наиболее эффективным является поэтапный подход: диагностика процессов – пилотные проекты – масштабирование успешных решений. Такой путь снижает риски и обеспечивает плавную интеграцию технологий.
- Перспективы связаны с синергией человека и ИИ: технологии выполняют вспомогательную и аналитическую функцию, а ответственность за принятие решений и стратегическое управление остается за специалистами.

Таким образом, цифровизация управленческого учета с применением искусственного интеллекта представляет собой естественный этап развития корпоративных систем управления. ИИ способен существенно повысить эффективность процессов, снизить рутинную нагрузку и сократить время подготовки отчетности. Однако его внедрение должно быть осторожным и поэтапным, с сохранением ключевой аналитической роли за специалистами. Будущее управленческого учета, вероятно, связано с синергией человека и технологий, где ИИ выполняет вспомогательную роль, а ответственность за принятие решений остается за человеком. По крайней мере, на сегодня.

Список используемых источников

1. KPMG. Global AI in Finance Report [Электронный ресурс]. – Режим доступа: https://kpmg.com/xx/en/our-insights/ai-and-technology/kpmg-global-ai-in-finance-report.html (дата обращения: 23.09.2025).
2. Mordor Intelligence. Artificial Intelligence in Accounting Market Report [Электронный ресурс]. – Режим доступа: https://www.mordorintelligence.com/industry-reports/artificial-intelligence-in-accounting-market (дата обращения: 23.09.2025).
3. Market Research Future. AI in Accounting Market Report [Электронный ресурс]. – Режим доступа: https://www.marketresearchfuture.com/reports/ai-in-accounting-market-22351 (дата обращения: 23.09.2025).
4. Accounting Today. AI implementation at firms grows 34% in one year [Электронный ресурс]. – Режим доступа: https:// www.accountingtoday.com/news/ai-implementation-at-firms-grows-34-in-one-year (дата обращения: 23.09.2025).
5. The Accountant Online. ACCA Report: AI adoption in accounting [Электронный ресурс]. – Режим доступа: https://www.theaccountant-online.com/news/acca-report-ai-adoption-accounting/ (дата обращения: 23.09.2025).
6. Deloitte & IMA. Survey on Cost and Profitability Management [Электронный ресурс]. – Режим доступа: https://www.internationalaccountingbulletin.com/news/deloitte-ima-survey-accounting/ (дата обращения: 23.09.2025).

7. PwC. Finance in the Digital Era: Global Survey [Электронный ресурс]. – Режим доступа: https://www.pwc.com/gx/en/services/people-organisation/workforce-of-the-future.html (дата обращения: 23.09.2025).
8. Ассоциация развития искусственного интеллекта (РАИИ). Доклад о развитии ИИ в России, 2024 [Электронный ресурс]. – Режим доступа: https://raii.ru/research (дата обращения: 23.09.2025).
9. Gartner. Future of Finance: AI and Automation Report [Электронный ресурс]. – Режим доступа: https://www.gartner.com/en/finance/insights/future-of-financ (дата обращения: 23.09.2025).
10. EY. The DNA of the CFO: Is the future of finance new technology? [Электронный ресурс]. – Режим доступа: https://www.ey.com/en_gl/dna-of-the-cfo (дата обращения: 23.09.2025).
11. European Commission. Artificial Intelligence Act (AI Act). Regulation proposal [Электронный ресурс]. – Режим доступа: https://artificialintelligenceact.eu/ (дата обращения: 23.09.2025).
12. OECD. OECD Principles on Artificial Intelligence [Электронный ресурс]. – Режим доступа: https://oecd.ai/en/ai-principles (дата обращения: 23.09.2025).
13. Jobin A., Ienca M., Vayena E. The global landscape of AI ethics guidelines // Nature Machine Intelligence. – 2019. – Vol. 1, № 9. – P. 389–399. – DOI: 10.1038/s42256-019-0088-2.
14. Floridi L., Cowls J. A Unified Framework of Five Principles for AI in Society // Harvard Data Science Review. – 2021. – Vol. 3, № 1. – DOI: 10.1162/99608f92.13bb3fe8.
15. Министерство цифрового развития, связи и массовых коммуникаций Российской Федерации (Минцифры РФ). Концепция развития регулирования искусственного интеллекта и робототехники до 2030 года [Электронный ресурс]. – Режим доступа: https://digital.gov.ru (дата обращения: 23.09.2025).

HR-бренд в портфеле брендов организации

Рувенный И. Я.
Уфимский университет науки и технологий
кафедра стратегического управления
Уфа, Россия
e-mail: ruvenny@mail.ru

Аннотация

Статья посвящена изучению современного рынка труда, характеризующегося высоким уровнем конкуренции за высококвалифицированных сотрудников. Рассматриваются ключевые элементы портфеля брендов организации, в частности, корпоративный и продуктовый бренды, личный и HR-бренды.

Основной целью исследования является выявление механизмов эффективного формирования интегрированной системы брендов, обеспечивающей рост результативности деятельности организации.

Автор анализирует существующие определения HR-брендинга, изучает зарубежный опыт и формулируют собственное представление о данном феномене. Используются подходы к разработке и оценке бренд-платформ, отражающие особенности бренда работодателя.

Определены недостатки традиционных представлений о HR-брендинге, предложены пути оптимизации системы брендов организации. Подчеркнута важность формирования позитивного образа работодателя как фактора успешного привлечения и удержания талантливых сотрудников.

Подтверждается ключевая роль правильного понимания роли каждого вида бренда в рамках общей стратегии организации. Эффективное использование HR-брендинга помогает решить задачи кадровой политики, укрепить позицию на рынке труда и обеспечить долговременное развитие организации. Это способствует улучшению конкурентоспособности, формированию позитивного корпоративного климата и усилению связи между брендом работодателя и другими брендами организации.

Ключевые слова: двойная конкуренция, HR-бренд, репутация работодателя, клиентоориентированность, бренд-платформа, Employee Value Proposition (EVP).

1. Введение

Современному рынку свойственен высокий уровень двойной конкуренции, не только непосредственно за клиентов, но и за высококвалифицированный и лояльный к организации персонал. С этой точки зрения, для организаций все возрастающее значение имеет система брендов, включающая два ключевых элемента: бренды бизнеса и бренды человека.

Бренды бизнеса представлены корпоративным брендом и продуктовым брендом. Корпоративный бренд включает представления, мнения и ценности организации в глазах внешних стейкхолдеров: партнеров, клиентов, общества и государства. Основными элементами корпоративного бренда являются видение, миссия, ценности и рыночная позиция организации [1]. Продуктовый бренд представляет собой отдельную единицу бренда организации, имеющую собственную идентичность, ценности и позиционирование на рынке. Основными элементами продуктового бренда являются: дифференциация продукта, воспринимаемое потребителем качество, пути формирования спроса на конкретный продукт, рост рыночной стоимости организации, связь с корпоративным брендом, коммуникация с потребителями, конкурентоспособность продукта.

Бренды человека представлены личным брендом и HR-брендом. HR-бренд предназначен для фиксации позитивного отношения к организации как к работодателю. Он нацелен на привлечение лояльных и талантливых сотрудников. Личный бренд представляет собой индивидуальный набор характеристик, компетенций и идентичностей, с помощью которых личность презентует себя обществу и рынку. Личный бренд важен в профессиональном росте, социальной активности, личной жизни и влияет на восприятие человека окружающими, принятие решений по выбору профессии, карьеры и личностного самоопределения [2]. Он включает профессиональные компетенции, личные качества, опыт, знания, навыки, поведение и репутацию человека.

Каждый из этих элементов брендинга решает свои собственные задачи и ориентирован на разные целевые группы, но все эти элементы тесно связаны друг с другом. Бренды бизнеса нацелены на улучшение восприятия организации клиентами и партнерами, а бренды человека сконцентрированы на укреплении позиций организации на рынке труда. Эффективные бренды бизнеса повышают доверие и улучшают репутацию организации, привлекают стратегических партнеров и формируют потребительскую лояльность. Успешные бренды

человека снижают текучесть кадров, повышают удовлетворенность сотрудников, производительность труда, способствует развитию компетенций и навыков сотрудников.

2. Публикация

Организациям важно учитывать связь между этими двумя направлениями брендинга, обеспечивать согласованность внутреннего и внешнего образов организации для роста результативности брендинговых усилий. В основе обоих направлений брендинга лежат миссия и ценности организации, поэтому инструменты интеграции этих направлений должны включать разработку четких норм корпоративного поведения, четкую координацию информационных потоков и выработку единых подходов к подаче информационного материала. Процесс формирования брендов начинается с глубокого понимания потребностей своей целевой аудитории и выявления факторов, влияющих на принятие решений. Система брендов организации представлена в таблице 1.

Таблица 1. – Интегрированная система брендов организации [составлено автором]

Параметр сравнения	Бренды человека		Бренды бизнеса	
	HR-бренд	Бренд личности	Корпоративный бренд	Продуктовый бренд
Целевая аудитория	Сотрудники кандидаты на вакансии	Общественность, поклонники, эксперты, работодатели	Инвесторы, партнеры, государственные органы	Потребители, пользователи
Цель	Привлечение и удержание талантов	Распространение личного авторитета, экспертности	Укрепление репутации, привлечение инвесторов	Увеличение продаж, завоевание доли рынка
Носители	Соцсети, внутренние ресурсы, ярмарки вакансий	Биография, интервью, социальные сети, выступления	Сайт, отчёты, официальные публикации	Упаковка, реклама, цифровые медиа, магазины
Каналы продвижения	Профессиональные форумы, соцсеть, личные встречи	Онлайн-конференции, мастер-классы, публичные выступления	Мероприятия, выступления, публикации в прессе	Интернет-маркетинг, ТВ, радио, наружная реклама
Элементы бренда	EVP, программы мотивации, карьера	История жизни, личный опыт, харизма, убеждения	Видение, миссия, корпоративные ценности	Название, упаковка, уникальный торговый знак (УТЗ)
Атрибуты	Символы принадлежности к компании, бейджи	Фото, цитаты, символика личной истории	Логотип, фирменный стиль	Внешний вид упаковки, рекламные образы, слоганы
Тональность	Лояльность, командный дух	Человеческие качества, история преодоления трудностей	Финансовая стабильность, инновации	Польза продукта, эмоция от потребления
Особенность	Поддержка карьерного роста	Харизма, неповторимый личный путь	Уникальное положение на рынке	Особые свойства, функциональность продукта
Результат	Улучшение корпоративной культуры, сокращение текучести кадров	Признание, авторитет, общественное влияние	Повышение стоимости акций, инвестиции	Рост прибыли, увеличение числа постоянных клиентов

Ролевая модель аудитории	Специалисты высокого класса, молодёжь	Известные лица, лидеры мнений, эксперты	Управленческий персонал крупных предприятий	Широкий круг потребителей
Характеристики эффективности	Открытость, уважение, забота о сотрудниках	Авторитет, вдохновение, уникальная личность	Надежность, успех, стабильность	Качество, новизна, удовлетворение нужд клиента

Основными отправными точками при формировании брендов бизнеса являются уникальные видение и миссия, конкурентные преимущества и стратегические перспективы. Методы продвижения брендов включают участие в социальных проектах, информационную открытость перед партнерами и инвесторами. Ключевыми моментами при формировании успешных брендов человека являются активное участие в профессиональной среде, совершенствование кадровых процедур, программы карьерного роста, мотивация сотрудников, личное развитие. Успешная интеграция всех брендов обеспечивает организацию преимуществами как на рынке труда, так и на рынке партнерских взаимодействий. Ключевым фактором успеха в интеграции является правильное понимание роли каждого вида бренда в эффективной реализация согласованной стратегии.

Современный рынок актуализирует необходимость привлечения и удержания организациями самых перспективных, творческих и амбициозных сотрудников из числа потенциальных кандидатов на вакантные должности. Эта тенденция свидетельствует о растущей значимости рационального подбора кадров, который должен осуществляться с применением инновационных инструментов, расширяющих рамки привычного рекрутмента. Одним из таких прогрессивных инструментов выступает HR-брендинг, предполагающий формирование позитивного имиджа организации как привлекательного работодателя.

Брендинг работодателя, как правило, связан с двумя направлениями его деятельности: маркетингом и управлением персоналом. Но, если посмотреть шире, брендинг работодателя тесно связан и со стратегическим менеджментом. Поэтому, необходимо воспринимать брендинг работодателя как важную часть стратегического управления персоналом, бизнес-стратегии и конкурентоспособности.

Выделяются следующие функции HR-брендинга [3]:

– комплексный подход к управлению всеми брендами организации (корпоративным брендом, HR-брендом, потребительским брендом, брендом личности);

– сопровождение внутриорганизационных изменений, необходимых для актуализации бренда работодателя;

– управление коммуникациями в организации на основе двухстороннего обмена информацией между работниками и клиентами;

– создание определенного стиля жизни работников в организации.

HR-бренд никогда не строится на пустом месте. Он, в том или ином виде, всегда уже существует в результате использования или, наоборот, неиспользования какой-либо стратегии. Основные элементы бренда работодателя формируются на основе его корпоративной идеологии, которая существует всегда. Если организация сознательно не формирует свой бренд работодателя, то она имеет ограниченное влияние на контактные аудитории только качеством своей продукции и не влияет с точки зрения ее восприятия как работодателя.

Исследование проблем брендинга в HR-сфере началось относительно недавно. Сам термин HR-брендинг предложил, а в дальнейшем и развил в своих работах S. Barrow [4]. Вопросы HR-брендинга активно рассматривали такие зарубежные ученые как B. Adams и C. Marshall [5], B. Minchington [6], М. Армстронг [7].

Рассмотрим имеющиеся в научной литературе интерпретации сущности HR-брендинга.

1. «Комплекс мер по созданию привлекательной репутации компании на рынке труда» [8].

2. «Работа с репутацией компании как работодателя» [9].

3. «Внимание к сотрудникам, качественная работа с персоналом внутри компании» [10].

4. «Комплекс целенаправленных мероприятий по формированию положительного имиджа работодателя с целью постоянного привлечения лучших из лучших специалистов в своей отрасли» [11].

5. «Передовой инструмент для привлечения сотрудников - основного актива компании» [12].

6. «Непрерывная работа, направленная на интеграцию бренда в персонал или наоборот персонала в бренд, которая продолжается на протяжении всей жизни компании» [13].

Встречаются и определения HR-брендинга, акцентирующие внимание на привлечении квалифицированных сотрудников нематериальными стимулами, которые зачастую являются не очень уловимыми, но эффективными в плане мотивации [14].

Проанализировав некоторые определения, можно обозначить систематизированное сложившееся представление о HR-брендинге. HR-брендинг представляет собой процесс, направленный на формирование и распространение позитивного имиджа организации как работодателя. Он представляет собой набор характеристик организации, способствующих привлечению и удержанию высококвалифицированных сотрудников. HR-брендинг реализуется через систему мероприятий и методов, способствующих повышению привлекательности организации для талантливых специалистов. Он основан на сбалансированном сочетании специального HR-образа и корпоративного бренда, обладающем потенциалом дифференциации организации в восприятии целевых аудиторий и стимулирования их долгосрочной заинтересованности в сотрудничестве. Опционально HR-брендинг может быть использован в качестве действенного инструмента мотивации работников, способствующего обнаружению и приросту их потенциала, направляемого на достижение командных и индивидуальных целей.

Анализ преобладающего понимания HR-брендинга позволил выявить ряд его недостатков.

1. Привлечения квалифицированных сотрудников недостаточно. Все-таки, основное – способствовать производительной работе персонала, в том числе, с помощью качественного подбора кандидатов с учетом соответствия их ценностей ценностям организации. Иными словами, HR-брендинг нацелен не просто на привлечение работников, а на повышение производительности их труда.

2. HR-брендинг относится к значимым аспектам корпоративной культуры, формируя внутреннюю и внешнюю коммуникацию. Он представляет собой постоянно идущий, непрерывный, а не однократный процесс. Однако существующие определения практически не затрагивают связь HR-брендинга с корпоративной культурой, недооценивая его долгосрочное влияние на нее.

3. Несмотря на признаваемую роль HR-брендинга в достижении миссии организации, не указывается его прямое влияние на взаимоотношения с клиентами. В то же время, успешный HR-брендинг ведет к увеличению доходов на основе улучшения взаимодействия с клиентами. Иногда HR-брендинг связывают лишь с рынком труда, но его полный функционал затрагивает всю систему отношений с клиентами и партнерами организации.

На основе проведенного анализа предлагается авторское определение: HR-брендинг – стратегический маркетинговый инструмент, содержащий комплекс определяемых корпоративной культурой формальных и неформальных мероприятий по управлению репутацией работодателя во внешней и внутренней среде и используемый в интегрированной деятельности подразделений маркетинга и управления персоналом для роста клиентоориентированности персонала и организации в целом.

Причины развития бренда работодателя.

1. Изменение внешней макро и микросреды бизнеса. В настоящее время внешняя среда бизнеса непредсказуемо меняется, и работодатели изыскивают новые возможности получения конкурентных преимуществ, позволяющих приспособиться к этим изменениям.

2. Глобализация. Многие организации зачастую осуществляют бизнес не только на региональном и национальном уровнях, но и на транснациональном уровне [15]. Инновационные бренд-технологии основываются на том, что создать HR-бренд лишь на отдельном рынке невозможно.

3. Трансформация структуры профессиональных компетенций и изменение отношения к работе создают новые вызовы в поиске и удержании ценных сотрудников.

4. Прозрачность. Сегодня все люди с помощью онлайн-каналов имеют возможность распространять свое мнение по разным вопросам на широкую аудиторию. Поэтому информация, поступающая от работодателя на рынок труда, обязательно должна быть достоверной, так как ее могут верифицировать нынешние и бывшие сотрудники.

5. Конкурентоспособность. В настоящее время многие работодатели рассматривают управление талантами ка важное условие построения конкурентоспособной организации в долгосрочной перспективе.

В современных организациях управление персоналом, во-многом, нацелено на привлечение и удержания сотрудников, являющихся человеческим капиталом, что в итоге повышает рыночную стоимость бизнеса. Можно отметить эволюционный характер бренда, который сейчас распространяется не только на продукцию, но и на людей. HR-бренд строится на множестве факторов, включающих уровень заработной платы, комфортные и безопасные условия труда, перспективы карьерного роста и возможности дополнительного образования. Подобные условия труда воспринимаются работниками и как основные, и как дополнительные мотивы для принятия положительного решения в отношении работодателя. Исследователи выделяют рациональное, эмоциональное и социальное направления формирования лояльности к бренду [16]. Приспособим указанные направления к бренду работодателя.

1. Рациональное направление связано с материальной стороной, по которой многие организации предлагают похожие условия, что минимизирует возможности влияния на потенциальных сотрудников.

2. Эмоциональное направление основано на опыте работы в организации и субъективных ощущениях, таких как удовлетворенность содержанием работы, уважение коллег, гармония с ценностями организации. Эмоциональное направление в значительной степени повышает привлекательность HR-бренда, способствуя распространению положительных отзывов.

3. Социальное направление затрагивает репутацию самого работодателя в обществе, что влияет на воспринимаемый имидж сотрудника, работающего в уважаемой организации, в кругу его общения.

Формирование сильного бренда работодателя требует наличия понятного для понимания сотрудниками формата адаптации ценностей и стратегических ориентиров. В HR-брендинге актуальным инструментом создания такого формата является ценностное предложения работодателя (EVP – Employee Value Proposition), которое описывает совокупность условий и преимуществ, представляющих в данной организации значимую выгоду от трудоустройства. Иными словами, это атрибуты работодателя, отличающие его от других организаций, набор ключевых преимуществ, предлагаемых сотруднику в обмен на его навыки, усилия и время (вознаграждения, система компенсаций и льгот, известность и положение организации на рынке, условия труда, местоположение, обучение и возможности карьерного роста, масштаб решаемых задач, отношения в коллективе, корпоративная культура, признание).

Конструкция HR-бренда, позволяющая сформулировать EVP, состоит из распространенных ключевых ценностей, значимых для большинства потенциальных соискателей.

1. Экономическая ценность (заработная плата выше среднеотраслевой, достойный социальный пакет, гарантия занятости, карьерный рост).

2. Ценность развития (признание заслуг, повышение самооценки, чувство уверенности при трудоустройстве в будущем).

3. Ценность интереса (интересная рабочая среда, применение современных методов работы, задействование творческого потенциала сотрудников).

4. Ценность применения (возможность для сотрудников применять все свои знания и развивать коллег в клиентоориентированной среде).

5. Социальная ценность (позитивная доброжелательная рабочая среда, благоприятная командная атмосфера).

Ценностное предложение работодателя (EVP) традиционно опирается на три классических направления.

1. Привлечение. Для привлечения используются разнообразные методы, такие как публикации в СМИ, ярмарки вакансий, корпоративные презентации, выступления руководителей на конференциях, личное общение с кандидатами и активные приглашения интересных специалистов.

2. Вовлеченность. Организация мероприятий и реализация инициатив, способствующих погружению сотрудников в жизнь организации: экскурсии на производство, дни «открытых дверей», встречи с топ-менеджерами, тематические семинары, конкурсы и викторины, посвященные корпоративной истории и ценностям.

3. Удержание. Применение внутрикорпоративных коммуникаций и традиций, таких как совещания, письменные отчеты о результатах, награждения и премии, наставничество, поддержка коммуникаций с руководителями, развитие сотрудников на курсах повышения квалификации и стажировках.

В ходе исследования, проведенного агентством SuperJob среди ведущих работодателей России были выявлены факторы, которые наиболее часто используют организации в своих объявлениях о вакансиях для привлечения соискателей:

– заработная плата по рыночным стандартам – 99%;

– предоставление достойного социального пакета – 99%;

– реальные перспективы карьерного и профессионального роста – 98%;

– гарантия стабильной занятости на продолжительный период времени – 90%;

– стабильное развитие организации в стратегической перспективе – 85%;

– наличие образовательных возможностей и курсов повышения квалификации – 85%;

– стремление демонстрировать высокий уровень профессионализма сотрудников – 80%;

– постановка увлекательных и сложных профессиональных задач – 75%;

– комфортабельные условия труда и оборудованное рабочее пространство – 70% [17].

Полученные данные подтверждают растущую конкуренцию за лучшие кадры на российском рынке труда. Согласно данному исследованию, свыше половины опрошенных организаций указывают минимум пять пунктов из перечисленных выше в объявлении о вакансии.

Чтобы выделиться на фоне конкурентов и выгодно позиционировать организацию как работодателя и необходим HR-брендинг. Дело в том, что у каждой организации, принявшей на работу сотрудников, автоматически формируется определенный образ в глазах как существующего персонала, так и потенциальных кандидатов. Этот образ уже сам по себе является HR-брендом. Однако систематически заниматься созданием и развитием своего HR-бренда пока готовы не более 30% российских организаций. Существует стереотип, что HR-брендинг актуален только для крупных организаций, но, на самом деле, такой подход к решению задач подбора качественного персонала применим для организаций любых размеров. По мере роста численности сотрудников лишь меняется фокус HR-задач: в организациях численностью до 100 человек остро стоит проблема повышения вовлеченности персонала, а в крупных организациях – проблема экономии средств на подбор персонала и повышения эффективности массовых наймов. Крупные российские организации тратят около 15% своего HR-бюджета на массовое привлечение сотрудников, в то время как малый бизнес вынужден тратить вдвое больше средств в расчете на закрытие одной вакансии из-за недостатка профессионального подхода к HR-брендингу.

Таким образом, работа над HR-брендом становится важной составной частью стратегии любого бизнеса, независимо от его масштаба, обеспечивая как привлечение и удержание ценных специалистов, так и укрепление конкурентных позиций организации на рынке труда.

Оптимизация структуры EVP осуществляется на основе бренд-платформы, выступающей ключевым механизмом выявления сущностных характеристик бренда, его дифференцирующих свойств и преимуществ [18]. Методика разработки бренд-платформы предусматривает оценку бренда работодателя по ряду критически важных параметров, позволяющих детально охарактеризовать и качественно описать бренд работодателя [19]. Эти критерии позволяют выявить основные достоинства бренда, продемонстрировать его уникальность и обосновать причины привлекательности организации для потенциальных сотрудников.

Представляется, что оценка бренда работодателя должна учитывать параметры, представленные в таблице 2.

Таблица 2. – Оценочные параметры HR-бренда [составлено автором]

Параметры оценки HR-бренда	Объект оценивания	Показатели оценивания
Воспринимаемая репутация	Внутренняя среда (организационная культура и ценностные установки, транслируемые организацией соискателям и сотрудникам)	Уровень привлекательности организации как работодателя Внутренние нормы и правила организации Корректировка нормативных установок в контексте корпоративной этики
Информационно-медийный фон	Контактная среда (деятельность HR-службы, направленная на информирование сотрудников и соискателей)	Каналы трансляции информационного поля организации Способность сообщений отразить основную идею корпоративной идеологии Взаимодействие HR-сервисов с медиа пространством
Практика HR-менеджмента в организации	Среда восприятия (отношение и восприятие сотрудников и соискателей)	Уровень осведомленности о миссии и целях организации Степень готовности рекомендовать организацию третьим лицам Отсутствие намерения смены работы в краткосрочном периоде Активность участия в процессах совершенствования организации

EVP лежит в основе формирования ассоциативной сети HR-бренда организации, которая представляет собой совокупность связей, формирующихся на основе ассоциаций между названием организации, ее брендом и другой информацией, находящейся в мозгу человека. Благодаря HR-брендингу, в сознании людей должны появиться положительные ассоциации, основанные на общечеловеческих и профессиональных ценностях. Труднее это осуществить в тех случаях, когда организация уже ассоциируется с негативом, так как отрицательные впечатления запоминаются быстрее и существуют устойчивее, чем положительные.

EVP представляет собой набор выгод и преимуществ, предлагаемых сотрудникам работодателем, включая материальные вознаграждения, возможности развития карьеры, корпоративную культуру и рабочие условия. Ассоциативная сеть HR-бренда формируется путем восприятия сотрудниками этих ценностей и дальнейшего распространения позитивных

ассоциаций среди потенциальных кандидатов и общественности. Ассоциативная сеть HR-бренда представляет собой сложную систему взаимосвязей, сформированных в сознании людей, основанную на воспоминаниях, эмоциях и впечатлениях, вызванных наименованием организации, ее продукцией и конкретными действиями. В процессе HR-брендинга формируются позитивные ассоциации, которые усиливают восприятие организации как достойного работодателя. Следует отметить, что создание подобных ассоциаций может оказаться проблематичным, если уже сформировался отрицательный образ. Поэтому важно правильно связать HR-бренд с самим названием организации. Для этого наименование организации обязательно должно присутствовать в рекламных и HR-материалах, презентациях и документах. Игнорирование этого принципа приведет к тому, что HR-бренд не будет ассоциироваться с конкретной организацией.

Рассмотрим набор ассоциаций, которые могут и должны возникать у кандидатов при упоминании бренда работодателя для усиления его привлекательности. Они представлены в таблице 3.

Таблица 3. – Основные категории ассоциативной сети HR-бренда [составлено автором]

Категория	Описание категории
Надежность	Организация зарекомендовала себя как стабильная и финансово устойчивая и занимает надежные позиции в своей отрасли. Она проявляет заботу о своих сотрудниках, обеспечивая необходимой поддержкой и защищая их интересы, даже, в кризисы
Перспективность	Организация предоставляет возможности для профессионального и карьерного роста, что позволяет сотруднику видеть свою перспективу. Сотрудники учатся новому, получают знания и опыт, полезные для карьеры. Деятельность создает ощутимые перспективы развития и признания
Профессионализм	Организация привлекает настоящих профессионалов, готовых делиться знаниями и помогать новичкам осваиваться в профессии. Коллеги обладают необходимыми навыками и квалификацией, поддерживают высокий уровень профессионализма. Проблемы решаются быстро и по-деловому
Сплоченность	Сотрудники чувствуют себя частью дружного коллектива, объединены общей целью и командным духом. Общий настрой на успех объединяет команду. Атмосфера дружбы и взаимопомощи создает ощущение уверенности и безопасности
Престижность	Работа в организации всеми сотрудниками оценивается как престижная, статусная и является символом профессионализма и признанием заслуг
Инновационность	Постоянное развитие сотрудников и их навыков, открытость к обучению и освоению новых технологий, что вдохновляет сотрудников на творчество и инициативу. Внедрение новых технологий позволяет организации опережать конкурентов
Социальность	Участие организации в жизни общества и поддержка своих сотрудников. Высокий уровень заботы о здоровье и комфорте сотрудников, включая медицинское страхование, спортивные залы, питание и отдых. Уделяется внимание социальной ответственности, благотворительности и экологической безопасности
Честность	Соблюдение организацией обязательств перед сотрудниками, открытость и прозрачность. Надежность и открытость руководителей создают ощущение спокойствия и предсказуемости
Законность	Соответствие законодательству РФ и добросовестность в ведении бизнеса, защита интересов сотрудников. Следование требованиям законодательства формирует образ организации как безопасной и ответственной

Кроме представленных в таблице 3 основных универсальных категорий, существуют и специфические категории ассоциативной сети, характерные для отдельных отраслей деятельности. Например, финансовые учреждения могут подчеркивать надежность вкладов и инвестиций, медицинский учреждения – высокий уровень помощи и безопасность пациентов, производственные предприятия – экологическую чистоту продукции и инновационные технологии.

В идеале ассоциация с брендом работодателя должна быть максимально полной и яркой, затрагивать все важнейшие аспекты привлекательности организации для работы в ней. Наличие хорошо структурированной ассоциативной сети позволяет организации уверенно конкурировать за лучшие кадры на рынке труда и при необходимости поддерживать стабильный приток талантливых сотрудников.

3. Заключение

Управление HR-брендом постепенно становится актуальной деятельностью для любой организации, стремящейся эффективно пользоваться имеющимся у нее человеческим капиталом. Преимущества, получаемые как организацией целом, так и отдельными сотрудниками от HR-бренда очевидны. Для организации сильный бренд означает уменьшение текучести ценных специалистов и снижение расходов на пополнение штата новыми кадрами. Сотрудники, работающие в организации с мощным HR-брендом, получают возможность для самореализации и гордости за принадлежность к известной организации. В результате такая двусторонняя польза способствует накоплению стратегического потенциала организации, укрепляя ее позиции на рынке и повышая уровень конкурентоспособности.

Список используемых источников

1. Бикметов Е.Ю., Рувенный И.Я. Стратегическое управленческое мышление: проблемы формирования и средства оценки // Инновационные технологии управления социально-экономическим развитием регионов России: Материалы XII Всероссийской научно-практической конференции с международным участием. Часть II. Уфа: ИСЭИ УФИЦ РАН, 2020. С. 94–99.

2. Бикметов Е.Ю., Рувенный И.Я., Шаяхметова Р.Р. Социальный маркетинг как средство социально ответственного стратегического управления бизнес-организациями // Вестник Пермского национального исследовательского политехнического университета. Социально-экономические науки. 2021. № 3. С. 183–196.

3. Бикметов Е.Ю., Кузнецова Е.В., Рувенный И.Я.: Современные коммуникационные технологии в маркетинге уч. пособие для вузов. Уфа: РИЦ УУНиТ, 2024.

4. Mosley R., Barrow S. The employer brand: Bringing the Best of Brand Management to People at Work. John Wiley & Sons Limited, 2019. 234 p.

5. Adams B., Marshall C. Give & Get Employer Branding: Repel the Many and Compel the Few with Impact, Purpose and Belonging. Houndstooth Press, 2020. 354 p.

6. Minchington B. Employer Brand Leadership. 2010. 250 p.

7. Армстронг М. Главный учебник HR в мире. М.: Эксмо, 2022.

8. Балашова Е. HR-брендинг: пустые траты или долгосрочные инвестиции? URL: http://www.top-personal.ru/issue.html?1782 (дата обращения: 15.08.2025).

9. Глебова Н. HR-брендинг. URL: http://www.myshared.ru/slide/109472/ (дата обращения: 15.08.2025).

10. Куликова К.С. Методический инструментарий HR-брендинга российских организаций. URL: http://www.rusnauka.com/5_SVMN_2013/Economics/6_116694.doc.htm (дата обращения: 15.08.2025).

11. Мансуров Р.Е. HR-брендинг. Как повысить эффективность персонала. Спб.: БХВ-Петербург, 2011.

12. Минчингтон Б. HR-бренд: как стать лидером. М.: Юнайтед Пресс, 2011.

13. Пять заблуждений об HR-брендинге. URL: http://www.advertology.ru/article111859.htm (дата обращения: 15.08.2025).

14. Рувенный И.Я. HR-брендинг как направление клиентоориентированного подхода к управлению компанией // Управление экономикой: методы, модели, технологии: международная научная конференция. Сборник научных трудов. Уфа: УГАТУ, 2014. С. 193–196.

15. Виноградов А.В., Кузнецова Е.В., Рувенный И.Я. Критерии и эффективные стратегии достижения технологического суверенитета России // Вестник БИСТ (Башкирского института социальных технологий). 2025. № 2 (67). С. 106–113.

16. Рувенный И.Я., Аввакумов А.А. Управление лояльностью потребителей: учебное пособие. Уфа: УГАТУ, 2015. 194 с.

17. Рынок труда в 2023: тенденции и прогнозы SuperJob. URL: https://www.superjob.ru/research/articles/114374/rynok-truda-v-2023 (дата обращения: 15.08.2025).

18. Шафранская И.Н., Катаев Е.П. Об одном подходе к оценке бренда работодателя // Маркетинг и маркетинговые исследования. 2015. № 3 (117). С. 228–237.

19. Касимова Э.Р., Кузнецова Е.В., Рувенный И.Я. HR-брендинг как инструмент стратегического маркетингового управления // Бизнес. Образование. Право: Вестник Волгоградского института бизнеса. 2024. № 2 (67). С. 89–94.

Моделирование предиктивного метода принятия решений для распределения ресурсов в вычислительных системах

Трушкин Е. С.
Пермский национальный исследовательский политехнический университет
Пермь, Россия
e-mail: egor.s.trushkin@gmail.com

Фрейман В. И.
Пермский национальный исследовательский политехнический университет
Пермь, Россия
e-mail: vifrejman@pstu.ru

Аннотация[1]

В статье рассматривается проблема принятия решений для распределения ресурсов в вычислительных системах с динамически изменяющейся нагрузкой. В качестве решения предлагается предиктивный метод, основанный на комбинации аналитического расчета и статистического анализа исторических данных о выполнении задач. Особое внимание уделяется использованию взвешенного прогноза, который позволяет оценить будущую загрузку узлов системы до момента назначения задач, что является ключевым отличием от классических подходов.

Для подтверждения эффективности метода разработана имитационная модель вычислительной системы с узлами различной производительности. Проведено сравнительное моделирование, демонстрирующее преимущество в эффективности предложенного подхода над классическими методами по ключевым метрикам: общему времени выполнения набора задач, равномерности распределения нагрузки и точности прогнозирования.

Помимо преимуществ, также рассматриваются ограничения метода, связанные с необходимостью предварительного накопления репрезентативной статистики для прогнозирования и вычислительной сложностью расчета взвешенных оценок для каждой новой задачи. Разработанный подход может быть применен при создании адаптивных систем управления для современных вычислительных инфраструктур.

Ключевые слова: принятие решений, распределение ресурсов, математическое моделирование, прогнозирование нагрузки, вычислительные системы.

1. Введение

Эффективное распределение вычислительных ресурсов остается одной из наиболее актуальных проблем в области управления современными информационными системами[1]. Рост объемов обрабатываемых данных и увеличение сложности вычислительных задач предъявляют повышенные требования к методам планирования и распределения ресурсов[2,3]. Большинство применяемых подходов, как правило, основываются на статических[4-6] или динамических алгоритмах[7-9], в которых решение принимается на основе текущего состояния системы или фиксированных правил. Эти методы имеют ряд существенных недостатков: они не способны заранее учитывать возможные изменения в нагрузке и структуре поступающих задач, не адаптируются к внезапным пиковым нагрузкам и, как следствие, могут приводить к перегрузке отдельных узлов или неравномерному использованию ресурсов[10].

В связи с этим возникает необходимость разработки новых методов принятия решений для распределения ресурсов, основанных на прогнозировании будущей нагрузки и превентивном принятии решений. Такие методы должны учитывать, как текущее состояние системы, так и статистическую информацию о выполнении предыдущих задач, что позволяет оптимизировать распределение ресурсов до возникновения потенциальных проблем.

2. Описание метода

Как отмечается в работах [11,12], в реальных вычислительных системах наблюдается определенная повторяемость и схожесть поступающих задач. В этом и заключается основная идея – задачи одного класса обладают определенной повторяемостью по времени выполнения и требуемым ресурсам. Это свойство позволяет

[1] Труды XI Международной научной конференции "Информационные технологии интеллектуальной поддержки принятия решений", 11-13 ноября, Уфа-Чандигарх-Фандербейлпарк, 2025

прогнозировать характеристики новой задачи до момента ее назначения на узел системы.

Предлагаемый метод можно рассматривать как инструмент поддержки решений для администратора системы или планировщика задач. Он использует комбинацию расчетных и статистических показателей, что позволяет адаптироваться к реальным условиям эксплуатации. Например, в случае обработки медиа-контента или потоковых данных, где нагрузка колеблется в течение суток, система способна заранее перераспределить ресурсы, снижая задержки и предотвращая простои.

2.1. Математическая модель

Будем считать, что выполнение задачи характеризуется определенным параметром, например, время обработки, объем выделяемой оперативной памяти или дискового пространства, пропускная способность коммуникационных ресурсов. Прогнозируемое значение параметра предлагается определять в зависимости от расчетного значения и реальных (статистических) значений обработки аналогичного параметра и выражается следующей формулой:

$$T_{\text{п}} = f(T_{\text{р}}, T_{\text{с}}),$$

где Tп – прогнозируемое значение; Tр – ряд расчетных значений; Tс – ряд статистических значений.

Чтобы справедливо оценивать ряд статистических данных предлагается ввести для каждого значения весовой коэффициент, с помощью которого можно было бы задавать поведение модели [13]. Таким образом, модель является адаптивной: например, сделать самые свежие данные более значимыми, а устаревшие – менее значимыми, или при недостатке статистики (например, в начале работы системы) больший вес имеет расчетное значение; по мере накопления статистики прогноз становится все более обученным и учитывает реальные условия работы.

Алгоритм метода включает последовательность шагов:

1) Задача поступает в систему, определяются ее ключевые характеристики.
2) Для задачи вычисляется аналитическая оценка требуемого ресурса Tр.
3) На основе базы накопленной статистики рассчитывается значение Tс.
4) Итоговое прогнозируемое значение Tп определяется взвешенной комбинацией.
5) Задача распределяется на узел с минимальной ожидаемой загрузкой.
6) После выполнения обновляется статистика, что повышает качество последующих прогнозов.

2.2. Имитационная модель

Для оценки эффективности предиктивного метода разработана имитационная модель вычислительной системы. Модель позволяет воспроизводить работу набора узлов различной производительности при обработке входящего потока задач.

Основные параметры модели:

- количество вычислительных узлов: 5, 10 и 15;
- производительность узлов варьируется (от низкой до высокой, с фиксированными коэффициентами быстродействия);
- задачи подразделяются на три класса сложности (простые, нормальные и сложные), различающиеся по требуемому времени обработки;
- число поступающих задач изменяется в диапазоне от 100 до 2000;
- структура входного потока моделируется за счет изменения доли задач каждого класса.

Проводилось сравнение трех методов распределения ресурсов:

1) Статический метод Round Robin – задачи назначаются поочередно на каждый узел без учета его загруженности[14].
2) Динамический метод Least Loaded – задача назначается на узел с минимальной текущей загрузкой[15,16].
3) Предиктивный метод – задача направляется на узел с учетом прогнозируемого времени ее выполнения.

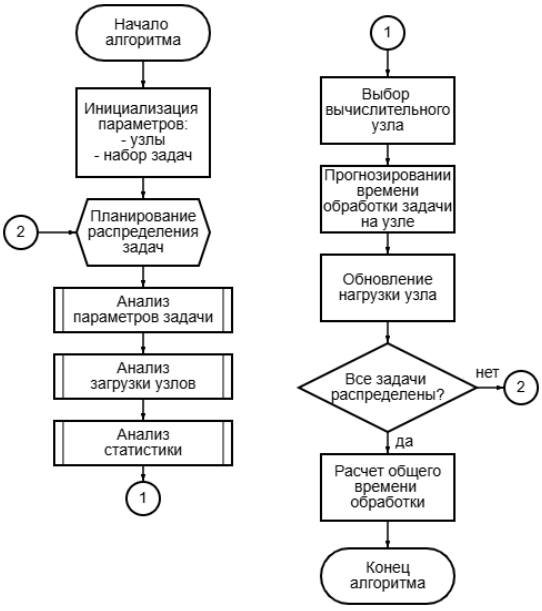


Рис. 1. Схема алгоритма модели

Модель реализована с возможностью изменения ключевых параметров, что позволяет проводить серию сценарных экспериментов для оценки эффективности каждого метода.

3. Сравнительный анализ

Результаты этого сценария позволят определить время обработки набора для каждого метода в зависимости от увеличения объема задач (рис. 2), а также максимальную и минимальную загруженность узлов системы при обработке набора задач.

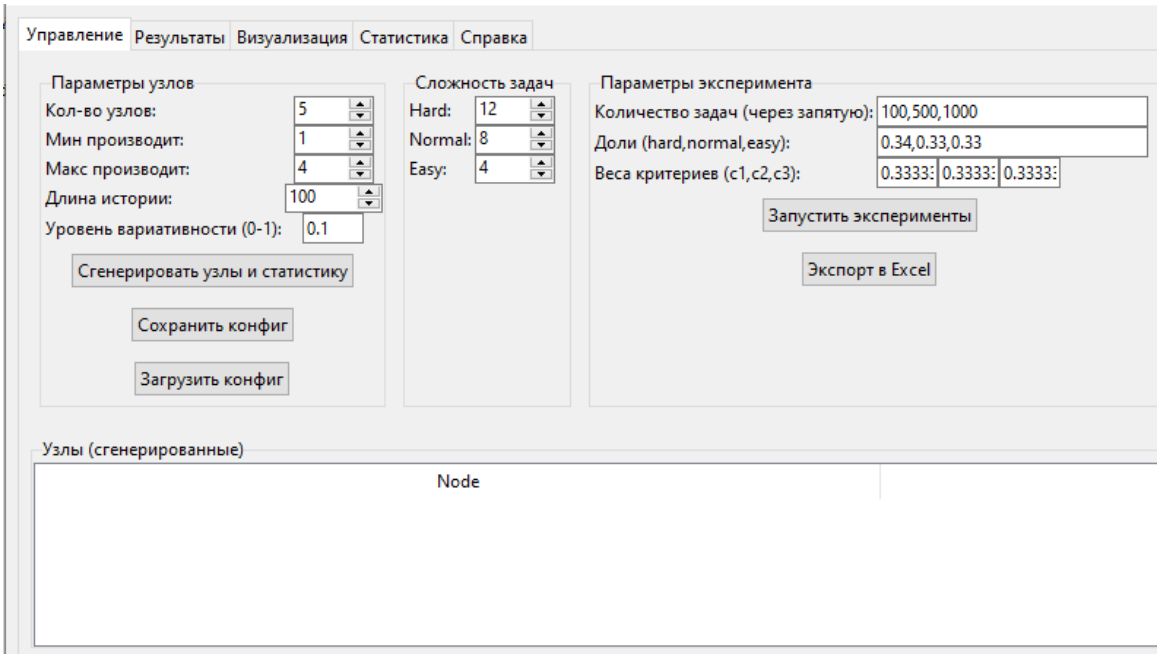


Рис. 2. Окно программной модели.

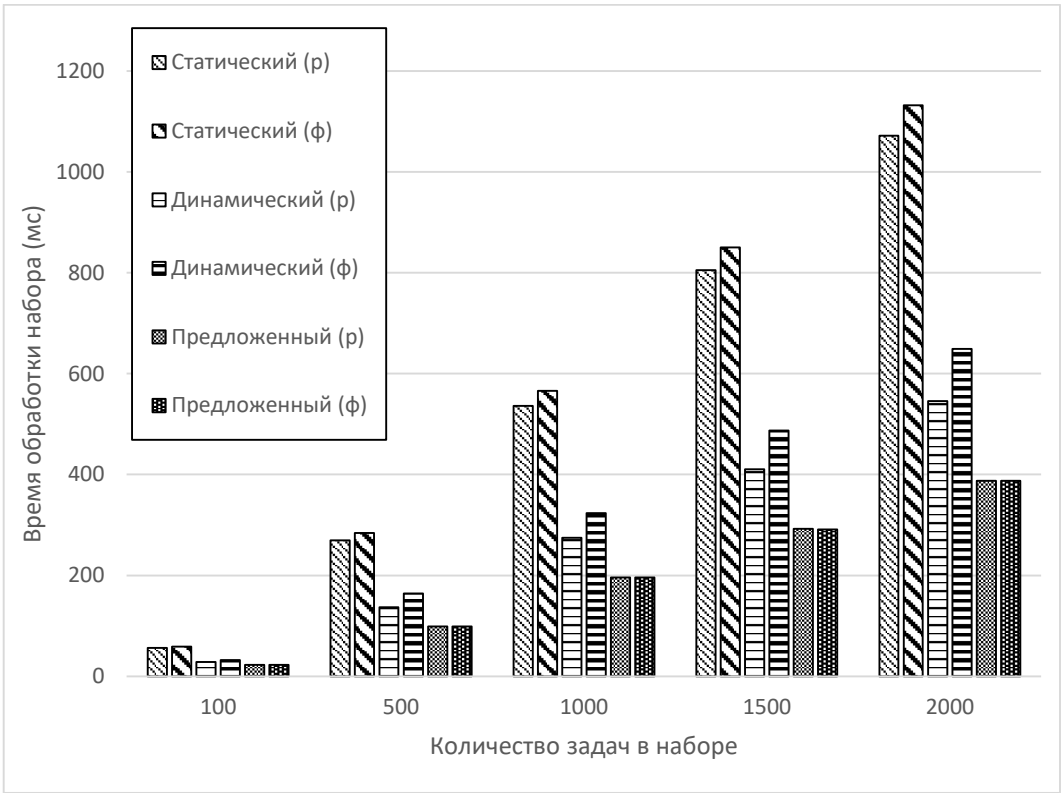


Рис. 3. Сравнение времени обработки набора задач, эксперимент для 10 узлов (меньше – лучше)

На рис. 3 показана гистограмма с группировкой по количеству задач в наборе для результатов моделирования различных методов: статический, динамический, предиктивный (предложенный). Кроме расчетных значений, определялись и фактические значения для времени обработки задач. Можно обратить внимание на разницу вершин расчетных и фактических значений. Она позволяет оценить, насколько точно методы выполняют расчеты в реальных системах.

Чтобы оценить темпы роста времени обработки набора задач при увеличении количества задач был построен график зависимости времени обработки задач от количества задач в наборе для двух методов: динамического и предиктивного (рис. 4). По рисунку видно, что темпы роста времени нового метода при увеличении количества задач меньше, чем у существующего.

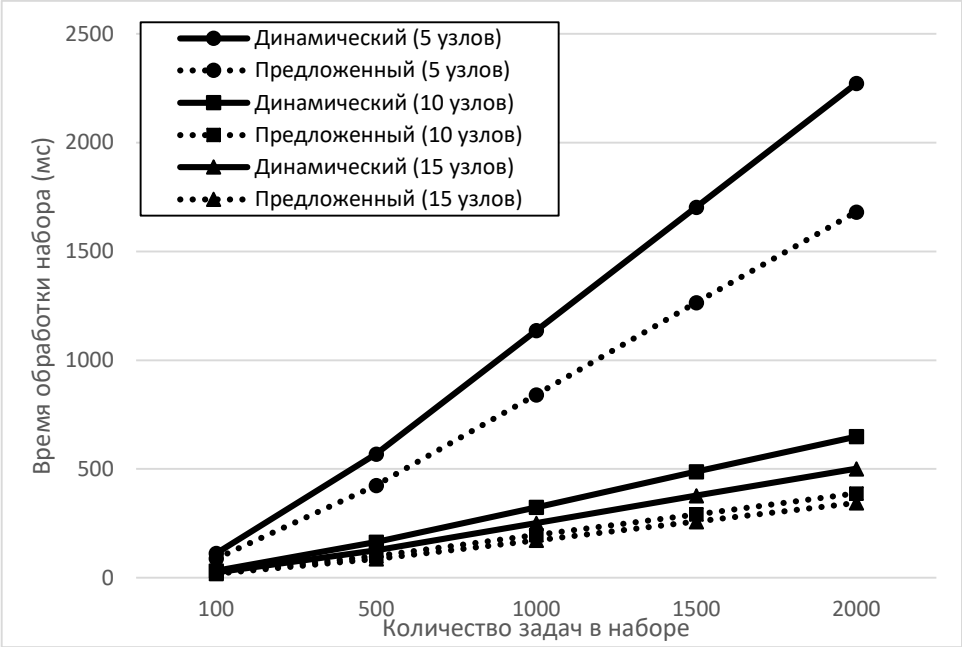


Рис. 4. График зависимости времени обработки задач от количества задач

На рис. 5 можно увидеть на сколько процентов отличается время работы самого загруженного узла и самого недогруженного узла системы при обработке задач. Этот параметр показывает на сколько эффективно метод задействует вычислительные узлы системы при выполнении распределения. Таким образом, можно сделать вывод, что предиктивный метод более оптимально использует ресурсы системы.

Далее для того, чтобы сравнить методы при динамически меняющейся нагрузке, рассмотрим результаты моделирования для разной пропорции задач по сложности. На рис. 6 представлены графики, которые показывают зависимость отличия времени обработки набора задач (между динамическим и предиктивным методами) от количества задач в наборе, при фиксированной пропорции (в процентах) групп задач по сложности (сложные/нормальные/легкие).

Были промоделированы следующие конфигурации:

1) Пропорция: 80% – сложных, 10% – нормальных, 10% – легких;
2) Пропорция: 10% – сложных, 80% – нормальных, 10% – легких;
3) Пропорция: 10% – сложных, 10% – нормальных, 80% – легких.

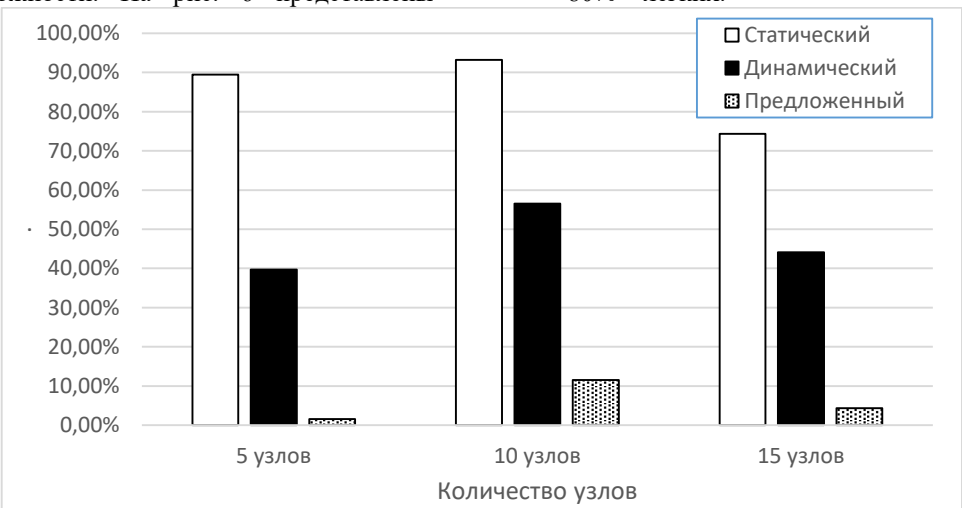


Рис. 5. Гистограмма среднего отличия времени работы самого загруженного узла и самого недогруженного от количества задач в наборе (меньше – лучше)

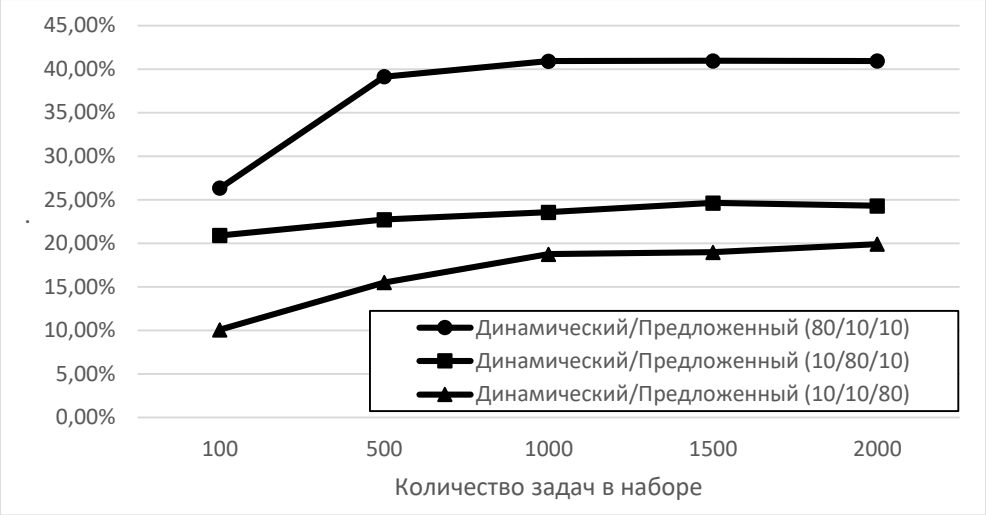


Рис. 6. График процентного отличия между динамическим и предиктивным методами при разных пропорциях

По рис. 6 видно, что эффективность предложенного предиктивного метода, относительно динамического, повышается в ситуации, когда средняя сложность всего набора увеличивается, т.е. преобладают более сложные задачи.

Для практического применения полученные результаты означают, что администраторы вычислительных систем могут использовать предиктивный метод как более эффективный по времени обработки. Даже в условиях отсутствия значительных вычислительных ресурсов метод способен повысить общую производительность без необходимости внедрения тяжелых интеллектуальных алгоритмов (например, нейросетей или эволюционных моделей).

4. Заключение

1) Предложен предиктивный метод принятия решений при распределения ресурсов в вычислительных системах. Его особенность заключается в комбинировании аналитических расчетов с данными о ранее выполненных задачах, что позволяет учитывать повторяемость и закономерности в поступающем потоке.
2) Сформирована математическая модель метода. Она описывает механизм вычисления прогнозируемого времени выполнения задач на основе взвешенной комбинации расчетного и статистического значений, а также определяет алгоритм обновления статистики при накоплении новых данных.
3) Реализована имитационная модель вычислительной системы. В ней варьировались количество узлов, их производительность, структура и объем потока задач. Это позволило воспроизвести различные сценарии функционирования системы и оценить поведение метода в разных условиях.
4) Проведено сравнение с классическими подходами. В качестве базовых алгоритмов использовались Round Robin (равномерное статическое распределение) и Least Loaded (динамический выбор наименее загруженного узла).
5) Экспериментально подтверждено сокращение времени выполнения набора задач. Предиктивный метод показал меньший рост времени обработки при увеличении числа задач по сравнению с динамическим методом.
6) Оценена равномерность загрузки узлов. В отличие от статического и динамического распределения, предлагаемый метод формирует более сбалансированное использование ресурсов, предотвращая перегрузку отдельных узлов.
7) Проведен анализ влияния структуры задач. При увеличении доли сложных задач эффективность метода возрастает, так как прогнозирование снижает вероятность перегрузки узлов высокими нагрузками.
8) Выявлены ограничения подхода. Для качественного прогнозирования требуется накопление репрезентативной статистики, а также необходимо учитывать дополнительные вычислительные затраты на расчет прогноза.

Список используемых источников

1. Trushkin E. S., Gavrilov A. V., Freyman V. I. Mathematical and Simulation Modeling for Assessing the Performance of Communication Devices of Computing, Information Control and Telecommunication Systems. Electrical engineering, information technology, control systems. – 2025. – № 53. – P. 129-156. – DOI 10.15593/2224-9397/2025.1.07

2. Yuqing Cheng, Zhiying Cao, Xiuguo Zhang, Qilei Cao, Dezhen Zhang. Multi objective dynamic task scheduling optimization algorithm based on deep reinforcement learning // The Journal of Supercomputing (2023) https://doi.org/10.1007/s11227-023-05714-1

3. Arun K., Nishant G., Aziz N. Bi-LSTM Based Deep Learning Algorithm for NOMA-MIMO Signal Detection System // Natl. Acad. Sci. Lett. https:// doi.org/10.1007/s40009-024-01516-y

4. Lyashov E.I. Resource-efficient algorithms for load balancing in distributed microservice architectures // Bulletin of Science No2 (83) vol. 1. Pp. 629 - 647. 2025. ISSN 2712-8849 // DOI 10.24412/2712-8849-2025-283-629-647

5. Cao, X., Chen, C., Li, S. et al. Research on computing task scheduling method for distributed heterogeneous parallel systems. Sci Rep 15, 8937 (2025). https://doi.org/10.1038/s41598-025-94068-0

6. Hwang K., Dongarra J. Distributed and Cloud Computing: From Parallel Processing to the Internet of Things. – Elsevier, 2012.

7. Bo Tian, M. A. Posypkin, I. Kh. Sigal, Load balancing based on estimates of the algorithmic complexity of subtasks // ITaCS, 2015, issue 1, 10–18

8. Kaurav, N.S., Yadav, P., 2019. A genetic algorithm-based load balancing approach for resource optimization for cloud computing environment. Int. J. Inf. Comput. Sci. 6 (3), 175–184. http://dx.doi.org/10.1016/j.protcy.2013.12.369

9. Moly M.I. Load Balancing Approach and Algorithm in Cloud Computing Environment // American Journal of Engineering Research (AJER), vol.8, no.04, 2019, pp.99-105

10. Saidi, K., Bardou, D. Task scheduling and VM placement to resource allocation in Cloud computing: challenges and opportunities // Cluster Comput 26, 3069–3087 (2023). https://doi.org/10.1007/s10586-023-04098-4

11. Kleiman L.A., Freyman V.I. The method of dynamic distribution of the diagnostic load between information and control systems elements // Proceedings of the 2021 IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering, ElConRus 2021. 2021. P. 952-955. https://doi.org/10.1109/ElConRus51938.2021.9396552

12. Tanenbaum E., Steen M. Distributed Systems: Principles and Paradigms. Moscow, Williams Publ., 2019.

13. Baranova T.P., Bugerya A.B., Efimkin K.N. Method of balancing the computational load for hybrid computing systems. 2021. V. 24. No. 1. https://doi.org/10.26907/1562-5419-2021-24-1-42-56

14. Li, J. Distributed data processing and task scheduling based on GPU parallel computing. Neural Comput & Applic 37, 1757–1769 (2025). https://doi.org/10.1007/s00521-024-10489-4

15. Chandra W., Rita W., Chin-Yin H., Jodi T., Chao-Tung Y. Load Balancing Algorithm in a Software-Defined Network Environment with Round Robin and Least Connections // Smart Grid and Internet of Things // 7th EAI International Conference, SGIoT 2023, TaiChung, Taiwan, November 18-19, 2023, Proceedings

16. Hulin Jin, Yong-Guk Kim, Zhiran Jin, Chunyang Fan, Yonglong Xu. Joint Task Ofoading Based on Distributed Deep Reinforcement Learning-Based Genetic Optimization Algorithm for Internet of Vehicles // J Grid Computing (2024) https:// doi.org/10.1007/s10723-024-09741-x

Интеллектуальный модуль для детекции сельскохозяйственных объектов региона, заросших хвойной растительностью

Максимова В. Н.
каф. «Информационные системы и технологии»
Южно-Уральский государственный университет
Челябинск, Россия
e-mail: maksimovavn@susu.ru

Макаровских Т. А.
Центр образовательных программ топ-уровня в сфере искусственного интеллекта «ВиртУм»
Южно-Уральский государственный университет
Челябинск, Россия
e-mail: makarovskikh.t.a@susu.ru

Жулев А. Э.
каф. «Системное программирование»
Южно-Уральский государственный университет
Челябинск, Россия
e-mail: zhulevae@susu.ru

А. С. Макаровских
каф. «Электронные вычислительные машины»
Южно-Уральский государственный университет
Челябинск, Россия
e-mail: maksvel@yandex.ru

Аннотация[1]

В настоящее время практически все отрасли подвержены процессу цифровой трансформации. В области сельского хозяйства и природопользования данный процесс идет достаточно медленно ввиду ряда объективных причин. Тем не мене, в отрасли назрело большое количество задач, которые невозможно решить без использования современных технологий. Одной такой задачей является детекция проблемных объектов землепользования. **Целью** работы является разработка подхода для детекции неиспользуемых земель сельскохозяйственного назначения, заросших хвойной растительностью. **Методы:** Разработанное программное обеспечение использует данные спутниковых снимков, информацию о границах объектов из базы данных системы УралГИС-Агро и данные о динамике вегетационного индекса для принятия решений. Предложенный подход привлекателен тем, что он не требует использования громоздких нейросетевых подходов, в его основе лежат статистические методы анализа данных, что позволяет получать решение приемлемой точности за сравнительно небольшой промежуток времени. Полученные **результаты** (перечень проблемных объектов, заросших хвоей) являются рекомендательными и требуют экспертной оценки при дальнейшем принятии решения. **Заключение:** Разработанный подход можно обобщить для решения смежных задач при определении инвестиционной привлекательности земель региона. Модуль встроен в систему УралГИС-Агро и позволяет ежегодно получать перечень проблемных объектов землепользования.

Ключевые слова: интеллектуальная система, комплекс программ, точное земледелие, принятие решений, мониторинг.

Keywords: intellectual system, software, precision farming, decision making, monitoring.

1. Введение

В настоящее время практически все отрасли подвержены процессу цифровой трансформации, которая стала уже не модной тенденцией, а необходимостью. Разработка интеллектуальных

[1] Труды XI Международной научной конференции "Информационные технологии интеллектуальной поддержки принятия решений", 11-13 ноября, Уфа-Чандигарх-Фандербейлпарк, 2025

методов управления и принятия решений для задач различных областей человеческой деятельности является актуальной задачей. В области сельского хозяйства и природопользования данный процесс идет достаточно медленно ввиду ряда объективных причин, начиная с проблем целеполагания и отсутствия единого стратегического плана, заканчивая сильнейшим сопротивлением изменениям со стороны специалистов, работающих в отрасли [1]. Тем не мене, в отрасли назрело большое количество задач, которые невозможно решить без использования современных технологий. Одной такой задачей является детекция проблемных объектов землепользования.

В [2] отмечено, что одной из значимых экономических и экологических проблем является выявление зарастания сельскохозяйственных земель. Земли, относящиеся к категории сельскохозяйственных, должны использоваться по назначению и по закону они не могут долго простаивать и не возделываться совсем: в таком случае, их изымают у собственника. Особый интерес в данном случае представляют объекты, заросшие хвойной растительностью, поскольку данный ресурс является наиболее ценным.

Проблема зарастания связана с тем, что учет неиспользуемых сельскохозяйственных земель не ведется более 30 лет. За это время поля, устроенные в лесной зоне, покрылись древесной растительностью. Особенно остро такая проблема стоит для регионов Урала и Сибири. Поскольку за это время на заброшенных объектах землепользования успел вырасти полноценный молодой лес, это значительно осложняет поиск такого объекта. А возможности его дальнейшего возделывания весьма сомнительны [3]. Так, в [3] предлагается подход, протестированный в Владимирской области для детекции залесенных земель региона, основанный на нейросетевых алгоритмах автоматического дешифрирования, и использующий открытые данные Росреестра. [3] является одним из первых материалов, где отмечено, что совместное использование космической съемки и кадастровых данных позволяет провести оперативную инвентаризацию сельскохозяйственных земель и выявить нарушения землепользования на больших площадях. Тем не менее, в статье не приводится рекомендаций по обобщению полученных результатов и возможности использования разработанного подхода для других регионов.

Рассматривая в качестве объекта исследования Челябинскую область, можно отметить, что на территории региона в 22 сельскохозяйственных муниципальных образованиях на учете стоит более миллиона угодий, информация о которых, как и во всей стране, не обновлялась последние 20-30 лет. В 2021 году Рослесинфорг начал работы по второму циклу госинвентаризации лесов. Это может позволить провести анализ зарастающих земельных участков для оценки потенциала ведения лесного хозяйства на таких участках [4, 5], тем не менее, информация о результатах данной работы после 2023 года так и не была опубликована. Согласно [5] Рослесинфорг оцифровывает данные о лесных насаждениях на зарастающих земельных участках с 2021 года, в ходе второго этапа государственной инвентаризации лесов (ГИЛ). За это время на территории 27 регионов (в соответствии с [5] – оцифровка начата на Дальнем Востоке и в Сибири) было выявлено 13,2 млн га пашен, которые заросли лесом. В этом году в 15 российских субъектах 85 полевых бригад Рослесинфорга обследуют 118,8 млн га, чтобы снова обнаружить древесно-кустарниковую растительность на сельскохозяйственных землях.

В научных работах данная проблема была обозначена и того раньше – в 2006 году [6], однако вопрос рассматривался не с точки зрения интеллектуальных методов детекции проблемных участков, а как статистическое исследование динамики зарастания. Большинство работ зарубежных авторов посвящено не столько проблеме зарастания, сколько проблеме определения выгоревших участков леса [7-9].

Поэтому разработка и внедрение в современные ГИС-системы интеллектуального модуля детекции проблемных объектов землепользования и выработка обобщенного подхода для решения смежных задач, является актуальной. К предпосылкам создания технологии автоматизированного определения оптимального использования проблемных объектов относятся:

- Реализация с 2018 г. по настоящее время программы «Цифровизация земель сельскохозяйственного назначения для рационального использования».
- Выполнение пилотных проектов по наполнению геоданными федеральных систем.
- Реализация научно-исследовательских работ для Министерства экологии и природопользования РФ и Челябинской области по обследованию водных объектов.
- Разработка интерактивной геоинформационной системы для Челябинской области.

С 2015 года в Южно-Уральском государственном университете создана и успешно развивается геопортальная система «УралГИС-Агро». Целью создания этой системы является информационно-аналитическое обеспечение рационального использования земель сельскохозяйственного назначения и оперативная межведомственная интеграция по вопросам, связанным с землями сельскохозяйственного назначения.

На рис.1 приведены источники данных для системы «УралГИС-Агро».

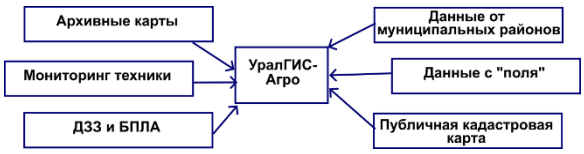


Рис. 1. Данные, используемые системой «УралГИС-Агро»

Система имеет модульный интерфейс (рис. 2) и позволяет подключать дополнительные интеллектуальные модули для решения ряда практических задач.



Рис. 2. Модули системы «УралГИС-Агро»

Рассмотрим более подробно модуль «Детекция проблемных объектов землепользования». В нашей предыдущей работе [2] рассмотрен подход, в котором авторами предлагается метод для детекции полей, заросших лесом. В этой работе лес не подразделяется на хвойный и лиственный и для детекции используются достаточно грубые оценки пороговых значений. В данной работе авторы предлагают различать хвойные и лиственные леса, разработан метод определения пороговых значений вегетационного индекса для объекта землепользования. Предложенные методы используют статистические подходы, что позволяет получить решение за сравнительно небольшое время для объектов всего региона. Отказ от использования нейросетевых методов обусловлен наличием в базе данных геопортальной системы «УралГИС-Агро» данных о контурах объектов интереса и отсутствием необходимости уточнения этих границ, поскольку целью работы интеллектуального модуля является фильтрация данных для проведения дальнейшей экспертной оценки.

2. Общая схема системы детекции проблемных объектов

Алгоритм детекции проблемного объекта землепользования в общем случае может быть сформулирован следующим образом (см. рис.3) [10].

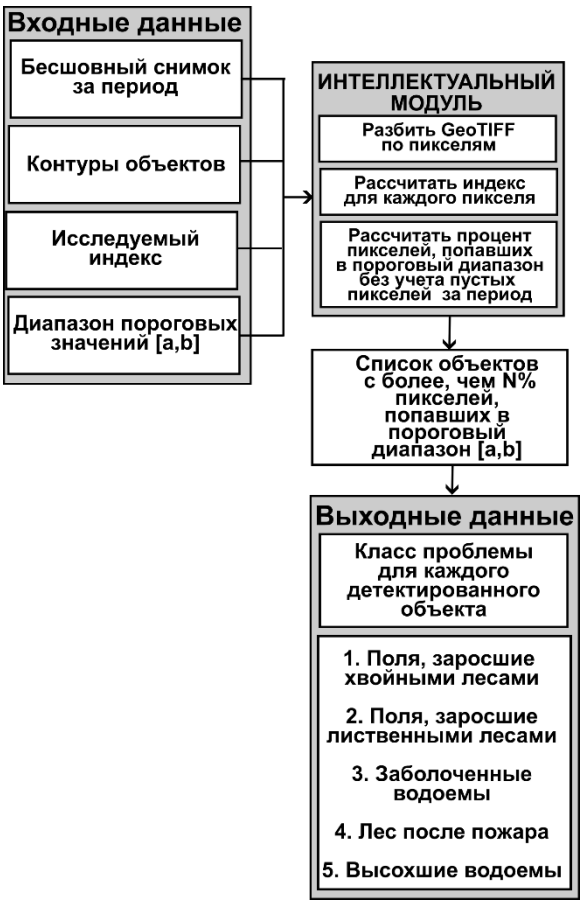


Рис. 3. Общая схема алгоритма детекции проблемных объектов

Алгоритм использует на входе бесшовные спутниковые снимки за анализируемый период. С помощью контуров объектов (векторные данные о границах объектов, сформированные из картографических материалов ГФД), из этих снимков может быть получена информация в формате GeoTiff, соответствующая объектам интереса. В зависимости от типа выявляемого дефекта определяется исследуемый индекс (вегетационный, водный или др.) указываются пороговые значения этого индекса, соответствующие выявляемой проблеме. Например, если требуется определить заросшие лесом пашни, то в качестве индекса используется NDVI (нормализованный относительный индекс растительности), а в качестве диапазона пороговых значений берутся все значения выше 0.4, что соответствует хорошо развитой растительности.

Независимо от класса проблемы, схема его детектирования остается неизменной: требуется попиксельное вычисление индекса для каждого объекта, затем проводится сравнение рассчитанных значений с пороговыми и вычисляется процент пикселей объекта, попавших в пороговый диапазон. В зависимости от полученного значения принимается решение о наличии детектируемой проблемы для рассматриваемого объекта. Например, согласно официальной документации

[11] залесенным считается объект, на котором кустарники и деревья занимают не менее 75 % площади (то есть, количество пикселей, имеющих повышенный вегетационный индекс в период с мая по сентябрь, превышает 75%). В [5] под новоявленным лесом понимается участок площадью больше 0,5 га, на котором кустарники и деревья высотой от 5 метров занимают не менее 75 % площади. Разработанный модуль детекции проблемных объектов не позволяет оценить высоту деревьев. Для этого предлагается воспользоваться либо экспертной оценкой на местах, либо разработать модуль на основе методов искусственного интеллекта для проведения более детального исследования выявленного проблемного участка. Для определения высоты деревьев также могут быть использованы результаты космической стереосъемки объектов под разными углами наклона целевой аппаратуры. Поскольку использование подобных методов требует больших вычислительных ресурсов, предварительный отбор проблемных участков необходим.

3. Определение величин для диапазона пороговых значений при детекции зарастания хвойными лесами

Рассмотрим задачу определения заброшенных объектов землепользования, заросших хвойными лесами. Для решения каждой задачи детекции проблемных участков отметим, что каждый объект в природе, будь то лес, пашня, озеро, река, в течение года проходят цикл развития. В зависимости от времени года, температуры воздуха, влажности и многих других параметров меняются и значения вегетационных индексов. Например, осенью вегетационный индекс для лиственных лесов понижается, поскольку в таких лесах отсутствует зеленая масса. И ее отсутствие в средней полосе России, на Урале, в Сибири приходится примерно на ноябрь – апрель. С другой стороны, для вечнозеленых хвойных лесов, представляющих особую ценность для бизнеса, вегетационный индекс остается повышенным в течение всего календарного года. Это наблюдение позволяет сделать вывод, что каждый интересующий нас тип объектов имеет некоторый паттерн изменения значений различных индексов в течение календарного года. И для детекции соответствующего типа на территории, занятой, например, пашнями, необходимо сопоставить эти паттерны. Под паттерном в данном случае будем понимать устойчивое повторяющееся сочетание индексов, которое формирует узнаваемый график на протяжении календарного года.

Для определения пороговых значений и дальнейшего формирования паттерна для хвойных лесов был разработан модуль расчета медианных значений индексов в течение календарного года. Программный модуль, который получает на вход файл *.csv, содержащий информацию о среднем значении вегетационного индекса для каждого лесного объекта в заданном районе. На выходе модуль позволяет получить векторы медианных, средних величин, дисперсии для дальнейшего их использования при анализе информации, классификации лесных объектов на хвойные и лиственные. Модуль позволяет вычислить медианное, среднее значения и дисперсию для имеющихся данных. Разработка программного модуля для анализа вегетационного индекса (NDVI) значительно повысит эффективность обработки данных за счет автоматизации расчетов описательной статистики и фильтрации объектов.

Для определения динамики NDVI в муниципальном районе используются именно медианные значения, поскольку распределение значений индекса по объектам в каждом месяце не подчиняется нормальному закону. Для определения значения индекса отдельного объекта используется среднее значение, поскольку попиксельные данные для отдельно взятых объектов распределены без значительных выбросов и удовлетворяют нормальному закону распределения.

Например, при построении графика распределения значений индекса NDVI в зимние месяцы была отмечена следующая закономерность (рис.4).

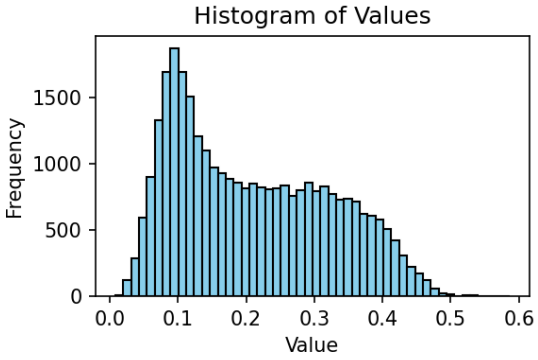


Рис.4. Распределение частот NDVI для лесных объектов по данным за 10 февраля 2025 г.

По графику видно, что имеется два пиковых значения частот, один – в промежутке [0; 0.2], второй – в промежутке [0.2;1]. Первое значение соответствует лиственной растительности (когда зеленая масса практически отсутствует), второе – хвойной, когда зеленая масса сохранена на протяжении календарного года. В весенние и летние месяцы, а также ранней осенью график распределения имеет более пологий первый пик, смещенный вправо (рис. 5).

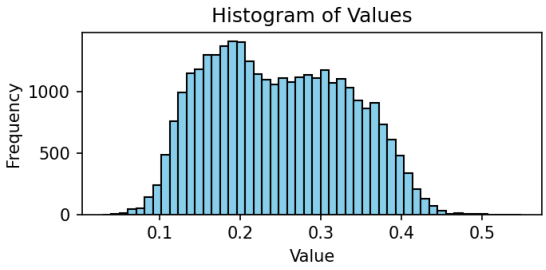


Рис.5. Распределение частот NDVI для лесных объектов по данным за 18 октября 2024 г.

Руководствуясь данной информацией, можно получить медианные значения для NDVI хвойных лесов за календарный год и построить паттерн динамики индекса в течение 12 месяцев (рис.6). К хвойным лесам будем относить объекты, для которых средний NDVI в каждом месяце отличается не более, чем на 5% от найденного медианного значения.

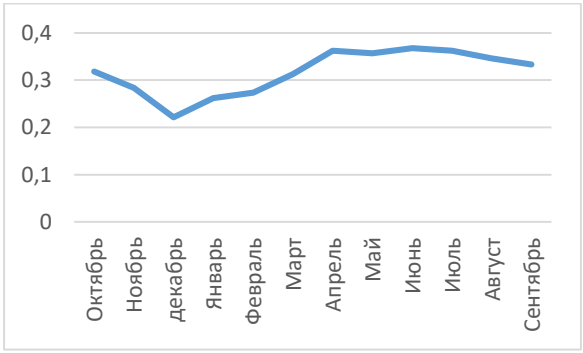


Рис.6. Паттерн динамики NDVI для хвойных лесов Челябинской области

Данный паттерн можно получить как для каждого муниципального района в отдельности, так и для региона в общем. Паттерны для районов являются более точными, поскольку один регион может располагаться в различных зонах (лес, лесостепь, степь).

В дальнейшем, используя данный паттерн для определения пороговых значений NDVI, можно выполнить детекцию хвойных лесов для невостребованных сельскохозяйственных земель региона. В соответствии с [11] заросших хвойной растительностью будем считать объекты, для которых доля пикселей со значением NDVI превышающем значения паттерна в каждом месяце, больше 75%.

Для детекции заросших хвойной растительностью невостребованных объектов землепользования воспользуемся попиксельным анализом, как и для детекции заросших объектов независимо от типа растительности. В отличие от предыдущей задачи [2], нет необходимости анализировать пять вегетационных периодов подряд. В данном случае определяется наличие зеленой массы в зимние месяцы, что соответствует наличию хвойной растительности на исследуемом участке. Поэтому объем вычислений будет значительно ниже. Для анализа зарастания в данном случае берутся данные за одну зиму. Выполняется проверка в течение 6 месяцев, чтобы снизить погрешность детекции, связанную с дефектами на снимках и отсутствием данных в связи с повышенной облачностью.

В результате запуска модуля детекции заросших хвойной растительностью невостребованных объектов землепользоваания для муниципальных районов Челябинской области были получены следующие результаты (таблица 1).

Таблица 1. – Данные по пяти наиболее заросшим хвойной растительностью муниципальным районам

№	S (га)	Заросших хвоей (га)	% заросших хвоей	Прибыль от передачи в лесфонд (млн.руб.)
1	69185,70	3852,11	6	3852,00
2	97905,53	4897,34	5	4897,00
3	96817,46	4590,88	5	4590,00
4	62801,70	1593,97	3	1593,00
5	12963,83	354,19	3	354,00
	369977,07	15288,49	4	15286,00

В таблице приводятся данные по пяти районам, в которых была детектирована наибольшая площадь зарастания хвойной растительностью. В таблице также приводится возможная прибыль для региона при передаче детектированных участков в лесной фонд. Результаты, представленные в таблице, позволяют оценить масштабность проблемы. За те годы, когда отсутствовал какой-либо мониторинг и какая-либо инвентаризация земель, порядка 10% сельскохозяйственных угодий заросло хвойной растительностью.

Отметим, что проведенный анализ является первичным. Он позволяет отфильтровать большой объем данных, не представляющий интереса при экспертной оценке. Результаты проведенной детекции в дальнейшем можно передать экспертам для проведения полевых экспериментов и осмотра детектированных объектов землепользования. Кроме того, целесообразно разработать интеллектуальный модуль на основе нейросетевой модели анализа изображения более высокого пространственного разрешения. Необходимость разработки модуля для грубой оценки объектов землепользования обусловлен большим объемом анализируемой информации и его целью является фильтрация объектов, представляющих интерес для дальнейшего анализа. Экспертный анализ, полевые эксперименты, более высокодетальные снимки

являются достаточно дорогими инструментами, и поэтому применяются точечно для объектов, представляющих интерес.

4. Приложение-анализатор для детекции проблемных объектов землепользования

Как было показано в статье [2], подход для первичного интеллектуального анализа большого объема геопространственной информации может быть масштабирован на решение большого спектра задач детекции проблемных объектов землепользования.

Приведем основные характеристики и опишем принципы работы приложения-анализатора для фильтрации объектов землепользования по заданной характеристике.

1. Приложение-анализатор представляет собой разработку на языке Python с оконным интерфейсом, которая позволяет пользователю указать местоположение (папку, возможно, с информацией в нескольких вложенных папках) анализируемых файлов и критерии оценивания и при нажатии на кнопку «Рассчитать» получить файл с расширением *.csv, в котором содержатся ID отфильтрованных объектов. Приложение используется как подключаемый модуль к системе «УралГИС-Агро» для построения перечней проблемных объектов землепользования, удовлетворяющих определенным критериям.

2. Входные данные для анализа:

- Директория, в которой хранятся спутниковые снимки муниципального района (файлы tiff) за интересующий период.

- Директория, в которой хранятся контуры объектов (файлы shp), используемых в качестве эталонных для получения медианных пороговых значений и их дисперсий.

- Директория, в которой хранятся контуры исследуемых объектов (файлы shp). Это объекты, среди которых требуется найти проблемные.

- Анализируемый индекс (перечень индексов, которые можно рассчитать по загруженным спутниковым снимкам – NDVI, NDWI и др.). По умолчанию установлен индекс NDVI.

- Процент содержания дефекта D (целое число, по умолчанию 75, в соответствии с нормативными документами [11]).

- Описание дефекта (строковое поле, где пользователь вводит, что именно он анализирует, например, «Зарастание хвоей»). Значение этого поля будет выведено как название сформированного слоя в «УралГИС-Агро».

3. Результат: список исследуемых объектов, для которых процент содержания дефекта не ниже D.

4. Алгоритм функционирования модуля:

- Ввести все исходные данные, проверить, чтобы в выбранных папках присутствовал хотя бы один файл указанного типа.

- При нажатии на кнопку «Рассчитать» для эталонных объектов выполнить обрезку имеющихся снимков по контуру, рассчитать среднее значение выбранного пользователем индекса для каждого объекта и сохранить в файл CSV.

- Запустить модуль получения средних значений для указанного пользователем индекса, выбрать диапазон изменения значений для объектов интереса и сформировать векторы медианных значений и их дисперсий индекса для всех снимков. Таким образом, будет получен паттерн медианных значений.

- Обрезать снимки по данным о контурах исследуемых объектов интереса (объекты, для которых определяется наличие проблемы). Рассчитать значение выбранного индекса для каждого пикселя каждого объекта и сравнить полученное значение с пороговым. Если значение индекса для пикселя выше порогового за вычетом возможного отклонения, признать пиксель проблемным. Подсчитать количество проблемных пикселей для каждого объекта. Если количество проблемных пикселей выше D, то признать объект проблемным по заданному критерию и сохранить его ID в список проблемных.

- Список проблемных записать в отдельный CSV-файл.

Полученный выходной файл можно использовать для формирования в «УралГИС-Агро» слоя, на котором отображены все проблемные объекты.

5. Заключение

Таким образом, наличие актуальной информации о сельскохозяйственных угодьях, заросших хвойной растительностью, будет определять дальнейшее развитие интеллектуальной информационной системы "УралГИС-Агро", в частности, в части расширения информационно-аналитического модуля "Рациональное использование заросших сельскохозяйственных угодий". Суть модуля основана на анализе на уровне критериев и интеграции данных о заросших сельскохозяйственных угодьях с информацией о климате, почве, населении, рельефе, гидрологии и многом другом. Результатом модуля "Рациональное использование заросших сельскохозяйственных угодий" станет автоматизированная оценка пригодности заросшей сельскохозяйственной территории с учетом ее потенциала, в частности: перевод земли в лесной фонд, использование ее под промышленные сады или изъятие из сельскохозяйственного использования для альтернативных видов деятельности, например, туризм и т.д.

Обнаружение сельскохозяйственных угодий, заросших хвойной растительностью, знаменует собой начало планомерной работы, направленной на восстановление заброшенных сельскохозяйственных угодий для дальнейшего надлежащего использования с учетом региональных особенностей и их природного ландшафтного потенциала. Сегодняшние результаты этого анализа могут помочь региону сэкономить порядка 15 288 миллионов рублей на переводе заброшенных земель в лесной фонд (таблица 1). Такой мониторинг необходим ежегодно, поскольку зарастание – это непрерывный процесс.

Помимо использования разработанного модуля для отображения проблемных участков землепользования, отфильтрованных по выбранному критерию, полученную информацию об объектах можно использовать в качестве входных данных для эконометрической модели определения экономической привлекательности невостребованных земель региона. Разработка и тестирование данной модели является направлением дальнейших исследований авторов. Еще одним направлением дальнейших разработок можно отметить создание нейросетевых моделей для проведения повторной оценки и фильтрации детектированных объектов с целью сокращения перечня обнаруженных объектов и, как следствие, уменьшение объема работы эксперта, проводящего оценку детектированных объектов и принятие решения о дальнейшем использовании этого объекта.

Список используемых источников

1. Логиновский О. В., Лясковская Е. А., Габдулин Р. Р. Проблемы цифровой трансформации субъектов Российской Федерации // Вестник ЮУрГУ. Серия: Компьютерные технологии, управление, радиоэлектроника. 2023. №3. С. 76-92.
2. Максимова В.Н., Макаровских Т.А., Жулев А.Э., Макаровских А.С. Интеллектуальная система мониторинга сельскохозяйственных земель региона (УралГИС-Агро) // В сборнике: Информационные технологии интеллектуальной поддержки принятия решений (памяти проф. Н.И. Юсуповой) ITIDS'2024. Труды X Международной научной конференции. В 2-х томах. Уфа, 2024. С. 134-140.
3. Лес наступает. Как зарастают сельскохозяйственные земли/ Геопространственная аналитика [Электронный ресурс] https://geonovosti.terratech.ru/nature/les-nastupaet/ (дата доступа 10.09.2025).
4. В 15 регионах России в 2023 году начнут поиск заросших лесом пашен [Электронный ресурс]// URL: https://tass.ru/ekonomika/16822533 (дата доступа 10.09.2025).
5. Рослесинфорг в поисках заросших лесом полей [Электронный ресурс] // https://dzen.ru/a/ZJKWAG1BgnT5of5K (дата доступа 10.09.2025).
6. Балашкевич Ю. А. Зарастание бывших сельскохозяйственных земель древесной растительностью // Актуальные проблемы лесного комплекса. 2006. №13. URL: https://cyberleninka.ru/article/n/zarastanie-byvshih-selskohozyaystvennyh-zemel-drevesnoy-rastitelnostyu (дата обращения: 10.09.2025).
7. Martinez J.L., Lucas-Borja, M.E., Plaza-Alvarez, P.A., Denisi, P., Moreno, M.A., Hernandez, D., Gonzalez-Romero J., Zema D.A. Comparison of Satellite and Drone-Based Images at Two Spatial Scales to Evaluate Vegetation Regeneration after Post-Fire Treatments in a Mediterranean Forest. Appl. Sci. No.11. ID 5423 (2021). https://doi.org/10.3390/app11125423
8. Andrada M.E., Russell D., Arevalo-Ramirez T., Kuang W., Kantor G., Yandun F. Mapping of Potential Fuel Regions Using Uncrewed Aerial Vehicles for Wildfire Prevention. Forests no. 14, ID 1601 (2023). https://doi.org/10.3390/ f14081601
9. Lu K., Xu R., Li J., Lv Y., Lin H., Liu Y. A Vision-Based Detection and Spatial Localization Scheme for Forest Fire Inspection from UAV. Forests no. 13, ID 383 (2022). https://doi.org/10.3390/f13030383
10. Максимова В.Н., Макаровских Т.А. технология определения оптимального использования проблемных объектов природопользования на основе управления геоинформацией в целях пространственного развития территории // Вестник ЮУрГУ. Сер.: Компьютерные технологии, управление, радиоэлектроника. 2025. Т. 25. № 3. С. 108-115.
11. Постановление Правительства РФ от 8 июня 2022 г. N 1043 «О внесении изменений в Положение об особенностях использования, охраны, защиты, воспроизводства лесов, расположенных на землях сельскохозяйственного назначения».

Анализ современных решений с использованием языковых моделей для поддержки работы проектного менеджера

Рафиков И. Р.
Институт информатики, математики и робототехники
Уфимский университет науки и технологий
Уфа, Россия
e-mail: rofrenius@gmail.com

Ризванов Д. А.
Институт информатики, математики и робототехники
Уфимский университет науки и технологий
Уфа, Россия
e-mail: ridmi@mail.ru

Разяпов Р. В.
Архитектурно-строительный институт (АСИ)
Уфимский государственный нефтяной технический университет
Уфа, Россия
e-mail: kpi1t@yandex.ru

Аннотация[1]

В статье рассматриваются современные подходы к применению языковых моделей (Large Language Models, LLM) для интеллектуальной поддержки работы проектного менеджера. Анализируются возможности использования архитектур типа Retrieval-Augmented Generation (RAG), сочетающих языковые модели с корпоративными базами знаний, а также перспективы интеграции с традиционными системами поддержки принятия решений (СППР). Отмечаются преимущества LLM при работе с неструктурированными данными проектной документации и коммуникаций. Обсуждаются ограничения, связанные с безопасностью, необходимостью адаптации к отраслевой специфике и интерпретируемостью решений. Делается вывод о перспективности гибридных систем, объединяющих LLM, онтологические модели и нечёткую логику для задач управления инженерными проектами.

Ключевые слова: Проектный менеджмент, языковые модели, интеллектуальная поддержка принятия решений, инженерные проекты, RAG, онтология, нечёткая логика.

1. Введение

Современный проектный менеджмент характеризуется высокой сложностью, большим объёмом документации и необходимостью быстрой обработки информации для принятия решений. Особенно остро эта проблема проявляется в инженерных и промышленных проектах, где менеджер координирует работу различных подразделений, анализирует технические отчёты и должен обеспечивать соответствие нормативным требованиям. В таких условиях возрастает значение интеллектуальных инструментов, способных поддерживать принятие решений на основе анализа больших объёмов текстовой и табличной информации.

Традиционные системы поддержки принятия решений (СППР), описанные в работах Э. Турбана и соавторов, ориентированы преимущественно на обработку структурированных данных и формализованных правил [1]. В отечественных исследованиях предлагаются модели на основе нечёткой логики и онтологий, обеспечивающие формализацию экспертных знаний [2]. Однако эти подходы требуют ручного обновления базы знаний и не адаптируются к новым источникам данных без участия специалиста.

Появление больших языковых моделей (Large Language Models, LLM) открыло новые возможности для автоматизированного анализа и обработки естественного языка. Такие модели, как GPT-4 [OpenAI, 2023] и LLaMA-3 [Meta AI, 2024], способны обрабатывать неструктурированные тексты, извлекать знания и формировать рекомендации в интерактивной форме. Наиболее перспективным направлением их применения

[1] Труды XI Международной научной конференции "Информационные технологии интеллектуальной поддержки принятия решений", 11-13 ноября, Уфа-Чандигарх-Фандербейлпарк, 2025

является архитектура Retrieval-Augmented Generation (RAG), которая объединяет языковую модель и корпоративную базу знаний, обеспечивая доступ к актуальной информации при генерации ответов [3].

Цель настоящей работы – провести анализ существующих решений, основанных на языковых моделях, и определить их потенциал для интеллектуальной поддержки деятельности проектного менеджера в инженерных проектах.

2. Классические подходы к интеллектуальной поддержке принятия решений

Системы поддержки принятия решений (СППР) применяются в управлении проектами более трёх десятилетий и представляют собой комплекс программных средств, обеспечивающих анализ данных, моделирование и выбор наилучших альтернатив.

Классическая архитектура СППР [1] включает четыре взаимосвязанные подсистемы:

1. Подсистема управления данными (Data management) – обеспечивает доступ к внутренним и внешним источникам информации, таким как ERP-системы, веб-ресурсы и корпоративные базы данных.

2. Подсистема управления моделями (Model management) – содержит математические и имитационные модели, используемые для анализа и прогнозирования, а также обеспечивает взаимодействие с внешними моделями (External models), реализованными во внешних программных средствах.

3. Подсистема управления знаниями (Knowledge-based subsystem) – хранит экспертные правила и логические зависимости, формирующие основу рассуждений.

4. Подсистема пользовательского интерфейса (User interface) – обеспечивает взаимодействие пользователя (менеджера проекта) с системой.

Эта структура представлена на рисунке 1 и остаётся методологической основой современных интеллектуальных систем поддержки решений.

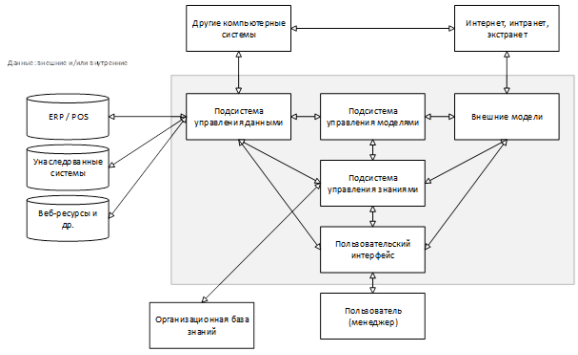


Рис. 1. Классическая структура системы поддержки принятия решений [1]

В инженерных и промышленных проектах такие СППР традиционно реализуются на основе экспертных систем, где знания формализуются в виде нечётких правил и онтологий предметной области. Применение нечёткой логики позволяет описывать слабо структурированные ситуации, типичные для управления проектами, где многие факторы имеют неопределённый характер (например, сроки поставки оборудования, надёжность подрядчиков, риски срыва графика) [2].

Онтологический подход обеспечивает представление знаний в виде взаимосвязанных понятий – процессов, ресурсов, рисков и действий. Пример фрагмента такой структуры приведён на рисунке 2.

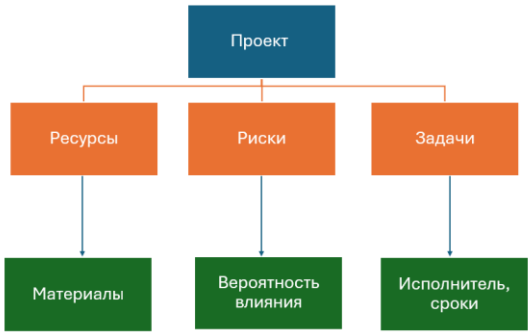


Рис. 2. Фрагмент онтологической модели предметной области проектного управления [2]

Такие системы эффективны для анализа заранее известных сценариев, но они не масштабируются при увеличении объёма данных и количества факторов. В большинстве случаев обновление правил и онтологий выполняется вручную экспертами, что ограничивает их применимость в динамичной среде современных проектов.

Как отмечают авторы [4], классические СППР обеспечивают прозрачность и объяснимость решений, однако требуют значительных ресурсов на сопровождение и не позволяют эффективно использовать неструктурированную информацию, содержащуюся в текстовых отчётах, переписке или технических спецификациях.

3. Эволюция систем поддержки решений и появление языковых моделей

Постоянный рост объёма проектной информации и развитие технологий обработки естественного языка (NLP) привели к эволюции систем поддержки решений в сторону более гибких и самообучающихся архитектур. Если традиционные СППР опираются на жёстко формализованные правила и экспертные базы знаний, то современные подходы стремятся использовать модели,

способные понимать контекст и смысл текстовых данных без предварительного кодирования правил.

Прорыв в этом направлении связан с появлением трансформерных архитектур, которые легли в основу всех современных языковых моделей [5]. Трансформеры позволили обучать нейронные сети на триллионах слов и обрабатывать длинные контексты, что ранее было невозможно. На базе этой архитектуры были созданы модели GPT (OpenAI, 2023), LLaMA (Meta AI, 2024) и Phi-3 (Microsoft Research, 2024), способные не только генерировать текст, но и анализировать сложные запросы и выводить логические зависимости.

В отличие от экспертных систем, языковые модели не требуют явного программирования правил: знания формируются в процессе обучения на больших текстовых коллекциях. Это позволяет LLM адаптироваться к различным предметным областям, в том числе к управлению проектами, где большое значение имеют документы, отчёты и коммуникации между участниками проекта.

Одним из ключевых направлений развития стала архитектура Retrieval-Augmented Generation (RAG) [3]. Её суть заключается в объединении языковой модели с внешней базой знаний (Knowledge Base, KB). При поступлении запроса система сначала выполняет поиск по документам, а затем формирует ответ, используя найденную информацию. Такой подход повышает точность ответов и позволяет работать с актуальными корпоративными данными.

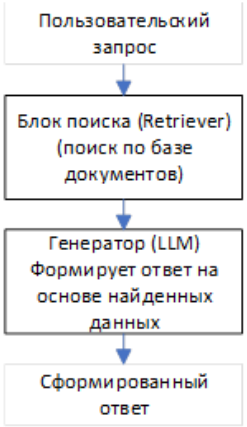


Рис. 3. Архитектура RAG (Retrieval-Augmented Generation) [3]

Благодаря RAG архитектуре, LLM могут действовать как интеллектуальные интерфейсы к корпоративным базам знаний, объединяя генеративные возможности и точные данные. В контексте проектного менеджмента это открывает новые возможности:

- анализ проектной документации (технические задания, P&ID-схемы, отчёты по монтажу);
- генерация итоговых протоколов совещаний и недельных отчётов;
- извлечение информации о рисках, отклонениях и сроках;
- автоматизация ответов на типовые запросы сотрудников.

Как отмечается в отчёте OpenAI [OpenAI, 2023], LLM способны выполнять комплексные задачи управления знаниями, включая анализ контекста и рекомендации. Таким образом, RAG и языковые модели становятся естественным развитием классических СППР, решая их главные ограничения: невозможность работы с неструктурированными текстами и низкую адаптивность к изменениям данных.

Современные решения, такие как ChatGPT Enterprise, Microsoft Copilot, Google Gemini и открытые платформы (LangChain, AnythingLLM), демонстрируют практическую реализуемость подхода RAG. Они позволяют организациям создавать внутренние интеллектуальные ассистенты, которые могут «понимать» проектную документацию, отслеживать ход работ и помогать руководителю принимать решения в реальном времени.

Таким образом, использование языковых моделей в управлении проектами можно рассматривать как следующий этап эволюции систем поддержки решений – от статических баз знаний к динамическим, самообучающимся интеллектуальным системам.

4. Современные решения для поддержки проектного менеджера на основе LLM

Развитие больших языковых моделей (LLM) привело к появлению целого класса программных решений, ориентированных на автоматизацию интеллектуальных функций управления проектами. Эти инструменты позволяют проектному менеджеру получать структурированные ответы на запросы, анализировать документацию и получать рекомендации по дальнейшим действиям.

4.1. Коммерческие решения

Наиболее известные промышленные реализации – ChatGPT Enterprise (OpenAI), Microsoft Copilot и Google Gemini. Они интегрируются с корпоративными системами (MS Teams, Office 365, Google Workspace) и обеспечивают генерацию отчетов, резюме совещаний, планирование задач и анализ переписки.

Модели работают в облачной среде и предлагают расширенные возможности безопасности – изолированное хранение данных и аудит обращений.

Однако такие решения зависят от внешних серверов и не всегда допускают загрузку конфиденциальной проектной документации [6].

4.2. Открытые и локальные решения

Для организаций, где критична конфиденциальность, всё чаще используются открытые фреймворки – LangChain, LlamaIndex, AnythingLLM, LocalAI.

Они позволяют развернуть RAG-архитектуру на локальном сервере и работать с внутренними документами без передачи данных во внешние облака.

Примером такого подхода является AnythingLLM, обеспечивающий загрузку PDF, DOCX, DWG и изображений с автоматическим индексированием текста и поиском по векторной базе.

В контексте инженерных проектов подобные системы могут выступать как ассистенты проектного менеджера, выполняя поиск технических инструкций, требований и нормативов, а также формируя краткие сводки и планы работ.

Таблица 1. – Сравнение популярных решений на основе LLM

Решение	Тип развертывания	Поддержка RAG / поиск по документам	Конфиденциальность данных	Типичные задачи в PM
ChatGPT Enterprise	Облачное	Частично (через API)	Высокая (ограничено облаком)	Анализ текста, отчёты, планирование
Microsoft Copilot	Облачное + MS 365	Интеграция с корп. данными (SharePoint)	Средняя (через MS Graph)	Формирование документов, сводок
Google Gemini	Облачное	Да (через Google Workspace)	Средняя	Анализ почты, таблиц, отчётов
LangChain	Локальное/гибридное	Да (встроенный RAG)	Полная контрольность	Индексация баз знаний, чат-боты
AnythingLLM	Локальное	Да (встроенный RAG и OCR)	Полная	Поиск по документам, отчёты и дневники проектов

4.3. Потенциал применения в проектном менеджменте

Языковые модели позволяют автоматизировать ключевые аспекты деятельности менеджера:

- анализ документации – поиск нужных инструкций, норм и чертежей;
- генерацию текстов – отчёты, итоги совещаний, планы недельных работ;
- управление коммуникациями – обработка переписки, выделение приоритетов;
- обучение персонала – создание интерактивных FAQ и чата по внутренним регламентам.

В работах отечественных авторов отмечается, что эффективная СППР должна учитывать разнородность данных и скорость их поступления [4] – именно эти требования наиболее полно реализуются в архитектурах RAG и LLM.

Внедрение таких систем в инженерных проектах позволяет проектному менеджеру сократить время на поиск информации, повысить точность решений и снизить нагрузку на экспертные группы.

Одновременно возникает ряд ограничений: необходимость локального развёртывания, учёт отраслевых онтологий и контроль качества ответов модели.

5. Перспективы интеграции LLM и традиционных СППР

Опыт эксплуатации классических систем поддержки принятия решений (СППР) показывает, что их главными достоинствами являются прозрачность логики, объяснимость рекомендаций и возможность формализации экспертных знаний. Однако такие системы плохо работают с неструктурированными данными — текстами отчётов, перепиской, технической документацией.

Напротив, современные языковые модели (LLM) превосходно справляются с анализом естественного языка, но страдают от непрозрачности reasoning-процессов и могут порождать ошибки интерпретации.

Решение заключается в построении гибридных архитектур, объединяющих преимущества обоих подходов. В такой модели языковая модель выполняет роль фронтенда — интеллектуального интерфейса для общения с пользователем, а за верификацию и интерпретацию выводов отвечает традиционный модуль СППР с онтологической базой и нечёткой логикой.

В соответствии с концепцией адаптивных нечетких систем, описанной в работе [7], использование нечётких правил позволяет учитывать неопределённость факторов, формализовать экспертные знания и описывать слабо структурированные процессы, характерные для управления инженерными проектами. Такие системы способны интерпретировать параметры, выраженные не в точных числах, а в лингвистических терминах («высокий риск», «средняя загрузка» и т.п.), что делает их удобными при моделировании сложных управленческих ситуаций.

Онтологический подход, представленный в исследовании [8], обеспечивает структурированное описание предметной области с определением взаимосвязей между её сущностями — оборудованием, задачами, рисками, сроками и другими компонентами проекта. Онтология выступает основой для интеграции данных из различных источников и формирования единого пространства знаний, которое может использоваться интеллектуальной системой для рассуждений и принятия решений.

Комбинация этих подходов с LLM обеспечивает гибридную интеллектуальную поддержку, где решения формируются на основе как

статистического анализа текста, так и формализованных экспертных знаний.

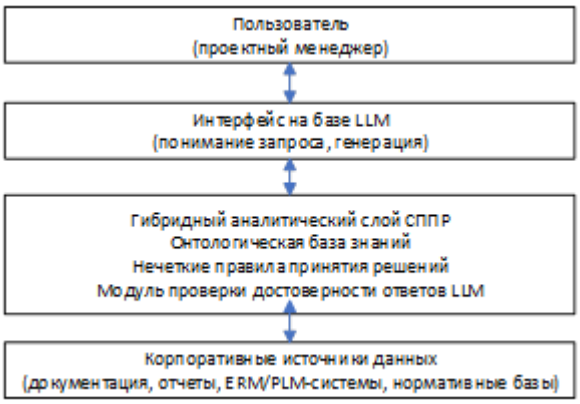


Рис. 5. Архитектура RAG (Retrieval-Augmented Generation)

Подобная архитектура соответствует принципу retrieval + reasoning: языковая модель выполняет роль генератора идей и интерфейса, а модули онтологий и нечёткой логики — роль фильтра, обеспечивающего достоверность и объяснимость результатов.

В перспективе такие гибридные решения позволят:

- интегрировать ассистента проектного менеджера с корпоративными системами (PLM, ERP, MES);
- использовать LLM для предварительного анализа данных, а СППР — для окончательного вывода;
- накапливать формализованные знания на основе диалогов и решений, принятых в проектах;
- повысить воспроизводимость и прозрачность управленческих решений.

Подобный подход формирует основу для интеллектуальных ассистентов нового поколения, которые смогут не только анализировать информацию, но и объяснять свои выводы, опираясь на внутреннюю модель проекта.

6. Заключение

Анализ современных исследований и решений показывает, что развитие языковых моделей (LLM) стало ключевым фактором эволюции систем интеллектуальной поддержки принятия решений. Если традиционные СППР опираются на формализованные правила и требуют ручного сопровождения, то LLM позволяют автоматизировать анализ больших объёмов текстовой информации и формировать рекомендации на естественном языке.

В области управления инженерными проектами применение языковых моделей открывает новые возможности для проектного менеджера: автоматизацию анализа документации, формирование отчётности, прогнозирование рисков и поддержку коммуникаций между участниками проекта. Архитектуры типа Retrieval-Augmented Generation (RAG) обеспечивают интеграцию языковых моделей с корпоративными базами знаний, а локальные решения (такие как AnythingLLM и LangChain) позволяют использовать эти технологии без угрозы утечки данных.

В то же время сохраняются проблемы, связанные с интерпретируемостью решений, необходимостью проверки достоверности выводов и адаптацией под отраслевые стандарты. Наиболее перспективным направлением является создание гибридных интеллектуальных систем, сочетающих преимущества LLM, онтологических моделей и нечёткой логики. Такие решения сохраняют объяснимость классических СППР и обеспечивают адаптивность современных языковых моделей.

В дальнейших исследованиях планируется разработка прототипа интеллектуального ассистента проектного менеджера для инженерных проектов, реализующего гибридную архитектуру «LLM + СППР» и ориентированного на практические задачи анализа проектной документации, планирования и управления знаниями.

Список используемых источников

Список используемых источников оформляется согласно ГОСТ Р 7.0.5 2008 «Библиографическая ссылка».

1. Turban E., Sharda R., Delen D. Decision Support and Business Intelligence Systems. – 10th ed. – Pearson Education, 2022. – 752 p
2. Мухаметьянова Р. И., Черняховская Л. Р., Ризванов Д. А. Разработка базы нечётких правил поддержки принятия решений при управлении проектом // Информационные технологии интеллектуальной поддержки принятия решений. – 2023. – С. 45–50.
3. Lewis P., Perez E., Piktus A. и др. Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks // Advances in Neural Information Processing Systems 33 (NeurIPS). – 2020. – С. 9459–9474.
4. Ризванов Д. А., Галеев А. Р. Система поддержки принятия решений при оценке имущества для целей залога // Информационные технологии интеллектуальной поддержки принятия решений. – 2024. – С. 112–118.
5. Vaswani A., Shazeer N., Parmar N., Uszkoreit J., Jones L., Gomez A. N., Kaiser Ł., Polosukhin I. Attention Is All You Need // Advances in Neural Information Processing Systems 30 (NeurIPS 2017). – 2017. – P. 5998–6008. – URL: https://arxiv.org/abs/1706.03762

6. OpenAI. GPT-4 Technical Report // arXiv:2303.08774. – 2023. – URL: https://arxiv.org/abs/2303.08774

7. Jang J.-S. R. ANFIS: Adaptive Network-Based Fuzzy Inference System // IEEE Transactions on Systems, Man, and Cybernetics. – 1993. – Vol. 23, No. 3. – P. 665–685.

8. Gruber T. R. Toward Principles for the Design of Ontologies Used for Knowledge Sharing // International Journal of Human-Computer Studies. – 1995. – Т. 43. – № 5–6. – С. 907–928.

Особенности реализации сервиса аналитического вычисления на графах

Жаркова О. А.
УУНиТ, каф. ТК
Уфа, Россия
e-mail: oksana.zharkova@yandex.ru

Насыров Р. В.
УУНиТ, каф. ТК
Уфа, Россия
e-mail: nrash@yandex.ru

Аннотация[1]

В данной статье рассматриваются особенности реализации сервиса для аналитических вычислений на графах с применением современной сервисной архитектуры. В условиях усложнения программных систем, авторы обосновывают целесообразность разделения приложения на отдельные сервисы и микросервисы, что повышает его доступность, безопасность и масштабируемость.

Проведен сравнительный анализ популярных фреймворков для создания веб-API, таких как Flask и FastAPI. По результатам сравнения, для реализации задачи был выбран FastAPI благодаря его высокой производительности, поддержке асинхронности, автоматической валидации данных и генерации документации.

Представлена модульная архитектура приложения, состоящая из интерфейса ввода, блока управления преобразования, сервиса формирования выражения и сервиса вычисления. Особое внимание уделяется способам хранения данных, включая использование чисел (float), двумерных массивов (матриц) и JSON-объектов для обмена информацией.

В заключительной части статьи проводится исследование эффективности сервисной реализации, подтверждающее ее преимущества перед локальным вариантом в части доступности, безопасности и перспектив развития, что делает предложенное решение масштабируемым и перспективным инструментом для аналитических вычислений на графах.

Введение

Задача вычисления на графах является важным инструментом при решении широкого круга задач как научного, так и практического характера. В качестве примеров типовых задачи можно привести следующее.

1. Деятельность специалиста по моделированию и оцениванию эффективности деятельности и применения приложений: Для данного специалиста сервис аналитического вычисления на графах открывает новые возможности в глубоком понимании взаимодействия компонентов приложений, поведения пользователей и влияния различных факторов на общую производительность. Анализируя графовые представления данных о производительности, логах, сетевых соединениях и взаимодействиях пользователей, специалист может более точно моделировать и оценивать эффективность как отдельных модулей, так и всего приложения в целом. Это позволяет выявлять "узкие места", прогнозировать потенциальные проблемы, оптимизировать архитектуру и ресурсы, а также обосновывать принимаемые решения по развитию и модернизации программного обеспечения на основе объективных аналитических данных.
2. Обучение на тренажерах разного рода: В контексте обучения на тренажерах, сервис аналитического вычисления на графах может быть использован для создания более реалистичных и адаптивных сценариев. Графовые модели позволяют представлять сложные взаимосвязи между действиями обучающегося, состоянием тренажера и последствиями принимаемых решений. Анализируя паттерны поведения пользователя в графовом пространстве, система может динамически корректировать сложность задач, предоставлять персонализированную обратную связь, моделировать различные нестандартные ситуации и оценивать степень освоения навыков, делая процесс обучения более эффективным и целенаправленным.
3. Оценка человеко-машинного интерфейса на основе сценариев: Для оценки юзабилити и эффективности человеко-машинного интерфейса (ЧМИ) сервис на графах позволяет моделировать и анализировать пользовательские сценарии. Представляя

[1] Труды XI Международной научной конференции "Информационные технологии интеллектуальной поддержки принятия решений", 11-13 ноября, Уфа-Чандигарх-Фандербейлпарк, 2025

последовательность действий пользователя, взаимодействия с элементами интерфейса и достижение конечной цели как графовую структуру, можно выявлять наиболее оптимальные и, наоборот, наиболее затрудненные пути выполнения задач. Анализ графовых метрик, таких как длина пути, количество шагов, время выполнения и частота ошибок, позволяет объективно оценивать, насколько интуитивно понятен и удобен интерфейс, предсказывать потенциальные проблемы с пользовательским опытом и предлагать конкретные рекомендации по его улучшению, основываясь на данных реальных или смоделированных сценариев использования.

Постановка задачи

Ранние проведенные исследования и разработки авторов позволили получить приложение, позволяющее проводить вычисления в аналитической форме. При этом осталось нереализованным разделение процедуры формализации, процедуры реализации и процедуры вычисления. Такое разделение обусловлено обеспечение необходимости создания разных вариантов интерфейсом для ввода данных. Такое разделение позволяет реализовать новых подход к программному обеспечению наукоемких процедур.

Такое усложнение приводит к разделению их на сервисы и микросервисы[1, 2, 3].

В настоящее время существует большое количество технологий в виде фреймворков и библиотек, которые позволяют создавать сервисы и микросервисы для решения конкретных задач. Например, Spring Boot, .NET Core, Django REST framework, Flask, FastAPI, Node.js (Express.js), Quarkus, Micronaut, Quarkus, Helidon.

В связи со сказанным целью настоящей работы является разработка методики и архитектуры распределенного вычисления на базе микросервисного подхода, обеспечивающего многопользовательского применение используемого метода.

Выбор технологии реализации

Для реализации поставленной задачи существуют следующие инструменты:

1. На языке Java:
   a) Spring Boot: Один из самых популярных фреймворков для создания RESTful веб-сервисов и микросервисов на Java. Предоставляет широкий набор инструментов для разработки, включая встроенный сервер, автоматическую конфигурацию и интеграцию с другими Spring-проектами.
   b) Quarkus: Высокопроизводительный фреймворк для JVM, оптимизированный для облачных сред и serverless-приложений. Отличается быстрым временем запуска и низким потреблением памяти.
   c) Micronaut: Современный фреймворк для создания микросервисов, который использует AOT (Ahead-of-Time) компиляцию для достижения высокой производительности и низкого потребления ресурсов.
2. На языке Python:
   a) Django REST framework: Мощное расширение для фреймворка Django, предназначенное для создания веб-API.
   b) Flask: Легковесный микрофреймворк для Python, который отлично подходит для создания простых RESTful API и микросервисов.
   c) FastAPI: Современный, быстрый (высокопроизводительный) веб-фреймворк для создания API с версии Python 3.7+, основанный на стандартных подсказках типов Python.

Для последующего выбора инструмента реализации было проведено сравнение между Flask и FastAPI. Сравнение представлено в таблице 1.

Таблица 1. Сравнение Flask и FastAPI

Достоинства	Недостатки
Flask	
Гибкость: Flask — микрофреймворк, предоставляющий большую свободу в выборе архитектуры и реализации; Простой в изучении: Flask легко освоить, что делает его привлекательным для новичков; Большое сообщество: Flask имеет обширное сообщество разработчиков, что означает доступность множества документации, библиотек и примеров.	Меньшая автоматизация: Flask требует больше ручного кода для реализации базовых функций; Потенциально более сложная архитектура: гибкость может привести к сложностям в больших проектах, если не соблюдать определенную структуру.
FastAPI	
Скорость: FastAPI известен своей высокой скоростью и производительностью, особенно для API; Автоматизация: FastAPI предоставляет множество функций "из коробки", таких как валидация данных и документация API; Асинхронность: FastAPI поддерживает асинхронное программирование, что повышает производительность в приложениях с высоким трафиком; Документация: FastAPI автоматически генерирует документацию API (Swagger UI), что упрощает тестирование и использование API.	Меньшая гибкость: FastAPI более структурирован, чем Flask, что может ограничивать некоторые архитектурные решения; Потенциальная сложность для новичков: FastAPI требует большего понимания асинхронного программирования и работы с API.

Исходя из данного сравнения, выбор пал на FastAPI.

Описание решения

Для разработки приложения была разработана архитектура, особенности которого представлены на рисунке 1.

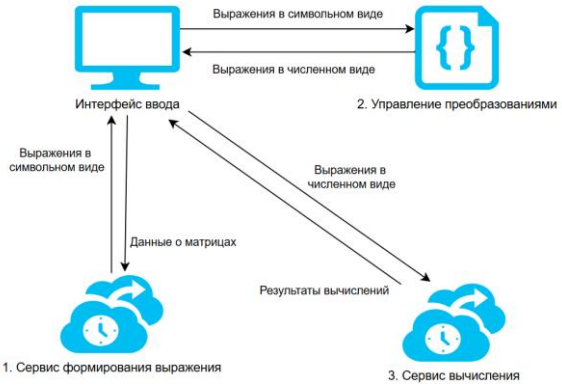


Рис. 1. Архитектура системы

Для удобства управления данными, их представления и преобразования данные хранятся в нескольких видах:

1. Числа: Значения из полей ввода, преобразованные в «float».
2. Двумерные массивы (матрицы): «matrixPData», «matrixMData», «matrixP», «matrixM». Это основная структура для представления матриц.
3. JSON-объекты: Для передачи данных на сервер и получения результатов от сервера.

В дальнейшем планируется улучшить данную технологию хранения.

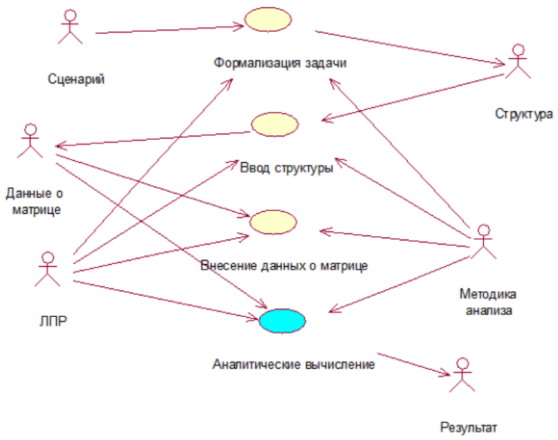


Рис. 2. Use-case диаграмма серверного приложения

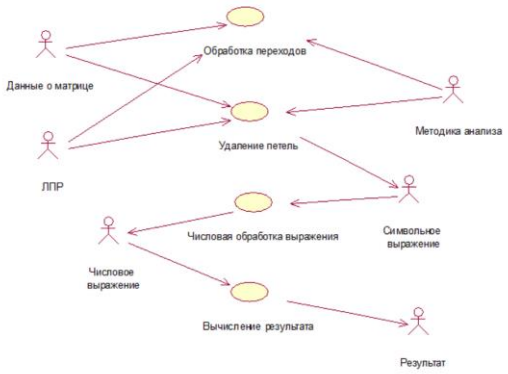


Рис. 3. Детализация use-case «Аналитическое вычисление»

Сценарий «Аналитическое вычисление»

1. Пользователь инициирует процесс вывода матриц;
2. Пользователь инициирует процесс символьной обработки матриц;
3. Пользователь инициирует числовую обработку символьных выражений;
4. Пользователь инициирует вычисление числовых выражений;
5. При желании пользователь может повторить пункты 3 и 4 с той же самой матрицей.

Особенности реализации

Поскольку требовалось решить несколько задач с принципиально разной функциональностью, в результате было создано гибридное приложение. Общая схема анализа представлена на рисунке 4.

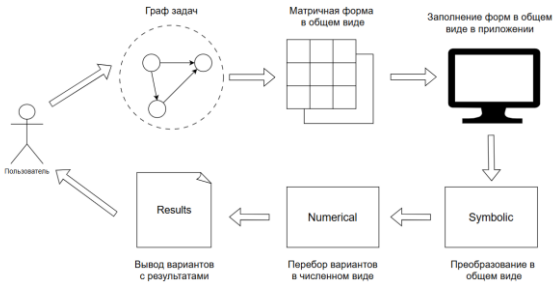


Рис. 4. Общая схема анализа

Этапы «Матричная форма в общем виде», «Заполнение форм в общем виде в приложении», и «Вывод вариантов с результатами» представлены с помощью html+javascript. Этапы «Преобразование в общем виде» и «Перебор вариантов в численном виде» реализованы в виде отдельных микросервисов с использованием языка Python и библиотеки pydantic для преобразования типов данных. Это позволило разделить контекст ввода данных и контекст вычислений по разным функциональным блокам.

Выводы и заключения

Выделение разных контекстов позволило разделить функциональность на отдельные сервисы. Разделение функциональности на отдельные сервисы (формирования выражения, вычисления) и использование FastAPI как основного инструмента для API обеспечивает не только высокую производительность, но и делает приложение более доступным, безопасным и перспективным для дальнейшего развития. Такой подход делает разработанный набор сервисов аналитического вычисления на графах масштабируемым решением.

Список литературы

1. Ньюмен С. Создание микросервисов. — СПб.: Питер, 2016. — 304 с.: ил. — (Серия «Бестселлеры O’Reilly»). — ISBN 978-5-496-02011-4.
2. Наркхе В. П., Суле А. П., Шрираме С. Д. Обзор микросервисов: шаг к распределенным системам (A survey of microservices: A step towards distributed systems) // В сб.: 2017 International Conference on Computer Communication and Informatics (ICCCI), 2017, с. 1-5. DOI: 10.1109/ICCCI.2017.8117478.
3. Чен П., Ванг Ю., Лю Дж. «Анализ масштабируемости и производительности приложений на основе микросервисов» (Scalability and performance analysis of microservices-based applications) // Journal of Cloud Computing, т. 9, № 1, с. 1-16, 2020.
4. Мартин, Роберт К. Быстрая разработка программ : Принципы, примеры, практика / Роберт К. Мартин, а также Джеймс В. Ньюкирк и Роберт С. Косс. - М. [и др.] : Вильямс, 2004 (ГПП Печ. Двор). - 739, [5] с. : ил.; 24 см.; ISBN 5-8459-0558-3
5. Пратт, Теренс У. Языки программирования : Разраб. и реализация : [Пер. с англ.] / Т. Пратт, М. Зелковиц. - 4. изд. - М. [и др.] : Питер, 2002. - 688 с. : ил.; 24 см. - (Серия Классика computer science).; ISBN 5-318-00189-0
6. Крачтен, Филипп. Введение в Rational Unified Process / / Филипп Крачтен; [Пер. с англ. и ред. А.В. Назаренко]. — 2. изд. — М. [и др.] : Вильямс, 2002. — 239 с. : ил., табл. : 24 см — (Серия "Объектные технологии").; ISBN 5-8459-0239-8
7. Бакусов, Л. М. Методы и модели причинно-структурного анализа в исследовании самоорганизующихся систем / Л. М. Бакусов. -

Москва : Машиностроение, 2005 (Уфа : Уфимский полиграфический комбинат). – 227 с. : ил., табл.; 21 см.; ISBN 5-217-03322-3

8. Гринберг Мигель. Разработка веб-приложений с использованием Flask на языке Python / пер. с англ. А. Н. Киселева. – 2-е изд., эл. – Москва : ДМК Пресс, 2023. – 274 с. – ISBN 978-5-89818-311-0
9. Соммервилл, Иан. Инженерия программного обеспечения / Иан Соммервилл; [Пер. с англ. А. А. Минько и др.]. – 6. изд. - М. [и др.] : Вильямс, 2002. – 623 с. : ил., табл.; 24 см.; ISBN 5-8459-0330-0
10. Ильин, Ф.В. Математические модели оценки эффективности инвестиций и принятия управленческих решений в условиях риска// Автореф. дисс. на соиск. уч. ст. канд. техн. Наук / Ф.В. Ильин. – Санкт-Петербург, 2017. – 16 с.
11. Есиков, Д.О. Оценка эффективности методов решения задач обеспечения устойчивости функционирования распределенных информационных систем / Д.О. Есиков // Программные продукты и системы. – 2015. – С. 241-256.
12. Исмагилова, Г.В. Некоторые подходы к оценке эффективности деятельности / Г.В. Исмагилова, М.С. Рухлядева, О.Г. Щемерова // Казанская наука №1. – 2015. – С. 75-77.
13. Прудникова, Л.В. Оценка инновационной деятельности и эффективности деятельности ОАО «Витебские ковры» в соответствии с этапами производственного цикла / Л.В. Прудникова, Е.В. Шаткова // УО «ВГТУ». – 2018. – С. 249-251.
14. Северцев, Н.А. Исследование эффективности функционирования технической системы специального назначения на основе марковского процесса / Н.А. Северцев // Научный вестник МГТУ ГА. – 2016. – стр. 61-68.
15. Плескунов, М.А. Задачи сетевого планирования: учебное пособие / Плескунов М.А. – Екатеринбург: Изд-во Урал. ун-та, 2014 – 92 с.
16. Кудрявцев, Е.М. Microsoft Project. Методы сетевого планирования и управления проектом / Е.М. Кудрявцев. – М.: Книга по Требованию, 2016. – 238 с.
17. Андриченко, А. Timeline — новое поколение технологических САПР. Семантический подход / А. Андриченко // САПР и графика. – 2011. – С. 80-84.

Концепция автоматизированной оценки графических интерфейсов корпоративных информационных систем с использованием синтетических пользовательских агентов

Кононова Е. М.
Уфимский университет науки и технологий
каф. АСУ
Уфа, Россия
e-mail: knnv.em@gmail.com

Кононов Н. А.
Уфимский университет науки и технологий
каф. АСУ
Уфа, Россия
e-mail: knnv.nkt@gmail.com

Кононов М. А.
Уфимский университет науки и технологий
каф. АСУ
Уфа, Россия
e-mail: knnv.mk@gmail.com

Аннотация

Актуальность оценки эргономики и работоспособности графических пользовательских интерфейсов (GUI) в условиях цифровой экономики сложно переоценить. Статья посвящена реинженерингу традиционного бизнес-процесса тестирования GUI, который характеризуется высокой трудоемкостью, наличием многочисленных итерационных циклов и значительными временными и финансовыми затратами. Авторы детально анализируют гипотетическую, но типичную модель процесса, включающую этапы от сбора требований до тестирования каркасного прототипа на фокус-группе, выявляя ключевые узкие места и скрытые издержки, связанные с простоем специалистов и дорогостоящими исправлениями ошибок на поздних стадиях.

В качестве решения предлагается автоматизация процесса с использованием синтетических пользовательских агентов (СПА) – программных алгоритмов, имитирующих поведение реальных пользователей. Описывается трансформированный процесс, в котором СПА проводят статический анализ макета на соответствие эвристикам юзабилити (например, Нильсена) и динамическое тестирование интерактивного прототипа по заданным сценариям, собирая объективные метрики. К ключевым преимуществам автоматизированного подхода авторы относят радикальное сокращение времени обратной связи (с недель до часов), значительное снижение затрат за счет устранения необходимости в фокус-группах, а также повышение объективности, воспроизводимости и масштабируемости тестирования. Делается вывод о том, что внедрение СПА является стратегическим шагом для минимизации рисков и перевода контроля качества пользовательского опыта на новый уровень.

Ключевые слова: графический пользовательский интерфейс; синтетические пользовательские агенты; автоматизация тестирования; бизнес-процесс; реинжиниринг; искусственный интеллект.

1. Введение

Актуальность вопросов, связанных с оценкой работоспособности и эргономики графических пользовательских интерфейсов (GUI), в современной цифровой экономике сложно переоценить. В условиях высокой конкуренции и стремительного развития технологий качество интерфейса становится ключевым фактором успеха любого программного продукта, оказывая прямое влияние на такие бизнес-показатели, как конверсия, удержание пользователей и лояльность к бренду [1]. Исторически сложилось, что оценка эта была трудоемкой и затратной, однако с учетом актуальных тенденций реинженеринга бизнес-процессов, направленных на оптимизацию и автоматизацию, на первый план выходят методики, основанные на использовании синтетических пользовательских агентов. Данная работа посвящена анализу традиционного подхода к оценке GUI, выявлению его системных недостатков и обоснованию перехода к автоматизированной модели с применением передовых технологий имитации пользовательского поведения.

2. Современные методы

На сегодняшний день существует два основных методологических подхода к оценке информационных систем с привлечением подобных технологий: эвристический и эмпирический.

Эвристические методы оценки фокусируются на экспертной проверке интерфейса на соответствие установленным отраслевым стандартам и принципам, таким как проверка на удобство использования (usability). Ярким примером служит оценка степени соответствия интерфейса знаменитым десяти эвристикам юзабилити, разработанным Якобом Нильсеном [2], которые охватывают такие аспекты, как видимость статуса системы, соответствие между системой и реальным миром, пользовательский контроль и свобода, единообразие и стандарты, предупреждение ошибок, распознавание вместо воспоминания, гибкость и эффективность использования, эстетичный и минималистичный дизайн, помощь пользователям в распознавании, диагностике и исправлении ошибок, а также наличие помощи и документации. Несмотря на свою формализованность, этот подход в значительной степени зависит от субъективного опыта и квалификации конкретного эксперта, что может приводить к пропуску нетипичных, но критичных для реальных пользователей проблем.

С другой стороны, эмпирические методы проверки основываются на сборе данных от реальных пользователей. К ним традиционно относятся тестирование ранней версии информационной системы на фокус-группе, открытое бета-тестирование программного продукта, сбор мнения пользователей с помощью анкетирования и опросов, а также предварительное моделирование и прототипирование интерфейсов для выявления их функциональности и удобства на ранних стадиях разработки. Эмпирические данные, получаемые таким образом, обладают высокой ценностью, так как отражают реакцию живых людей, однако их сбор сопряжен с целым рядом организационных и методологических трудностей. Несмотря на свою доказанную эффективность, бизнес-процесс, лежащий в основе всех описанных подходов, требует от компании-разработчика значительных финансовых, временных и репутационных вложений, что заставляет искать более эффективные пути. Именно синтез объективности эвристик и масштабируемости эмпирического подхода, но без присущих им человеческих ограничений, предлагают современные автоматизированные системы.

Актуальность настоящего исследования обусловлена необходимостью разработки строгих методов оценки и проектирования графических пользовательских интерфейсов (ГПИ) корпоративных информационных систем (КИС). Научная новизна работы заключается в формулировке формального метода описания ГПИ, который обеспечивает полную и однозначную спецификацию его структурных и поведенческих аспектов. Обоснование метода включает доказательство его полноты и адекватности для решения задач комплексной оценки качества интерфейсов.

Практическая значимость исследования раскрывается в двух ключевых направлениях. В **прикладном аспекте** предлагаемый формализм может быть положен в основу инструментов автоматизированного анализа. В частности, он открывает путь к реализации интеллектуального A/B-тестирования, где нейросетевые модели, опираясь на формальное описание интерфейса, способны не только количественно сравнивать варианты дизайна, но и генерировать семантически насыщенные отчеты с конкретными рекомендациями по оптимизации пользовательских сценариев в рамках реинжиниринга бизнес-процессов.

В **исследовательском аспекте** применение метода создает основу для формирования репрезентативных корпусов данных о структуре и поведении интерфейсов. Статистический анализ таких корпусов позволит выявить фундаментальные паттерны эффективного дизайна, что является необходимым условием для перехода от эвристических подходов к построению строгой, верифицируемой теории проектирования пользовательских интерфейсов для КИС.

3. Существующий бизнес-процесс

Гипотетическая модель существующего бизнес-процесса оценки качества графического интерфейса представлена на рисунке 1.

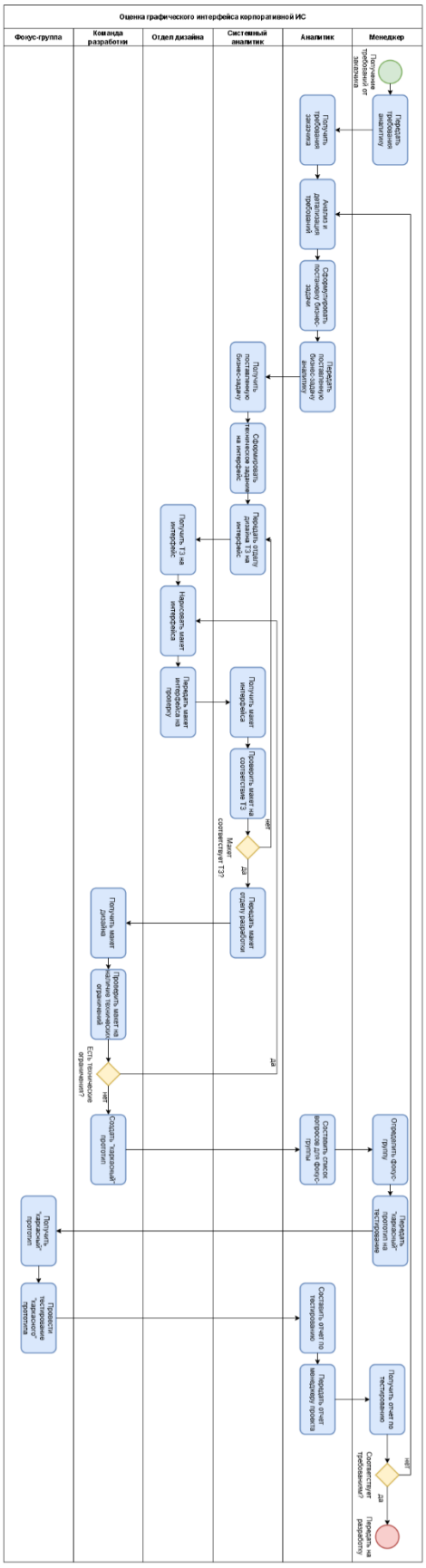


Рис. 1. Бизнес-процесс до реинжиниринга

Для наглядности представим гипотетически возможный, но весьма типичный бизнес-процесс в IT-компании, ориентированный на оценку качества графического интерфейса корпоративной информационной системы. Построим его динамическую модель, чтобы проиллюстрировать все этапы и потенциальные узкие места.

Процесс инициируется с получения менеджером проекта первоначальных требований от заказчика. Менеджер осуществляет первичную обработку этих требований, структурируя их, после чего передает документ бизнес-аналитику. Бизнес-аналитик, в свою очередь, проводит глубокий анализ и детализацию полученных требований, формулируя на их основе четкую постановку бизнес-задачи. Этот документ передается далее – системному аналитику.

Функция системного аналитика заключается в трансляции бизнес-требований в техническую плоскость. На основе поставленной задачи он формулирует детализированное техническое задание (ТЗ) на проектирование интерфейса, которое затем передается в отдел дизайна. Сотрудники отдела дизайна, получив техническое задание, приступают к созданию визуального макета будущего интерфейса. Завершенный макет передается обратно системному аналитику для проведения валидации.

Системный аналитик проверяет макет на строгое соответствие техническому заданию. Здесь возникает первая критическая точка принятия решения. Если макет не соответствует ТЗ, то отделу дизайна дается распоряжение на его переработку и доработку в соответствии с исходными требованиями. Если же макет успешно проходит проверку, он направляется в отдел разработки. Отдел разработки проводит собственную экспертизу, но уже на предмет наличия технических ограничений и осуществимости предложенных дизайнерских решений. Если технические ограничения выявлены, макет с соответствующими комментариями вновь возвращается дизайнерам на доработку. Если же макет признается технически реализуемым, команда разработки создает так называемый «каркасный» интерактивный прототип. Создание такого прототипа само по себе является ресурсоемкой задачей, требующей привлечения frontend-разработчиков и отвлекающей их от задач основной разработки.

На основе этого прототипа бизнес-аналитик составляет детальный список вопросов и сценариев для последующего тестирования. Этот этап включает в себя поиск релевантных респондентов, согласование их участия, бронирование специально оборудованного помещения (usability-лаборатории)

и планирование сессии, что может занять от нескольких дней до нескольких недель. Менеджер проекта в это время решает организационную задачу – определяет состав фокус-группы, привлекаемой для тестирования. После формирования группы «каркасный» прототип передается ей на тестирование. Участники фокус-группы проводят тестирование, следуя составленному списку вопросов и моделируя поведение конечных пользователей. По завершении этой сессии аналитик составляет сводный отчет по результатам тестирования и передает его менеджеру проекта. Данный отчет часто представляет собой компиляцию субъективных мнений, видеозаписей и ручных заметок, анализ которых требует значительного времени и интерпретации. Менеджер, изучая отчет, проверяет, соответствует ли прототип исходным требованиям заказчика. В случае положительного вердикта прототип передается на следующую стадию – полноценную разработку. В случае отрицательного – процесс может вернуться на один из предыдущих этапов.

3.1. Обоснование необходимости автоматизации

В описанном бизнес-процессе явно прослеживается наличие множества этапов, на которых создаются так называемые шлюзы (gateways) – точки принятия решений. В негативных сценариях, когда макет или прототип не проходят проверку, рабочий поток возвращается на более ранние стадии, образуя замкнутые циклы переделок и согласований. Такие итерационные «циклы» могут многократно увеличивать общее время выполнения всего бизнес-процесса, что, в силу закона экономики времени, напрямую и в геометрической прогрессии повышает итоговую стоимость проекта. Каждая такая итерация не только требует прямых затрат на оплату труда сотрудников, но и создает значительные альтернативные издержки, связанные с отвлечением команды от других проектов и замедлением выхода продукта на рынок.

Внешние, видимые затраты на организацию оценки, такие как оплата работы фокус-группы, могут на первый взгляд показаться незначительными в масштабах всего проекта. Однако скрытые издержки, связанные с простоем ключевых специалистов (аналитиков, дизайнеров, разработчиков) в процессе ожидания результатов тестирования и организации повторных циклов, являются существенными. Более того, ошибки и недочеты, заложенные в графическом интерфейсе на ранних этапах, по мере продвижения по жизненному циклу программного продукта обходятся все дороже. стоимость исправления дефекта, обнаруженного на этапе тестирования, в 15 раз выше, чем стоимость его исправления на этапе проектирования, а после выпуска продукта — уже в 100 раз [3]. Эффект «снежного кома» проявляется здесь в полной мере: недоработка, не выявленная на стадии проектирования, может привести к необходимости переписывания значительных объемов кода на стадии реализации или, что еще хуже, к кардинальным изменениям после релиза. Существует также не insignificant вероятность обнаружения критической ошибки в самой концепции интерфейса на поздних этапах разработки, что может привести к кратно большим финансовым вложениям, значительному удлинению сроков сдачи проекта, а в наиболее пессимистичном сценарии – и к полной отмене программного продукта.

4. Автоматизация с помощью синтетических пользовательских агентов

Автоматизация описанного бизнес-процесса видится закономерным и эффективным ответом на выявленные вызовы. Ключевым инструментом такой автоматизации выступают синтетические пользовательские агенты (СПА) – это специализированные программные алгоритмы, способные с высокой точностью имитировать поведение реального пользователя в рамках интерфейса. Они запрограммированы на выполнение предопределенных сценариев взаимодействия, сбор объективных метрик и их последующий анализ на соответствие заданным критериям производительности и удобства использования. По своей сути, СПА представляют собой виртуальных роботов, которые могут неограниченно и круглосуточно тестировать интерфейс в автоматическом режиме, освобождая человеческие ресурсы для более творческих и стратегических задач.

Рассмотрим, как трансформируется гипотетический бизнес-процесс при интеграции СПА. Начальные этапы остаются без существенных изменений: менеджер по-прежнему получает требования от заказчика, аналитик проводит их детализацию и формулирует бизнес-задачу, системный аналитик создает техническое задание, а дизайнеры разрабатывают макет, который проходит валидацию. Ключевое изменение происходит после одобрения макета системным аналитиком. Вместо передачи макета в разработку для создания трудоемкого «каркасного» прототипа, техническое задание и утвержденный макет загружаются в систему автоматизированного тестирования, поддерживающую работу СПА.

Далее, системный аналитик или привлеченный QA-инженер осуществляет настройку синтетического пользовательского агента. Этот процесс включает в себя несколько важных шагов. Во-первых,

производится формализация критериев оценки на основе ТЗ и лучших практик (тех же эвристик Нильсена). Во-вторых, описываются типичные пользовательские пути (user journeys), которые должен имитировать агент, например, «регистрация нового пользователя», «поиск товара в каталоге», «оформление заказа». В-третьих, определяются ключевые количественные показатели (метрики) для оценки: время отклика интерфейса, количество кликов для достижения цели, процент успешно завершенных сценариев, частота возникновения ошибок и т.д. Этот этап конфигурации является инвестицией, которая многократно окупается за счет последующей полной автоматизации проверок.

После настройки СПА приступает к работе. На первом этапе он может проводить статический анализ макета (а в последствии – интерактивного прототипа) на соответствие заданным эвристикам. Алгоритм способен автоматически проверять контрастность цветов, размеры шрифтов, соответствие стандартам доступности (например, WCAG) [4], логику расположения элементов навигации и многое другое. Результатом такого анализа является детализированный отчет, который в автоматическом режиме отправляется дизайнерам для исправления выявленных недочетов. Это позволяет устранить значительный пласт базовых ошибок еще до начала фазы активной разработки, что является воплощением принципа «найди и исправь дефект как можно раньше».

На следующем этапе, уже на основе интерактивного прототипа, СПА начинает выполнение пользовательских сценариев. Агент способен имитировать действия тысяч виртуальных пользователей одновременно, причем с разной скоростью взаимодействия, в различной последовательности и даже с моделированием неидеальных условий, таких как сетевые задержки. Отчеты о тестировании формируются системой в режиме, близком к реальному времени. Если агент обнаруживает критическую ошибку, например, невозможность завершить ключевой сценарий из-за заблокированной кнопки, процесс может быть автоматически остановлен, а задача – возвращена на соответствующий этап (дизайнерам или разработчикам) без необходимости длительных межфункциональных согласований с менеджером и аналитиком. Это кардинально сокращает время обратной связи и стоимость внесения изменений.

5. Преимущества автоматизированного подхода

Наиболее очевидным и весомым преимуществом перехода на автоматизированную модель является беспрецедентный рост операционной эффективности. Традиционный процесс с привлечением фокус-группы по своей природе дискретен и отнимает недели из-за необходимости решения логистических задач (поиск и согласование участников, аренда помещения), подготовки материалов и последующего трудоемкого анализа субъективных качественных данных. СПА трансформирует этот прерывистый процесс в непрерывный и практически мгновенный поток данных. Тестирование, которое раньше было редким и значимым событием, становится постоянным фоновым процессом, интегрированным в конвейер разработки (CI/CD) [5]. Любое изменение в коде или дизайне может быть немедленно протестировано, что обеспечивает непрерывный контроль качества на протяжении всего жизненного цикла проекта. Это приводит к радикальному сокращению времени обратной связи – с нескольких недель до считанных часов или даже минут. Такой выигрыш в скорости позволяет команде проводить итерации значительно чаще, повышая общую гибкость (agility) разработки и сводя к минимуму риски накопления технического долга и критических ошибок.

Вторым ключевым преимуществом выступает ярко выраженная экономическая эффективность. Прямые и весьма существенные затраты на организацию работы фокус-групп – гонорары участникам, аренда помещений, административная работа – полностью устраняются. Затраты компании сводятся к стоимости вычислительных ресурсов, потребляемых системой автоматизации, которые, как правило, предсказуемы, прозрачны и на порядок ниже. Однако главная экономия лежит не в прямой экономии на тестировании, а в области управления рисками. СПА обеспечивают раннее выявление ошибок, сдвигая его на самые начальные стадии проектирования макета. Обнаружение фундаментальной проектной ошибки (fundamental flaw) на этапе wireframe, а не в готовом продукте, экономит колоссальные ресурсы, которые пришлось бы потратить на перепроектирование и переделку больших объемов кода, и предотвращает серьезный репутационный ущерб, связанный с выпуском некачественного продукта. Таким образом, автоматизация действует как страховой полис, минимизирующий потенциальные убытки.

Третье важнейшее преимущество – это объективность, воспроизводимость и масштабируемость получаемых данных. Фокус-группа, безусловно, поставляет ценные качественные инсайты, однако эти данные субъективны по своей природе и часто подвержены влиянию групповой динамики, личных предпочтений модератора или отдельных доминирующих участников. СПА же оперирует строгими, не зависящими от человеческого фактора количественными метриками: время отклика интерфейса, количество кликов до цели, процент успешных завершений сценариев, коэффициент

ошибок. Это обеспечивает статистически значимую, измеримую и воспроизводимую основу для принятия управленческих решений (data-driven decisions) [6]. Кроме того, синтетические агенты могут одновременно имитировать поведение не 10-15, а десятков тысяч виртуальных пользователей с разными характеристиками, географическим расположением и типами устройств, что позволяет проводить масштабное нагрузочное тестирование и выявлять проблемы производительности и устойчивости, которые принципиально невозможно обнаружить в условиях малой фокус-группы. Это дает беспрецедентное понимание того, как интерфейс будет вести себя в условиях, максимально приближенных к реальной эксплуатации.

6. Предлагаемый метод формального описания ГПИ

Авторы полагают, что автоматизированная оценка графического пользовательского интерфейса (ГПИ) корпоративной информационной системы невозможна без предварительной формализации его описания и перехода на более высокий уровень абстракции, учитывающий семантические аспекты интерфейса. Переход к автоматизированной оценке графического интерфейса КИС возможен лишь при условии его репрезентации в виде формальной онтологии, которая позволяет абстрагироваться от визуальной реализации и описать функционально-семантические отношения между элементами.

Следуя данному утверждению, графический интерфейс КИС можно рассмотреть, как множество компонентов (*C*), объединенных в представление в определенный момент времени t при мощности множества компонентов равного n – $GUI_t = < C_1^t; \dots; C_n^t >$. Данная запись подчеркивает темпоральный аспект интерфейса: его состояние не статично и может изменяться с течением времени, что отражается в изменении состава или свойств множества *C*.

Для детального анализа необходимо формализовать и сами компоненты. Каждый компонент C_t может быть описан как упорядоченная пара, состоящая из двух множеств параметров: $C_t = < I_t, O_t >$. Множество $I_t = \{I_1^t; I_2^t; \dots; I_k^t\}$ определяет входные параметры компонента в момент t, где k — мощность множества входов. Эти параметры могут включать как данные, полученные извне, так и внутренние свойства, определяющие состояние компонента. Множество $O_t = \{O_1^t; O_2^t; \dots; O_m^t\}$ определяет исходящие параметры компонента в момент t, где m — мощность множества выходов. Это множество описывает реакции компонента, которые могут быть переданы другим элементам интерфейса или системе (например, события, команды, новые данные и прочие реакции графического интерфейса, соответствующие прикладной логике).

Следовательно, предложенная формальная модель позволяет рассматривать графический интерфейс как иерархическую структуру, состоящую из взаимодействующих компонентов, состояние которых эволюционирует во времени, что создает основу для его автоматизированного анализа с учетом семантических аспектов. Тогда состояние интерфейса в момент t будет представлено как $S(t) = \{(c,\ I(c,t),\ O(c,t)) \mid c \in C(t)\}$.

В свою очередь, для описания синтетического пользователя (респондента) представляется следующее формальное описание: $A_b = < G_b, P_b, K_b, E_b >$. Рассмотрим каждый из компонентов модель.

Множество G представляет собой множество целей, которые пользователь пытается достичь в ходе эксплуатации графического интерфейса. В прикладном аспекте, они напрямую зависят от участия пользователя в бизнес-процессах. Например, изменить статус обрабатываемой заявки. При этом $g \in G$, следует отметить, что каждая цель соответствует некоторому состоянию интерфейса $S(g)$, который позволяет реализовать цель g, следовательно, ведет к цели определенной бизнес-операции.

Множество K – множество правил и эвристик, которыми оперирует синтетический пользователь (агент) $a_1 \in A$ для выбора действия и формирования нового состояния $S(t+1)$ на основе текущего состояний интерфейса $S(t)$ и своей цели g. Следовательно, $K: S(t) \times G \to A$.

Множество E – множество траекторий, представляющее собой множество картежей $e_1 = < S_{t_0} a_{t_0}; \dots .; S_{fin} a_{fin} >$, где fin – финальное, конечное, терминальное состояние.

Формальный метод описания графических пользовательских интерфейсов (ГПИ), предложенный авторами, открывает перспективы для интеграции с различными типами нейронных сетей. Такая интеграция позволяет проводить комплексную мультимодальную оценку качества интерфейса, анализируя его по множеству взаимосвязанных критериев. В качестве примера могут быть рассмотрены следующие метрики:

1. Количество успешных траекторий, т.е. траекторий достигших целей из определенного стартового состояния S_0^t. Пусть $T(S_o^t) = \{\tau_1, \tau_2, \dots, \tau_n\}$ – множество всех возможных траекторий, начинающихся в стартовом состоянии, где каждая траектория $\tau_i = (S_0, S_1, S_2, \dots S_k)$ при длине траектории k представляет собой последовательность состояний. Тогда количество успешных состояний определяется как: $N_{success}(S_o) = |\{\tau \epsilon\, T(S_o^t) : S_k \epsilon\, G\}|$, где $G \subseteq S$ – множество целевых состояний интерфейса.

2. Длина успешной траектории. Для каждой успешной траектории $\tau \in E$, где E – множество успешных траекторий, то длина вычисляется как количество переходов между состояниями $L(\tau_i) =$

$|\tau_i| - 1 = k$, при $|\tau_i|$ – количество состояний в траектории τ_i.

3. Робастность (устойчивость к ошибкам), подсчет множества ошибочных состояний интерфейса. Пусть $\Omega \subseteq S$ – множество ошибочных состояний интерфейса. Тогда метрика робастности для множества траекторий T может быть определена как $R = 1 - \frac{|\{S \in \Omega : \exists \tau \in T,\ S \in \tau\}|}{|\{S: \exists \tau.\ S \in \tau\}|}$. В рассматриваемом контексте робастность показывает способность интерфейса минимизировать попадание в ошибочные состояния, где значение к 1 указывает на высокую устойчивость.

Перспективы применения данной методологии лежат в области как прикладных задач, так и фундаментальных исследований в области Human-Computer Interaction (HCI).

7. Заключение

Таким образом, в данной работе предложен метод формального описания ГПИ, который составляет теоретический фундамент для реинжиниринга бизнес-процесса оценки пользовательских интерфейсов. Ключевым результатом является разработка формального метаязыка, позволяющего непротиворечиво специфицировать структурно-семантические характеристики ГПИ в отрыве от его конкретной визуальной реализации.

Предлагаемый метод обеспечивает необходимую абстракцию и формальную строгость для реализации продвинутых сценариев автоматизированного тестирования с использованием синтетических пользовательских агентов (СПА). Этот подход представляет собой не операционную оптимизацию, а стратегическую трансформацию, обеспечивающую:

1) повышение эффективности за счет непрерывного, масштабируемого тестирования;

2) снижение затрат путем минимизации ручного труда на всех стадиях жизненного цикла ПО;

3) минимизацию рисков благодаря выявлению дефектов пользовательского опыта и семантических несоответствий на ранних этапах разработки.

Несмотря на объективную потребность в первоначальных инвестициях в инструментарий и адаптацию процессов, долгосрочные преимущества — воспроизводимость, полный охват тестовых сценариев и объективность оценки — обосновывают устойчивую экономическую целесообразность данного подхода.

В перспективе конвергенция разработанного метода с системами искусственного интеллекта открывает путь к созданию самооптимизирующихся интерфейсов. Формальная модель служит основой для генеративных моделей, способных не только оценивать, но и синтезировать варианты дизайна, соответствующие выявленным фундаментальным паттернам. Это утверждает автоматизацию, основанную на формальных методах, в качестве нового отраслевого стандарта для создания доказуемо эффективных пользовательских интерфейсов.

Список используемых источников

1. Сорокина Ю.В., Баклакова В.В., Свинарева Ю.А. Влияние юзабилити сайтов на повышения конкурентоспособности услуг фирмы // Теория и практика современной науки. - 2016. - №8. - С. 883-886.
2. 10 Usability Heuristics for User Interface Design // nngroup URL: https://www.nngroup.com/articles/ten-usability-heuristics/ (дата обращения: 10.09.2025).
3. Boehm B. W. Software Engineering Economics. - Prentice Hall, 1981. - 784 с.
4. Web Content Accessibility Guidelines (WCAG) 2.2 // W3C URL: https://www.w3.org/TR/WCAG22/ (дата обращения: 14.09.2025).
5. Fowler M. Continuous Integration / M. Fowler // MartinFowler.com. – 2006. – URL: https://martinfowler.com/articles/continuousIntegration.html (дата обращения: 27.08.2025)
6. Трофимов В.В., Трофимова Л.А. О концепции управления на основе данных в условиях цифровой трансформации // Петербургский экономический журнал. - 2021. - №1-2. - С. 149-155.

Информационная поддержка процесса формирования документов о высшем образовании с использованием интеллектуальных технологий

Бублик Т. В.
Институт информатики, математики и робототехники
Уфимский университет науки и технологий
Уфа, Россия
e-mail: bubliktimur@gmail.com

Сарванов С. С.
Институт информатики, математики и робототехники
Уфимский университет науки и технологий
Уфа, Россия
e-mail: sergejsarvanov@gmail.com

Ковтуненко А. В.
Институт информатики, математики и робототехники
Уфимский университет науки и технологий
Уфа, Россия
e-mail: klimova.av@ugatu.su

Аннотация[1]

В статье рассматривается информационная поддержка процесса формирования документов о высшем образовании в вузе. Предложена архитектура решения, объединяющая технологии Resource Description (RDF) для хранения и структурирования нормативной и учебной информации и подход Retrieval-Augmented Generation (RAG) для генерации ответов с использованием больших языковых моделей (Large Language Models, LLM). Проектируемое решение предоставляет возможность поиска и обработки данных с указанием источников, что повышает прозрачность и обоснованность принимаемых решений в рамках управления образовательным процессом.

Ключевые слова: управление образовательным процессом; RDF-технологии; Retrieval-Augmented Generation, RAG; Large Language Models, LLM; цифровая трансформация; университет.

1. Введение

Современные процессы цифровой трансформации государственного управления и образования в России регулируются рядом ключевых нормативных документов.

В соответствии с Распоряжением Правительства Российской Федерации [1] приоритетным направлением государственной политики является создание интеллектуальных платформ, использующих искусственный интеллект и технологии работы с данными для повышения эффективности управленческих решений и предоставления цифровых услуг гражданам.

В сфере образования определено стратегическое направление [2], предусматривающее внедрение цифровых инструментов для поддержки всех этапов образовательного процесса и управления вузами.

Одновременно, согласно Указу Президента Российской Федерации [3], одним из приоритетов государственной политики является использование интеллектуальных систем анализа данных в научно-образовательной сфере.

Реализация данных стратегических инициатив требует перехода образовательных учреждений к новым моделям информационного взаимодействия, в которых решения принимаются на основе анализа нормативных и учебно-методических данных с применением технологий искусственного интеллекта. Однако, в большинстве университетов значительная часть рутинных операций осуществляются с применением

[1] Труды XI Международной научной конференции "Информационные технологии интеллектуальной поддержки принятия решений", 11-13 ноября, Уфа-Чандигарх-Фандербейлпарк, 2025

разрозненных программных средств и частично автоматизированных процессов, что затрудняет комплексную обработку информации. Процесс формирования документов о высшем образовании является одним из сложных и трудоемких процессов, требующих больших трудозатрат со стороны сотрудников университета.

В современных вузах широко внедряются и используются различные программные решения и технологии, направленные на цифровизацию образовательных и административных процессов. Так, в Казанском федеральном университете функционирует система «Электронное обучение КФУ» [4] на базе LMS Moodle, обеспечивающая проведение онлайн-курсов и прокторинг; в Нижегородском государственном университете им. Н. И. Лобачевского действует портал e-learning.unn.ru [5], предоставляющий доступ к дистанционным дисциплинам; в Уфимском государственном нефтяном техническом университете применяется специализированная платформа «Нефтегазовое образование», ориентированная на профессиональные онлайн-курсы [6] . Особый интерес представляет Информационная система университета (ИСУ) [7], реализованная в Уфимском университете науки и технологий (УУНиТ) на основе приобретённого программного решения, интегрирующая личные кабинеты студентов и сотрудников, расписания, электронные зачётные книжки и сервисы справочной поддержки.

Однако, несмотря на наличие у перечисленных решений развитых инструментов для управления учебным процессом и контентом, они в меньшей степени ориентированы на интеллектуальную обработку информации и анализ нормативных документов в рамках управления образовательным процессом. Отсутствие механизмов семантического поиска и генерации обоснованных ответов на естественные запросы пользователей создаёт функциональный разрыв между традиционными системами управления образовательным процессом и задачами поддержки управленческих решений. Таким образом, использование интеллектуальных технологий для управления образовательным процессом является актуальной задачей.

В статье рассматривается информационная поддержка процесса формирования документов о высшем образовании с использованием технологии RDF-хранилищ для обеспечения хранения и обработки нормативной и учебной информации [8], предлагается интеллектуальный поиск с использованием семантических связей и онтологических моделей [9], формирование ответов на естественные запросы с применением технологии RAG[10] и интеграцией с LLM[11], а также обязательное подкрепление каждого ответа ссылками на источники данных.

2. Обзор языковых моделей

Современные большие языковые модели представляют собой ядро большинства интеллектуальных систем, основанных на машинном обучении и генерации текста. В последние годы на рынке сформировался ряд крупных открытых и коммерческих решений, различающихся по качеству рассуждений, размеру параметров, скорости работы, лицензионным ограничениям и возможности локального развёртывания.

В рамках реализации информационной поддержки процесса формирования документов о высшем образовании был проведён анализ современных открытых больших языковых моделей (Таблица 1).

Таблица 1 - сравнение моделей LLM для локального развёртывания

Модель (семейство)	**Сильные стороны**	**Ограничения**
DeepSeek-R1 (reasoning-ориентированная)[12]	Открытая лицензия MIT, доступность весов в открытом доступе, высокая способность к рассуждениям, оптимизирована под задачи поддержки принятия решений	Версия 70B требует значительного объёма видеопамяти (от 128 ГБ VRAM); версии 32B и 14B подходят для рабочих станций, но имеют более ограниченные когнитивные возможности
DeepSeek-V3 (SMoE-архитектура)[13]	Повышенная эффективность обработки токенов за счёт механизма Sparse Mixture of Experts, высокая точность генерации в задачах анализа документов	Требует сложной настройки и значительных вычислительных ресурсов; ориентирована преимущественно на серверные среды с GPU-классом A100/H100
Qwen 2.5 (Alibaba)[14]	Хорошее соотношение производительности и ресурсов, качественная обработка текстов на нескольких языках, строгий JSON-	Некоторые версии распространяются частично под ограниченными лицензиями и требуют облачного доступа

	вывод для интеграции с внешними системами	
Gemma 2 (Google)[15]	Эффективность вычислений, простота локального развёртывания и обслуживания	Основная оптимизация под англоязычные задачи; экосистема инструментов менее развита в сравнении с DeepSeek и Qwen

Результаты исследования показали, что выбор конкретной модели должен определяться не только её вычислительными характеристиками, но и открытостью лицензии, качеством рассуждений и возможностью локального развёртывания. В контексте рассматриваемого проекта приоритет отдаётся решениям, обеспечивающим высокий уровень интерпретируемости и прозрачности работы алгоритмов, а также устойчивости при обработке специализированных текстов нормативно-правового характера.

Кроме того, важное значение имеет соблюдение требований к защите персональных данных, что делает предпочтительным использование моделей, способных функционировать в полностью автономном режиме без обращения к внешним облачным сервисам. Выбранная конфигурация позволяет реализовать гибкую и безопасную систему генерации управленческих решений, адаптированную к инфраструктуре университета.

С учётом совокупности этих факторов целесообразным представляется использование модели DeepSeek-R1, сочетающей открытость архитектуры, развитые возможности рассуждения и гибкость развёртывания.

3. Процесс формирования документов о высшем образовании

В университете данный процесс (Рис. 1) представляет собой последовательность взаимосвязанных действий, выполняемых сотрудниками структурных подразделений (деканатов, институтов и др.). В рамках итоговой аттестации обучающихся подразделение получает сводные данные об успеваемости выпускников, включая средние оценки по защите выпускной квалификационной работы (ВКР). На основании этих данных формируется сводная таблица с перечнем студентов, их персональными данными и номерами дипломов, выделяемых Министерством науки и высшего образования Российской Федерации. Далее сотрудник вручную дополняет таблицу сопутствующими сведениями — заявлениями, характеристиками, информацией, собранной из различных источников.

Для каждого выпускника подготавливается индивидуальный проект приказа о выдаче диплома, часть информации при этом может автоматически подтягиваться из информационной системы университета, однако значительная доля данных по-прежнему требует ручной проверки и корректировки. После утверждения приказа о выдаче дипломов запускается процесс формирования документа об образовании и приложения к нему. На данном этапе осуществляется автоматическое извлечение данных из информационной системы университета, проверка корректности сведений и формирование шаблона документа. Существенную сложность вызывает и отсутствие унифицированных шаблонов сопроводительных справок, поступающих из различных структурных подразделений и внешних организаций, что требует дополнительных временных затрат на их адаптацию и проверку.

Таким образом, процесс формирования дипломов характеризуется высокой трудоёмкостью, значительными объёмами ручной обработки данных и высокой вероятностью ошибок при согласовании информации. Эти особенности подтверждают необходимость разработки интеллектуального решения, обеспечивающего автоматизацию анализа и генерации документов на основе структурированных источников данных и семантических связей между ними.

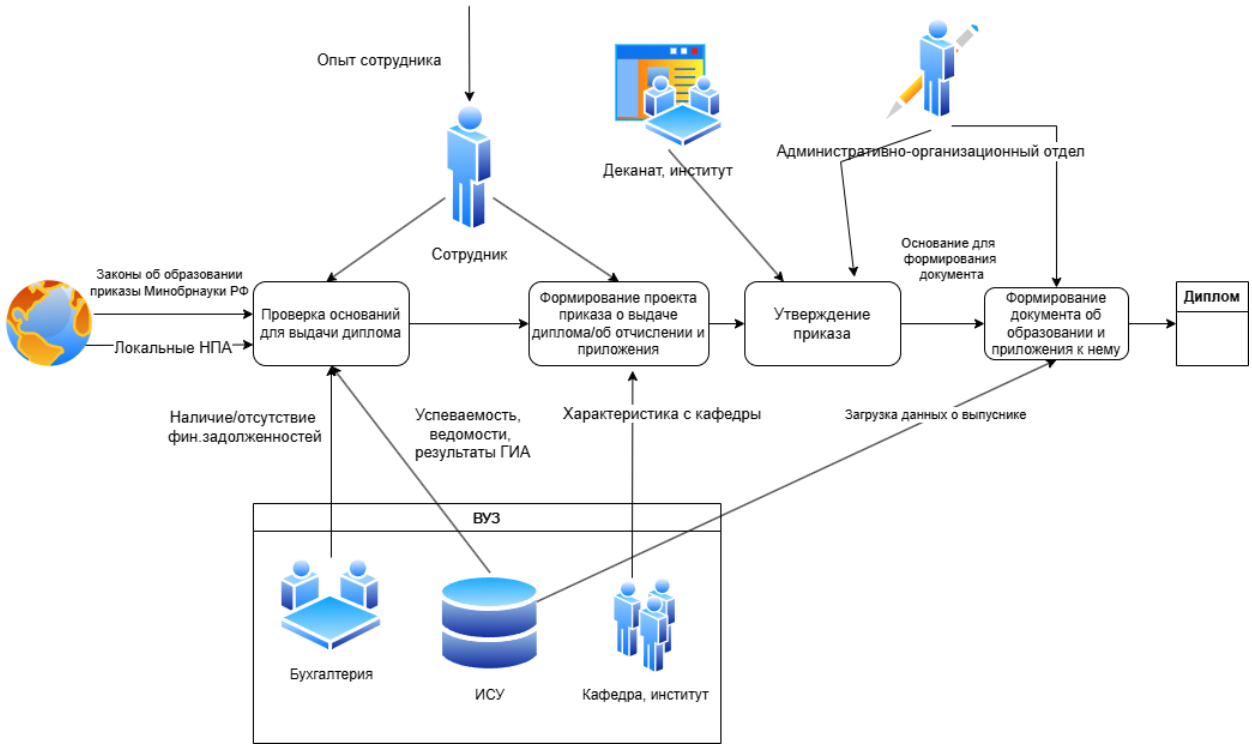


Рис. 1 - Мнемосхема процесса формирования документов о высшем образовании

4. Постановка задачи

В результате анализа процесса формирования документов о высшем образовании в университете, а также рассмотрения современных подходов к автоматизации управленческих процессов, предлагается разработка программного решения для реализации информационной поддержки при формировании документов о высшем образовании с использованием технологий семантического анализа и генерации обоснованных ответов с применением RDF-моделей данных и больших языковых моделей (LLM).

К решению предъявляются следующие функциональные требования:

- автоматическая проверка наличия оснований для выдачи диплома (данные из ИСУ, бухгалтерии, кафедр, библиотеки, общежития);
- интеллектуальный поиск нормативных документов и регламентов;
- автоматическое выявление академических и финансовых задолженностей;
- формирование проекта приказа с указанием ссылок на источники данных;
- проверка корректности и повторная верификация данных;
- автоматическая передача приказа в систему электронного документооборота;
- ведение журнала решений и источников данных
- формирование документа об образовании и приложения к нему.

Реализация указанных функций позволит обеспечить полную автоматизацию процессов подготовки, проверки и оформления документов об образовании.

5. Информационное и алгоритмическое обеспечение

Информационная модель (Рис. 2) основана на онтологическом подходе и описывает ключевые сущности системы: Студент, Приказ, Диплом, Приложение, Задолженность, Документ, Источник и Шаблон документа. Между объектами определены семантические связи, формализованные в формате RDF, что обеспечивает интеграцию данных из различных информационных систем вуза — учебной, бухгалтерской, библиотечной и общежитской подсистем. Обработка запросов и поиск взаимосвязей осуществляются с использованием SPARQL, что позволяет выполнять контекстный анализ данных и прослеживать происхождение информации (трассируемость решений).

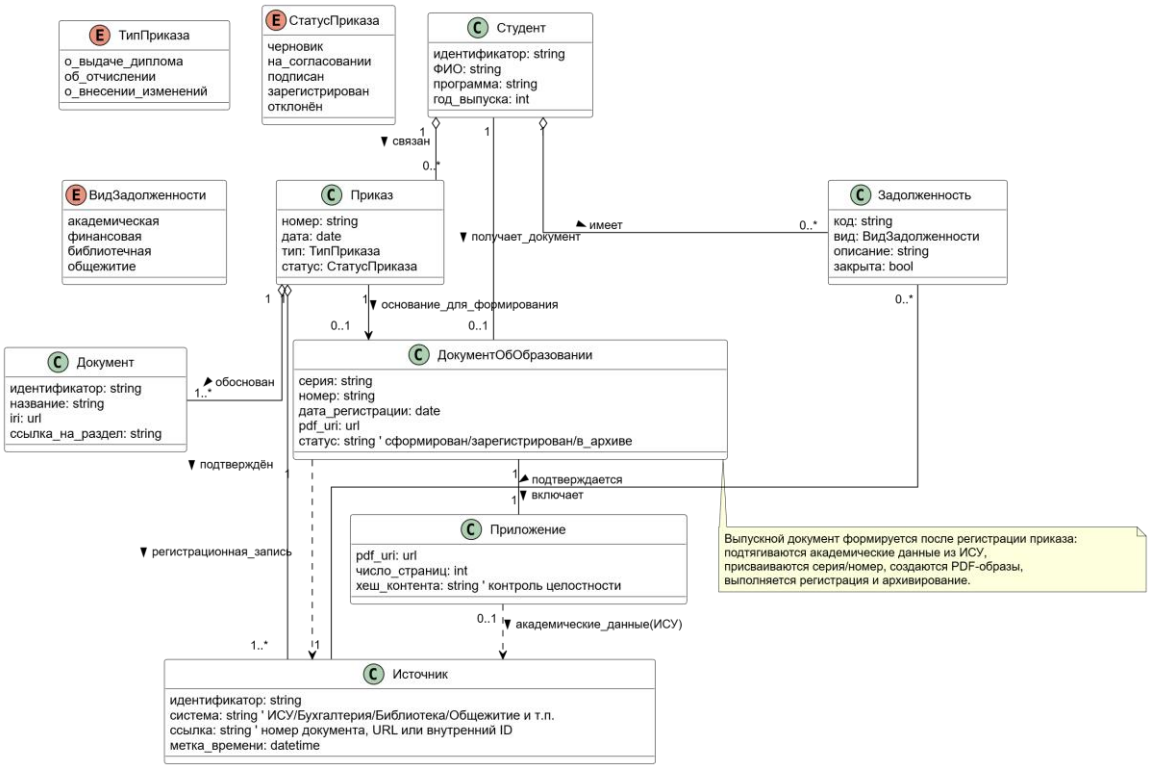


Рис. 2 - Информационная модель

Данная модель обеспечивает интеграцию данных из разнородных источников и трассируемость принятых решений.

Автоматизированный процесс формирования приказа и документов об образовании начинается с запроса на формирование диплома (Рис.). После получения запроса происходит извлечение данных о студенте из информационной системы университета, включая данные об успеваемости и наличии задолженностей. При отсутствии задолженностей выполняется обращение к семантическому хранилищу нормативно-правовых документов и языковой модели DeepSeek, которая на основе полученных данных формирует проект приказа о выдаче диплома.

После утверждения приказа в системе электронного документооборота осуществляется автоматическая генерация диплома и приложения на основании шаблонов, интегрированных с ИС университета.

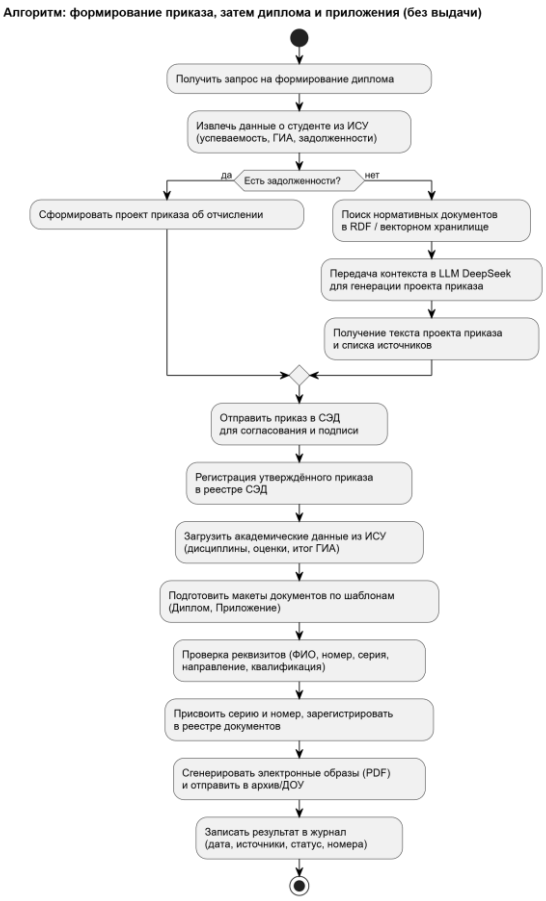


Рис. 3 - Алгоритм формирования документов о высшем образовании

Сформированный документ направляется в систему электронного документооборота для регистрации, а все действия фиксируются в журнале для последующего аудита и занесения документов в архив.

6. Архитектура программного обеспечения

Архитектура модуля (Рис.) построена по трёхуровневой схеме и включает уровень представления (веб-интерфейс сотрудника), прикладной уровень (сервер приложений, оркестрация RAG-процесса) и уровень данных/интеграций (ИСУ/СУБД, RDF/SPARQL, векторное хранилище и LLM). На уровне представления располагается веб-интерфейс сотрудника, обеспечивающий формирование запросов и получение результатов обработки.
Прикладной уровень реализует сервер приложений, включающий оркестрацию RAG-процесса, сервис формирования приказов, модуль генерации диплома и приложения, а также компоненты валидации и аудита.
На уровне данных и интеграций функционируют внешние подсистемы — информационная система университета (ИСУ/СУБД), RDF-хранилище, векторное хранилище и LLM.

Процесс взаимодействия реализуется следующим образом: пользователь инициирует оформление через веб-интерфейс, сервер приложений агрегирует сведения из ИСУ, выполняет семантический поиск регламентов (RDF + векторный индекс), формирует контекст и обращается к языковой модели для генерации проекта приказа. Сформированный документ передаётся в систему электронного документооборота для регистрации и подписи, а все действия протоколируются в журнале аудита.

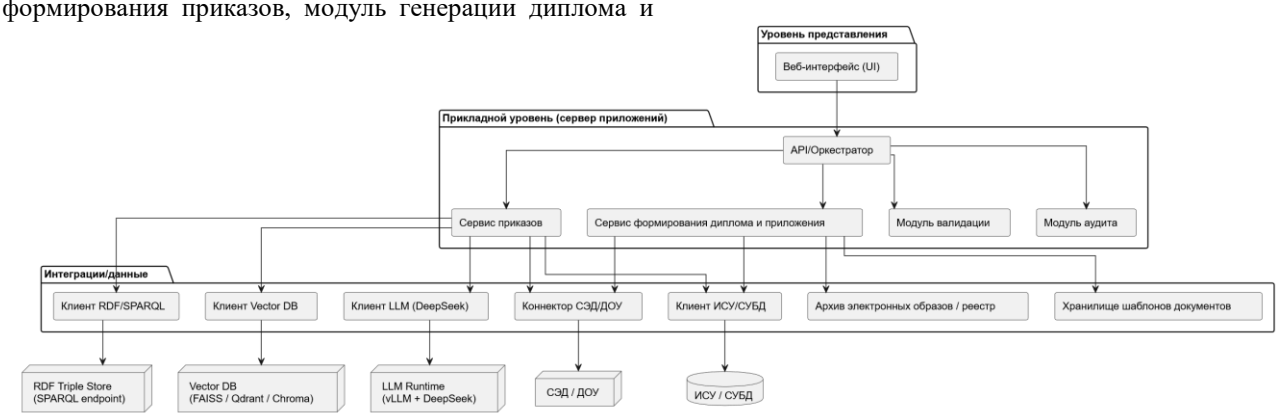


Рис. 4 - Архитектура программного обеспечения модуля (компонентная)

Такое разделение обеспечивает масштабируемость, изоляцию ответственности и соответствие требованиям защиты персональных данных за счёт локальной обработки и чётких границ интеграции.

7. Заключение

В статье предложен подход к автоматизации процесса формирования документов об образовании с использованием интеллектуальных технологий. Реализованный модуль опирается на сочетание семантического поиска в RDF-хранилище и генерации текстов на основе больших языковых моделей, что обеспечивает возможность интеллектуальной обработки нормативной информации и формирование обоснованных решений.

Разработанная архитектура отражает современный тренд на интеграцию искусственного интеллекта в процессы управления вузом. Использование RAG-подхода (Retrieval-Augmented Generation) и онтологической модели данных позволяет повысить прозрачность, точность и воспроизводимость решений, а также сократить трудоёмкость подготовки выпускной документации.

Представленные решения могут быть использованы при модернизации существующих информационных систем вузов, а также при построении комплексных платформ цифрового управления образовательным процессом. Дальнейшее развитие работы связано с расширением функционала модуля, внедрением механизмов адаптивного обучения моделей и интеграцией с внешними сервисами в рамках государственной стратегии цифровой трансформации образования.

Список используемых источников

1. Распоряжение Правительства Российской Федерации от 16 марта 2024 г. № 637-р «Об утверждении стратегического направления в области цифровой трансформации государственного управления» [Электронный ресурс]. — URL: https://www.consultant.ru/document/cons_doc_LAW_474269/ (дата обращения:

06.10.2025). (нормативно-правовой документ).

2. Министерство науки и высшего образования Российской Федерации. Стратегия цифровой трансформации науки и высшего образования [Электронный ресурс]. — URL: https://digital.gov.ru/ru/activity/directions/1064/ (дата обращения: 06.10.2025). (электронный ресурс).
3. Указ Президента Российской Федерации от 10 октября 2019 г. № 490 «О развитии искусственного интеллекта в Российской Федерации» [Электронный ресурс]. — URL: http://publication.pravo.gov.ru/Document/View/0001201910100032 (дата обращения: 06.10.2025). (нормативно-правовой документ).
4. Казанский (Приволжский) федеральный университет. Информационная система вуза edu.kpfu.ru [Электронный ресурс]. — URL: https://edu.kpfu.ru/ (дата обращения: 06.10.2025). (электронный ресурс).
5. Нижегородский государственный университет им. Н. И. Лобачевского. Электронная образовательная среда e-learning.unn.ru [Электронный ресурс]. — URL: https://e-learning.unn.ru/ (дата обращения: 06.10.2025). (электронный ресурс).
6. Уфимский государственный нефтяной технический университет. Система дистанционного обучения do.rusoil.net [Электронный ресурс]. — URL: https://do.rusoil.net/ (дата обращения: 06.10.2025). (электронный ресурс).
7. Уфимский университет науки и технологий. Информационная система университета isu.uust.ru [Электронный ресурс]. — URL: https://isu.uust.ru/ (дата обращения: 06.10.2025). (электронный ресурс).
8. Климова А. В., Ковтуненко А. С., Парфенов Д. В. Формирование учебного плана с использованием технологии RDF-хранилищ // Системная инженерия и информационные технологии. 2022. Т. 4. № 2(9). С. 49-55. EDN QMYQDT.
9. Ковтуненко А.В., Мустафин В.М. Информационная поддержка процесса проектирования образовательных программ в вузе с использованием онтологической модели данных. // В сборнике: Информационные технологии интеллектуальной поддержки принятия решений (памяти проф. Н.И. Юсуповой) ITIDS'2024. Труды X Международной научной конференции. В 2-х томах. Уфа, 2024. С. 57-62. EDN FGYCNE.
10. Ковтуненко А.В., Лактионов А.Е., Орашанов М.С., Ковтуненко А.С. STUDY CRATE // Свидетельство о регистрации программы для ЭВМ RU 2024663726, 10.06.2024. Заявка № 2024661706 от 24.05.2024. EDN TLSMHG.
11. Зорин, В. Н. Средства построения онтологий и графов знаний с помощью LLM / В. Н. Зорин // Современные информационные технологии. – 2025. – № 41(41). – С. 6-10. – EDN OTGNRC.
12. DeepSeek AI. API Documentation — DeepSeek R1 [Электронный ресурс]. —URL: https://api-docs.deepseek.com/ (дата обращения: 06.10.2025). (электронный ресурс).
13. deepseek-ai / DeepSeek-V3 [Электронный ресурс]. — GitHub.— URL: https://github.com/deepseek-ai/DeepSeek-V3?utm_source=chatgpt.com (дата обращения: 06.10.2025). (электронный ресурс).
14. Best open source LLMs in 2025 / Koyeb Blog [Электронный ресурс]. — URL: https://www.koyeb.com/blog/best-open-source-llms-in-2025?utm_source=chatgpt.com (дата обращения: 06.10.2025). (электронный ресурс).
15. Google. Gemma 2 / blog.google [Электронный ресурс]. — URL: https://blog.google/technology/developer s/google-gemma-2/?utm_source=chatgpt.com (дата обращения: 06.10.2025). (электронный ресурс).

Система поддержки принятия врачебных решений прогнозирования уровня риска сердечно-сосудистых заболеваний на основе взвешенных нечетких правил

Махметова К. М.
Институт математики
и информационных технологий
Оренбургский государственный
университет
Россия, Оренбург
e-mail: maksena00@gmail.com

Боровский А. С.
Аэрокосмический институт
Оренбургский
государственный университет
Россия, Оренбург
e-mail: borovski@mail.ru

Аннотация[1]

В статье рассматривается разработка системы поддержки принятия врачебных решений, основанной на нечеткой логике, для прогнозирования риска сердечно-сосудистых заболеваний. Предложенная система, использующая взвешенные нечеткие правила, позволяет учитывать неопределенность в медицинских данных и предоставляет врачам поддержку при принятии решений. Экспериментальные результаты показали, что система эффективно прогнозирует наличие или отсутствие сердечно-сосудистого заболевания. Предполагается, что система может повысить точность диагностики и улучшить качество лечения пациентов с сердечно-сосудистыми заболеваниями, особенно в условиях ограниченной информации.

Ключевые слова: нечеткая логика; система поддержки принятия врачебных решений; прогнозирование; риск сердечно-сосудистого заболевания; взвешенные нечеткие правила.

1. Введение

В последние десятилетия сердечно-сосудистые заболевания (далее – ССЗ) продолжают оставаться одной из ведущих причин смертности в мире. Согласно данным Всемирной организации здравоохранения [1], ССЗ ежегодно уносят жизни 17,9 миллионов человек.

Статистические данные показывают, что за последние 30 лет число случаев ССЗ в мире практически удвоилось – с 271 миллиона до 523 миллионов. В ближайшие десятилетия ситуация будет только ухудшаться. Прогнозируется, что в период с 2025 по 2050 год распространенность ССЗ увеличится на 90%, а общая смертность вырастет на 73%. Ожидается, что в 2050 году от ССЗ погибнет 35,6 миллиона человек, что значительно превышает 20,5 миллиона смертей в 2025 году.

Таким образом, приведенные данные подчеркивают необходимость в разработке и внедрении эффективных методов прогнозирования и ранней диагностики сердечно-сосудистых заболеваний. Стоит отметить, что на сегодняшний день традиционные методы диагностики часто не обеспечивают необходимой точности, что стимулирует разработку более современных инструментов. Одним из возможных вариантов решения данной проблемы является создание систем поддержки принятия врачебных решений [2-4].

2. Анализ похожих работ

Многие ученые из России и зарубежных стран представили ряд исследований по разработке систем поддержки принятия врачебных решений с использованием искусственного интеллекта и методов интеллектуального анализа данных. Среди них есть ряд работ, направленных на помощь в прогнозировании риска сердечно-сосудистых заболеваний. Ниже приведен обзор наиболее значимых трудов.

Так, в исследовании Т.Г. Вайханской [5] рассматривается применение нечеткой логики для прогнозирования неблагоприятных сценариев заболеваний сердечно-сосудистой системы. Предложенный подход направлен на создание системы, которая бы использовала экспертные знания и медицинские данные, и при этом могла бы адаптироваться к различным ситуациям и изменениям в данных, что позволяет улучшить качество диагностики и прогнозирования, особенно

[1] Труды XI Международной научной конференции "Информационные технологии интеллектуальной поддержки принятия решений", 11-13 ноября, Уфа-Чандигарх-Фандербейлпарк, 2025

в тех случаях, когда данные не могут быть точными или полными.

В исследовании [6] авторы Т.А. Джилани, Х. Ясин, М. Ясин, К. Ардиль применили методы интеллектуального анализа данных для выявления факторов, оказывающих влияние на риск развития острого коронарного синдрома. В качестве зависимой переменной был выбран диагноз, принимающий бинарные значения, отражающие наличие или отсутствие заболевания.

Экспериментальная база данных была сформирована на основе информации, полученной из двух кардиологических больниц города Карачи (Пакистан). Всего в выборке было представлено шестнадцать переменных, из которых одна рассматривалась как зависимая, а остальные пятнадцать – как независимые. Для уменьшения размерности данных и повышения точности и надежности модели прогнозирования был применен метод главных компонент. В результате процедуры отбора признаков для дальнейшего анализа были использованы 14 наиболее информативных факторов. Для оценки влияния независимых факторов на зависимую переменную использовалась множественная регрессия.

С.Б. Патиль и Ю.С. Кумарасвами [7] предложили нейросетевую систему прогнозирования сердечных приступов. Данные из базы заболеваний сердца предварительно обрабатывались и кластеризовывались с помощью K-средних для выделения групп, связанных с инфарктом. Затем алгоритм MAFIA выявлял часто встречающиеся паттерны, из которых отбирались наиболее значимые для прогнозирования. Обучение многослойного перцептрона осуществлялось методом обратного распространения ошибки.

Анализ проведенных исследований показывает, что в области разработки систем поддержки принятия врачебных решений существует широкий спектр подходов. Однако большинство решений ограничены малыми наборами данных и недостаточно адаптивны к их неполноте или изменчивости, что подчеркивает необходимость разработки более гибкой системы, способной интегрировать разнообразные источники информации и повысить точность прогнозирования.

3. Описание набора данных с заболеваниями сердца

3.1. Специфика сердечно-сосудистых заболеваний

Сердечно-сосудистые заболевания – это группа заболеваний, поражающих сердце и сосудистую систему. Этиология ССЗ многогранна и включает как генетические, так и экологические факторы. Одним из механизмов развития многих сердечно-сосудистых заболеваний является атеросклероз, характеризующийся отложением липидов и образованием атеросклеротических бляшек в стенках сосудов, которые ведут к нарушению нормального кровообращения, что, в свою очередь, способствует развитию ишемической болезни сердца и других заболеваний.

К основным факторам риска относят артериальную гипертензию, дислипидемию, сахарный диабет, ожирение, курение и малую физическую активность. Такие факторы способствуют ускорению атеросклеротических процессов и ухудшают эндотелиальную функцию, что способствует снижению сосудистого тонуса и увеличивает вероятность тромбообразования. Важным аспектом является влияние воспалительных процессов на развитие ССЗ, где повышение уровня воспалительных маркеров, таких как С-реактивный белок, может служить предсказателем сердечно-сосудистых осложнений [8].

Клинические проявления ССЗ зависят от вида заболевания. Ишемическая болезнь сердца часто проявляется стенокардией, болями в груди, одышкой, а также сердечной недостаточностью. Инсульт, как острое нарушение мозгового кровообращения, может привести к параличу, нарушениям речи и координации. Хронические заболевания, такие как гипертония, могут долгое время не проявляться симптоматически, но со временем приводят к повреждению сосудов и сердца.

Диагностика ССЗ основывается на анамнезе, физикальном обследовании, лабораторных тестах (определение уровня холестерина, сахара в крови, С-реактивного белка) и инструментальных исследованиях (электрокардиография, эхокардиография, коронарная ангиография и магнитно-резонансная томография).

Лечение ССЗ направлено на замедление прогрессирования заболевания и предотвращение осложнений с помощью медикаментозной терапии (антигипертензивные препараты, статины, антиагреганты и антикоагулянты), а также хирургического вмешательства (коронарное шунтирование или стентирование). Важными аспектами лечения являются коррекция образа жизни (питание, физическая активность, отказ от курения) и контроль факторов риска (гипертония и дислипидемия) [9].

Профилактика ССЗ включает первичную профилактику, направленную на предотвращение возникновения заболевания (здоровый образ жизни, контроль за артериальным давлением и уровнем холестерина) и вторичную профилактику, которая направлена на предотвращение обострений и рецидивов у пациентов с уже установленными диагнозами ССЗ.

3.2. Набор данных

Кливлендский набор данных о заболеваниях сердца является одним из наиболее популярных в области машинного обучения и медицины для задачи предсказания риска сердечно-сосудистых заболеваний [10]. Он был собран в 1980 году и используется для анализа факторов риска, предсказывающих наличие или отсутствие сердечных заболеваний у пациентов. Содержит 303 экземпляра (пациентов) и 14 атрибутов (характеристик), включая метки о наличии заболевания. К основным атрибутам можно отнести:

1) Age – возраст пациента;

2) Sex – пол пациента;

3) Cp – тип боли в груди:

4) Trestbps – кровяное давление в покое;

5) Chol – уровень холестерина в крови;

6) Fbs – уровень сахара в крови натощак;

7) Restecg – результат электрокардиограммы в покое;

8) Thalach – максимальный пульс при физической нагрузке;

9) Exang – наличие стенокардии при физической нагрузке;

10) Oldpeak – депрессия сегмента ST;

11) Slope – наклон сегмента ST;

12) Ca – количество крупных сосудов, отображаемых на флюороскопии;

13) Thal – состояние талассемии;

14) Target – наличие сердечного заболевания.

Распределение по классам: у 54% отсутствуют сердечные заболевания, у 46% – присутствуют.

4. Предлагаемая система поддержки принятия врачебных решений на основе взвешенных нечетких правил

Для более точного прогнозирования уровня риска используется нечеткая логика, где решение, принимаемое по заболеваниям сердца у пациентов, опирается на взвешенные нечеткие правила, от которых и зависит эффективность работы системы. Как правило, эксперты в соответствующей области предоставляют нечеткие правила для решения задачи прогнозирования. Однако, здесь они генерируются автоматически на основе исторических данных. Кроме того, нечеткие правила взвешиваются в зависимости от их важности с использованием весов атрибутов. Архитектура предложенной системы поддержки принятия врачебных решений на основе нечеткой логики показана на рисунках 1 и 2.

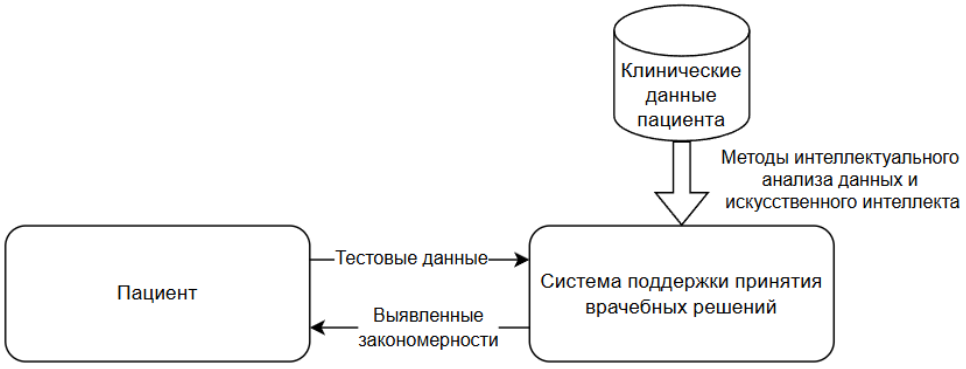


Рис. 1. Система поддержки принятия врачебных решений

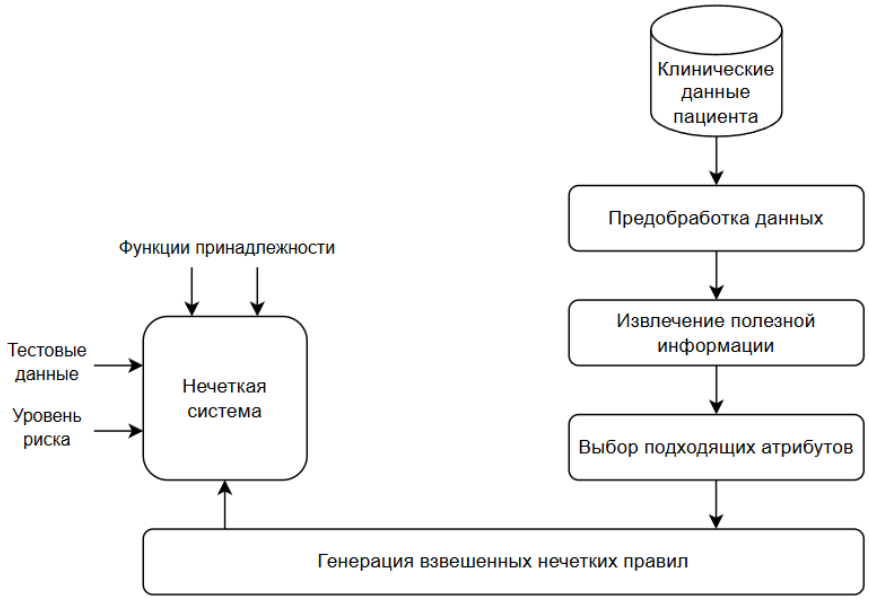


Рис. 2. Архитектура системы поддержки принятия врачебных решений на основе нечеткой логики

4.1. Предобработка данных

Цель предварительной обработки данных – извлечение полезной информации из сырых данных, а затем преобразование этих данных в формат, необходимый для прогнозирования. Из-за наличия нерелевантной информации в исходных данных, они не могут быть непосредственно использованы в процедуре предсказания, поэтому на этапе предварительной обработки данные должны быть очищены, проанализированы и преобразованы для дальнейших шагов. Каждая строка представляет собой информацию о пациенте, а столбец – это список атрибутов (биомаркеров). Последний столбец содержит метку класса, соответствующую уровню риска у пациентов с заболеваниями сердца.

4.2. Классификация обучающего набора данных по уровню риска

После обработки данных, обучающий набор классифицируется на два подмножества в зависимости от метки класса: отсутствие сердечного заболевания и наличие сердечного заболевания. Класс 0 обозначает отсутствие заболевания, а класс 1 – его наличие.

4.3. Автоматизированный подход для генерации взвешенных нечетких правил

В процессе прогнозирования риска развития сердечно-сосудистых заболеваний часто используется большое количество атрибутов, однако не все из них оказывают значительное влияние на точность прогноза. Для выбора наиболее информативных признаков применяется частотный анализ. На основании полученных частот и соответствующих весов автоматически генерируются нечеткие правила, составляющие основу будущей системы поддержки принятия врачебных решений [11].

На первом этапе из набора данных извлекаются часто встречающиеся категории атрибутов для каждого признака, что позволяет рассчитать частоту появления каждой категории в соответствующем классе. Для непрерывных признаков используется процедура дискретизации, преобразующая их в категориальные данные.

На следующем этапе осуществляется отбор наиболее значимых атрибутов для прогнозирования влияния на риск возникновения ССЗ. Для этого сравниваются минимальные и максимальные отклонения каждого признака, что позволяет выделить наиболее значимые характеристики и исключить менее важные.

Далее, на основе полученных векторов отклонений формируются нечеткие правила, описывающие влияние фактора на риск возникновения сердечного заболевания у пациента. Каждое правило взвешивается в зависимости от его значимости.

Создание таких правил на основе числовых и бинарных данных требует наглядности и простоты интерпретации. Для решения этой задачи используется теория нечетких множеств, которая позволяет строить нечеткие правила с использованием функций принадлежности.

Разработка системы поддержки принятия решений на базе нечеткой логики включает несколько этапов. На этапе фазификации входные данные преобразуются в нечеткие множества с помощью функций принадлежности, что позволяет интерпретировать данные в терминах нечеткой логики. Далее система применяет базу правил для получения нечетких выводов, которые затем подвергаются дефазификации и преобразуются в точные значения, используемые для принятия обоснованных решений и дальнейшего анализа рисков (рисунок 3).

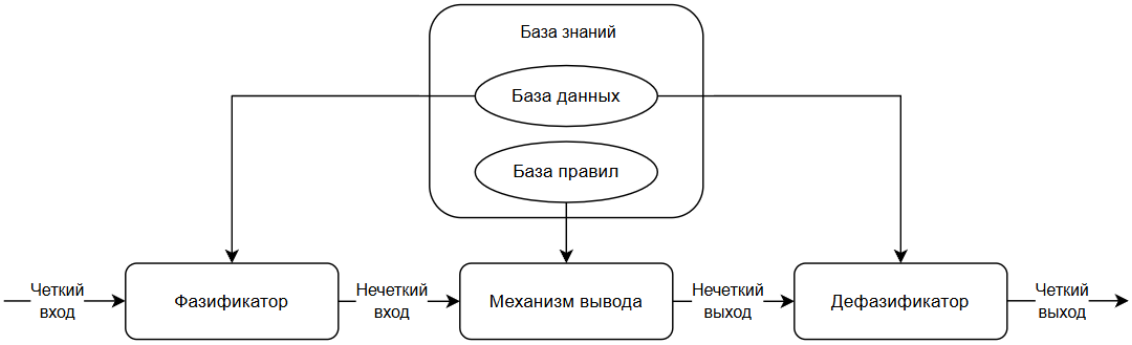


Рис. 3. Нечеткая система вывода

Предложенная система поддержки принятия врачебных решений, представленная на рисунке 4, включает в себя m входных параметров и один выход, где входы ассоциированы с m атрибутами, а выход представляет собой метку класса, отображающую уровень риска [12, 13]. В качестве основы для реализации данной системы используется модель нечеткого вывода Мамдани с дефаззификацией, выполненной с применением метода центроида области.

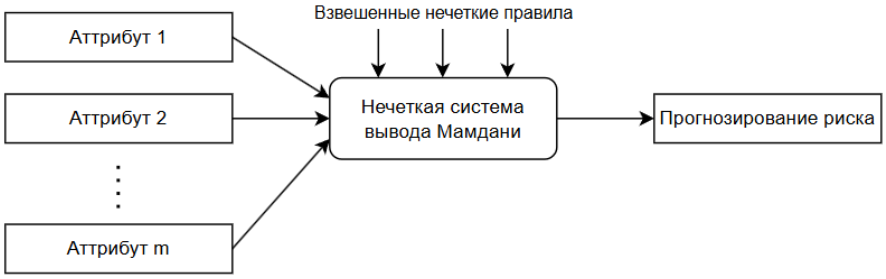


Рис. 4. Нечеткая система вывода на основе взвешенных нечетких правил

5. Полученные результаты

Для эксперимента набор данных был разделен на обучающую (202 экземпляра) и тестовую (101 экземпляр) выборки.

Обучающая выборка используется для генерации взвешенных нечетких правил, а тестовый выборка – для анализа производительности предложенной системы.

Выбранные атрибуты: Age, Trestbps, Chol, Thalach, Oldpeak, Thal.

Пример нечетких правил: «ЕСЛИ (артериальное давление в покое высокое), ТО (Класс – Class 1) (0,22438)», «ЕСЛИ (возраст очень низкий), ТО (Класс – Class 0) (0,14237)», «ЕСЛИ (уровень сахара в крови натощак средний), ТО (Класс – Class 1) (0,68943)».

В этих правилах плавающие значения указывают на вес полученных нечетких правил.

В таблице 1 приведены метрики оценки системы.

Таблица 1. – Значения метрик оценки системы

Класс	Метрика	Обучающий набор	Тестовый набор
0 (риск ниже 50%)	Точность	0,52453	0,65672
	Чувстви-тельность	0,75123	0,79678
	Специ-фичность	0,28547	0,48758
1 (риск выше 50%)	Точность	0,55068	0,68567
	Чувстви-тельность	0,30437	0,50437
	Специ-фичность	0,78475	0,81473

6. Заключение

В ходе исследования была рассмотрена проблема разработки систем поддержки принятия врачебных решений для прогнозирования риска сердечно-сосудистых заболеваний на основе методов нечеткой логики.

Экспериментальные исследования, выполненные на основе Кливлендского набора данных, показали, что предложенная система демонстрирует удовлетворительные результаты по показателям точности, чувствительности и специфичности, что

свидетельствует о возможности ее практического применения.

Предложенный подход может быть доработан путем добавления дополнительных факторов риска развития ССЗ и внедрения более сложных гибридных моделей, что позволит повысить точность прогнозирования и обеспечить эффективную работу системы с крупными медицинскими массивами данных.

Таким образом, предложенная система поддержки принятия врачебных решений может рассматриваться как перспективный инструмент для здравоохранения, способствующий снижению уровня смертности от сердечно-сосудистых заболеваний за счет своевременного выявления пациентов с повышенным риском.

Список используемых источников

1. World health statistics 2025 : World Health Organization. URL: https://data.who.int/ (дата обращения: 13.08.2025).
2. Novoa K.C., Dunn T., Curry A., Froude R., Simpson S.A. Limitations of Traditional Models for Medical Decision-Making Capacity and Ethical Clinical Practice in Light of the SARS-CoV-2 Pandemic // Cureus. 2021. Vol. 13. № 4. P. 27-31.
3. Molloy M., Hagedorn P., Dewan M. Why Does Current Clinical Decision Support Frequently Fail to Support Clinical Decisions? // Pediatr Crit Care Med. 2022. Vol. 23. № 8. P. 670-672.
4. Castaneda C., Nalley K., Mannion C., Bhattacharyya P., Blake P., Pecora A., Goy A. Clinical decision support systems for improving diagnostic accuracy and achieving precision medicine // Journal of Clinical Bioinformatics. 2015. Vol. 5. P. 4.
5. Frolov A.V., Martsenyuk M.A., Vaikhanskaya T., Poljakov V.B. Fuzzy logic in prediction of adverse scenarios of the cardiac diseases // Klinical Informatics and Telemedicine. 2017. Vol. 12. P. 35-41.
6. Jilani T.A. et al. Acute coronary syndrome prediction using data mining techniques-an application // World Academy of Science, Engineering and Technology. 2009. Vol. 59. № 4. P. 295-299.
7. Patil S.B., Kumaraswamy Y.S. Intelligent and effective heart attack prediction system using data mining and artificial neural network // European Journal of Scientific Research. 2009. Vol. 31. № 4. P. 642-656.
8. Богачевская С.А., Киселев С.Н. Анализ сердечно-сосудистой смертности в России и ДФО в рамках реализации федеральных и региональных программ «Борьба с сердечно-сосудистыми заболеваниями»: целевые показатели, прогноз и факты / С.А. Богачевская, С.Н. Киселев // Дальневосточный медицинский журнал. – 2024. - № 1. – С. 44-50.
9. Каушан К.С., Титаренко О.П., Серцова А.М., Пешкова Н.Г. Сердечно-сосудистые заболевания как фактор снижения трудоспособности и продолжительности жизни: современные тенденции и пути преодоления / К.С. Каушан, О.П. Титаренко, А.М. Серцова, Н.Г. Пешкова // Биология и интегративная медицина. – 2025. – № 2 (74). – С. 303-319.
10. Heart Disease Cleveland UCI : Kaggle. URL: https://www.kaggle.com/datasets/cherngs/heart-disease-cleveland-uci/ (дата обращения: 22.08.2025).
11. Миловидова А.А. Формирование нечеткой системы оценки для системы поддержки принятия решений при управлении процессом переработки в условиях неопределенности информации о качестве материала / А.А. Миловидова // Прикаспийский журнал: управление и высокие технологии. – 2020. – № 1 (49). – С. 33-48.
12. Шабанов К.Б., Алексеев В.В. Применение методов интеллектуального анализа данных для повышения качества принятия решений при управлении ресурсами информационной медиасистемы / К.Б. Шабанов, В.В. Алексеев // Вестник Тамбовского государственного технического университета. – 2021. – Т. 27, № 1. – С. 14-19.
13. Вертакова Ю.В., Шульгина Ю.В. Применение аппарата нечеткой логики при принятии кадровых решений / Ю.В. Вертакова, Ю.В. Шульгина // Теория и практика сервиса: экономика, социальная сфера, технологии. – 2021. – № 3 (49). – С. 9-12.

Финансовая модель образовательной программы как основа для принятия решений о реализации

Зубарев А. Е.
Тихоокеанский государственный университет, институт экономики и управления
Хабаровск, Россия
e-mail: 000008@togudv.ru

Кулик И. В.
Тихоокеанский государственный университет, институт экономики и управления
Хабаровск, Россия
e-mail: 003187@togudv.ru

Аннотация[1]

В статье обсуждаются вопросы построения финансовых моделей основных образовательных программ (далее ООП) в университетах Российской Федерации. Среди вопросов, которые поднимаются авторами, формирование моделей расходов и доходов ООП, построение на их основе модели финансового результата ООП, определение управляемых факторов, способных изменять совокупные доходы и/или расходы ООП, тестирование (апробация) прогностических способностей построенных моделей доходов и расходов. При моделировании учтены факторы инфляции, приведения контингента разных форм обучения, доходы от разных категорий обучающихся. Практика использования полученных финансовых моделей ООП в ФГБОУ ВО «Тихоокеанский государственный университет» (далее ТОГУ) позволяет принимать взвешенные решения относительно возможности открытия новых ООП, продолжения в неизменном виде реализации существующих ООП, закрытия или существенного переформатирования нерентабельных ООП, а также ряд других сопутствующих вопросов, связанных с функционированием вуза как самостоятельной финансово-хозяйственной системы в условиях конкурентного рынка образовательных услуг высшего образования.

Ключевые слова: финансовый результат, прибыль, убыток, рентабельность, математическая модель, переменные расходы, постоянные расходы, порог рентабельности, точка безубыточности, образовательная организация, основная образовательная программа.

1. Введение

Современное состояние высшего образование предполагает использование разнообразных источников финансирования основной деятельности университетов. Тем не менее при детальном рассмотрении образовательной деятельности следует выделить в качестве основных источников финансовых поступлений бюджетные ассигнования, выделяемые из федерального бюджета в виде госзадания (дотации) на подготовку студентов-бюджетников, и внебюджетные поступления, аккумулируемые вузами от поступлений по договорам с юридическими и физическими лицами за подготовку студентов, обучающихся на платной основе (студенты-договорники). Общая сумма поступлений характеризует совокупные финансовые притоки (доходы) университета от основной образовательной деятельности. С другой стороны университет, как хозяйственная система генерирует поток финансовых оттоков в виде разнообразных текущих затрат для успешной реализации образовательных программ. Очевидно, что сальдо поступлений и оттоков будет говорить о степени «финансовой успешности» образовательной деятельности университета. При положительном сальдо или положительном финансовом результате в коммерческой организации говорят о наличии прибыли, при отрицательном – об убытке. В условиях бюджетных организаций это не совсем уместно (ибо нет основной цели – извлечения прибыли), однако аналогия вполне заслуживает внимания. Положительный финансовый результат образовательной деятельности вуза позволяет проводить модернизацию самих образовательных программ, имущественного комплекса, осуществлять дополнительные расходы для мотивации преподавательского состава и студентов, и, вообще реализовывать разные инвестиционные инициативы и проекты, направленные на

[1] Труды XI Международной научной конференции "Информационные технологии интеллектуальной поддержки принятия решений", 11-13 ноября, Уфа-Чандигарх-Фандербейлпарк, 2025

повышение как образовательно-воспитательной, так и научной деятельности университета. Отрицательный финансовый результат образовательной деятельности вуза «убыток», очевидно не позволяет сделать ничего из вышеперечисленного, а, наоборот, требует изыскания финансовых средств для покрытия недостатка в возмещении каких-то обязательных текущих расходов. Проблема принятия решений об открытии новых образовательных программ, а также проблема анализа эффективности уже реализуемых образовательных программ считается авторами актуальной и насущной для многих образовательных организаций в России.

Учитывая, что основу образовательной деятельности в вузах РФ составляют основные образовательные программы (далее ООП), реализующие четыре уровня направлений подготовки (бакалавриат, специалитет, магистратура, аспирантура), авторы работы сосредоточили внимание именно на них, исключив (в данной работе) из рассмотрения программы дополнительно профессионального образования. При анализе последних могут быть использованы приёмы, изложенные авторами для ООП, учитывая ряд особенностей, не рассматриваемых в этой работе.

В предлагаемом исследовании описана апробированная в Тихоокеанском государственном университете финансовая модель основной образовательной программы университета, позволяющая проводить финансовый анализ комплекса показателей для принятия решений в области целесообразности и возможности её реализации.

2. Цели и задачи

Гипотеза работы сформулирована нами в виде тезиса о возможности использования полученной авторами математической модели финансового результата ООП для принятия решений:

- относительно существующей ООП в части продолжения её реализации (с изменением или без) или в части закрытия ООП;

- относительно новой (планируемой к открытию) ООП, собственно, о её открытии (как есть) либо об открытии, но с доработкой в отношении снижения состава расходов (или роста поступлений), либо об отказе в открытии при невозможности доработать, либо о принятии решения об открытии убыточной ООП с погашением убытков за счёт рентабельных ООП.

Цель исследования – получение финансового инструмента для принятия управленческих решений руководством университета в отношении целесообразности реализации ООП.

Для достижения указанной цели в исследовании сформулированы задачи:

1. обосновать обобщённую модель бюджета университета,
2. выделить среди совокупности поступлений часть, относящуюся к реализации ООП вуза,
3. определить классификацию расходов на ООП, с выделение расходов, находящихся в прямой зависимости от численности обучающихся,
4. предложить математические и графические модели доходов и расходов ООП,
5. построить модель финансового результата ООП,
6. провести анализ чувствительности модели финансового результат ООП к изменению управляемых факторов.

3. Материалы и методы

Основным методом, применяемым авторами в исследовании, является CVP-метод (Costs – Volume – Profit), отслеживающий зависимость финансового результата образовательной услуги от издержек и объёмов реализации услуги (численности контингента).

Хотя моделирование может быть изначально проведено с разным целеполаганием, основными задачами построения финансовых моделей по мнению исследования [1] являются: анализ событий, сценарный анализ, анализ чувствительности, анализ точки безубыточности, план-факторный анализ, базовый финансовый анализ деятельности. Практически все эти задачи способны решать построенная нами и предлагаемая в статье финансовая модель ООП.

Среди проблем финансирования высшего образования можно выделить работу [2], где автор классифицирует и анализирует присутствующие в секторе высшего образования организации на два вида: государственные и негосударственные. По нашему мнению, для построения финансовой модели ООП принципиальных отличий в структурном описании первых и вторых организаций, нет. Отличия с точки зрения учредителя безусловно важны и даже имеют первостепенное значение, поскольку преследуют разное целеполагание. В качестве целей государственных организаций в сфере образования лежит удовлетворение общественной потребности в производстве и потреблении образовательных услуг, в то время как для учредителя негосударственного образовательного учреждения (коммерческой организации) целью является извлечение прибыли. В этой связи может показаться, что для негосударственной образовательной организации проблема финансового моделирования более актуальна, поскольку позволяет в полной мере учесть факторы внутренней и внешней среды для определения точки безубыточности деятельности в условиях рынка образовательных услуг.

Однако, как указано ранее во введении, для государственных образовательных организаций построение финансовой модели ООП не менее важно. В текущем исследовании не делается различий между этими видами образовательных организаций. Оба вида считаются одинаковыми, и функционирующими по общим экономическим законам, типичным для самостоятельных хозяйствующих субъектов рыночных отношений.

Многие авторы научных работ в сфере экономики высшего образования изучают правовую и экономическую сущность образовательной услуги, как объекта соответствующего рынка. Подлежат изучению бюджетные основы финансирования деятельности образовательных учреждений, даётся анализ зарубежного опыта использования финансовых механизмов управления организациями высшего образования [3, 4, 5]. Практически не встречаются исследования в сфере оценки финансовой эффективности основного образовательного процесса в организациях ВО, не выявлено исследований рентабельности отдельных образовательных программ (основных или дополнительных), методик определения точки безубыточности, порога рентабельности (рентабельного контингента), операционного рычага или иных финансовых метрик ООП образовательных организаций.

Интересный подход к формированию сбалансированных расходов на ООП авторов исследований [6, 7] опирается на нормативные расходы по финансированию ООП. При этом выделяется только часть затрат по реализации ООП, приходящаяся на фонд оплаты труда. Методика интересна тем, что авторы демонстрируют механизм, мотивирующий центры финансовой ответственности (кафедры) не «раздувать» нагрузку и штаты, а оптимизировать выделяемые фонды на оплату труда для целей обоснованного и эффективного стимулирования работников своих подразделений. Однако, авторы методики [6, 7] так и не доходят до полной калькуляции себестоимости, хотя пытаются привести расчёты финансовой эффективности образовательной программы, используя исключительно прямые расходы по реализации ООП, представленные фондами оплаты труда. Вопрос определения и распределения накладных расходов, связанных с реализацией ООП, в этих исследованиях не рассматривается.

Укажем работу [7], в которой автор поднимает вопрос финансового управленческого учёта затрат на образовательные программы организации. Автор показывает, что для принятия правильных решений в отношении продолжения реализации образовательных программ целесообразно использовать популярную в производстве систему финансового учёта неполной себестоимости «директ-костинг». В рамках предлагаемого маржинального анализа безубыточности ООП, работа [7] интересна обсуждением классификации затрат вуза с выделением постоянных и переменных расходов.

В данной работе авторы предлагают авторские методические подходы к формированию моделей: притока финансовых средств от реализации ООП; оттоков финансовых средств в виде расходов на реализацию ООП; финансового результата (прибыли/убытка) ООП.
Полученные модели успешно используются в ТОГУ в рамках ежегодной приёмной кампании при принятии решений о численности минимального необходимого контингента студентов первого курса ООП, обеспечивающего достижение порога рентабельности ООП.
К авторскому методу, выносимому в дальнейшем на обсуждение, относится «проецирование» информации о структуре доходов вуза на обобщённые данные о совокупных расходах вуза, для получения пропорциональной структуры расходов ООП. Этот подход в условиях отсутствия информации о постатейном управленческом учёте затрат на ООП позволяет относительно точно и быстро получить картину о структуре расходов ООП, описанную далее в п. 4.2. Этот подход позволяет также получить полную себестоимость образовательной услуги ВО.
В предварительных исследованиями нами сформулировано следующее определение ООП [8]: «**ООП – это образовательная программа высшего образования с общим направлением подготовки (специальностью) и общей направленностью (профилем / специализацией), учитывающей все курсы обучения и все формы обучения (очную, очно-заочную и заочную).**» Приведём, примеры ООП ТОГУ согласно данному определению:

- 38.03.01 Экономика, профиль финансы и кредит;
- 15.04.02 Технологические машины и оборудование, направленность оборудование нефтегазопереработки;
- 40.05.01 Правовое обеспечение национальной безопасности, специализация государственно-правовая;
- 5.8.1 Педагогика, направленность общая педагогика, история педагогики и образования.

Укажем, что приведение контингента (а значит и затрат) в разрезе форм обучения производится с помощью коэффициентов приведения численности МИНОБРНАУКИ в соотношении: один студент очной формы - один студент очно-заочной формы - один студент заочной формы = 1 - 0,25 - 0,01. При расчёте поступлений по ООП важно учитывать численность студентов в разрезе форм обучения и курсов обучения одновременно. Это связано с колебаниями цен за обучение для разных лет набора и разных форм обучения одной и той же ООП, при

одновременной ежегодной индексации на уровень инфляции.

Практика использования полученных финансовых инструментов для принятия решений о реализации ООП показала возможности успешного простого масштабирования объекта исследования. Под масштабированием мы понимаем, например, возможности объединения всех ООП в единую УГСН (по своего рода общей отраслевой привязке выпускников) для сопоставления рентабельностей ООП разных профессий («масштабирование вверх»). Уместно и обратное «масштабирование вниз» с выделением для целей анализа, например, только одну форму обучения для сравнения между собой рентабельности одной ООП в разрезе разных форм обучения.

4. Результаты

Характеристика финансовой модели ООП сводится к её выделению из общей финансовой модели университета (рис. 1). Слева на рис. 1 показаны притоки (доходы), в центре – некий механизм бюджетирования (перераспределения поступлений) и справа оттоки (расходы) университета, классифицированные по разным основаниям для целей исследования.

К источникам финансовых поступлений вузов (согласно уставной деятельности) относятся: доходы от реализации образовательных программ (основных и дополнительных); доходы от НИОКР; доходы от иной доход приносящей деятельности (далее ДПД).

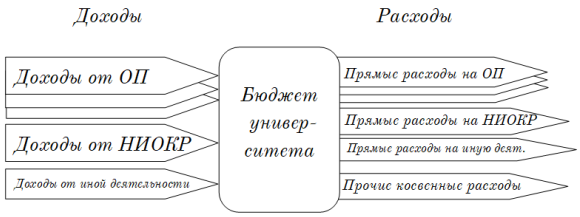


Рис. 1. Финансовая модель университета

Среди расходов для целей макроанализа финансовой модели вуза мы выделяем прямые расходы на образовательные программы, прямые расходы на НИОКР, прямые расходы на иную ДПД, прочие (косвенные) расходы, относимые на все образовательные и иные продукты вуза пропорционально некоторых наиболее целесообразных и обоснованных баз.

Модель финансового результата ООП, выделенная из финансовой модели университета, может быть описана простой формулой

$$ФР_{ООП} = Д_{ООП} - Р_{ООП}, \quad (1)$$

где, $ФР_{ООП}$ – финансовый результат ООП, руб./год; $Д_{ООП}$ – доходы ООП, руб./год; $Р_{ООП}$ – расходы на реализацию ООП, руб./год

Анализируя структуру расходов университета рис. 1, можно отметить, что совокупные расходы на реализацию ООП (полная годовая себестоимость ООП) могут быть выражены уравнением

$$Р_{ООП} = П + К, \quad (2)$$

где $П$ – прямые расходы на реализацию ООП. В своей основной массе они представлены расходами на оплату труда преподавателей, привлечённых к реализации ООП. Однако среди них есть и иные виды расходов (материальные, оплата услуг и др.), «жёстко» связанные с конкретной ООП; $К$ – это часть прочих косвенных расходов университета, относимая на ООП для целей калькуляции полной себестоимости ООП.

Положительное или отрицательное значение $ФР_{ООП}$ будет свидетельствовать о рентабельной (как минимум самоокупаемой) или нерентабельной ООП. При расчёте относительного показателя «рентабельности продаж» по формуле

$$R = (ФР_{ООП} / Д_{ООП}) \cdot 100\%, \quad (3)$$

можно говорить об уровне относительной прибыльности каждой анализируемой ООП, как величине прибыли в составе совокупных доходов от реализации ООП (в %).

Для справки, доходы ТОГУ от реализации ООП составляли в 2024 году 76,36% всех доходов университета [9]. Общее количество реализуемых ООП в номенклатуре ТОГУ в 2024 году насчитывало 259 программ согласно определению ООП, данному в п. 3.

Следуя теории финансового менеджмента, если организация в целом рентабельна, но какая-либо продукция (из ассортимента) нерентабельна, то её убытки покрываются за счёт прибыли от другой (рентабельной) продукции. Такой убыточный продукт следует либо постепенно “выводить” из ассортимента (заменять рентабельными видами), либо “дорабатывать” его так, чтобы он стал рентабельным. Эти положения полностью применимы и к номенклатуре основных образовательных программ университета. Однако, применительно к классическому университету полный отказ от нерентабельных образовательных программ невозможен по причинам государственного заказа, имиджевой составляющей классического университета, но главное – вследствие экономических последствий от перераспределения “центра тяжести” условно-постоянных расходов. Радикальный отказ от убыточных ООП распределит всю сумму косвенных условно-постоянных расходов между оставшимися, реализуемыми ООП так, что в итоге некоторые из них могут превратиться из рентабельных в нерентабельные.

В ходе исследования, основываясь на принципах классического финансового менеджмента, авторами проведён маржинальный анализ состава и структуры доходов и расходов ООП, построены финансовые модели и проведён сценарный подход к

реализации ООП, посредством влияния на управляемые факторы ООП вуза.

4.1. Доходы ООП.

Вся совокупность поступлений по отдельной ООП складывается в настоящее время в университете из трёх основных источников:

- федеральный бюджет – в виде субсидии на образовательную деятельности в соответствии с контрольными цифрами приёма;
- внебюджетные источники в виде поступлений от оказания платных образовательных услуг для граждан РФ;
- внебюджетные источники в виде поступлений от оказания платных образовательных услуг для иностранных граждан;

Учитывая эти три источника, доходы основной образовательной программы в рамках одной формы обучения и одного курса представлены авторами исследования уравнением вида

$$Д_{ООП} = (Ц_{КЦП} \cdot Ч_{КЦП}) + (Ц_{РФ} \cdot Ч_{РФ}) + \\ + (Ц_{Ин} \cdot Ч_{Ин}), \quad (4)$$

где $Ц_{КЦП}$ – нормативные затраты подушевого финансирования учредителем контрольных цифр приёма (далее КЦП, бюджетные места), руб./(чел.·год);

$Ч_{КЦП}$ – численность обучающихся ООП в пределах КЦП, чел. Показатель отнесён к числу управляемых факторов, в части обеспечения сохранности и повышения рейтинга вуза в конкурсных процедурах для получения б*о*льших КЦП в очередном учебном году по направлениям подготовки. Например, в структуре доходов от реализации ООП в 2024 году в ТОГУ доля доходов из федерального бюджета составляла году 69,7%;

$Ц_{РФ}$ – “цена обучения” в год для российских абитуриентов, руб./(чел.·год). Данным фактором можно (нужно) управлять, путём изменения цены с учётом инфляции и конкурентной ценовой борьбы между вузами;

$Ч_{РФ}$ – численность студентов РФ, обучающихся в вузе на условиях договора полной оплаты стоимости обучения с вузом, чел. Это частично управляемый фактор, поскольку университет при правильной организации работы с абитуриентами на рынке образовательных услуг может обеспечивать прирост численности абитуриентов, будущих студентов-договорников. В структуре доходов ТОГУ на доходы от договорников РФ в 2024 году приходилось 30,3% [9];

$Ц_{Ин}$ – “цена обучения” в рамках ООП для иностранных граждан, руб./(чел.·год). Данный фактор – управляемый по аналогии с показателем $Ц_{РФ}$;

$Ч_{Ин}$ – численность студентов-иностранцев, обучающихся на условиях договора полной оплаты стоимости обучения с вузом, чел. Данный фактор управляемый по аналогии с показателем $Ч_{РФ}$. Справочно, удельный вес численности иностранных студентов на ООП в ТОГУ в 2024 году составлял 10,35% [9].

Полученная математическая модель (4) позволяет выявить управляемые факторы модели доходов ООП: контингент обучающихся ООП; цена обучения.

Как было ранее показано модель должна одновременно учитывать источник поступления, форму обучения и курс обучения (в силу различий цен для разных лет набора). Поэтому итоговая авторская модель доходов выражается формулой

$$Д_{ООП} = \sum_{l}\sum_{j}\sum_{k} Ц_{ljk} \cdot Ч_{ljk}\,, \quad (5)$$

где l – источник дохода, j – форма обучения, k – курс обучения.

В качестве промежуточного итога укажем, что весьма удобно доходы ООП представить произведением численности ($Ч$) контингента обучающихся ООП на величину средневзвешенной цены ($Ц$). С учётом такого упрощения модель доходов принимает следующий вид

$$Д_{ООП} = Ц \cdot Ч. \quad (6)$$

Графическая модель доходов (6) представлена на рис. 2 в виде прямой в системе координат: контингент – доход.

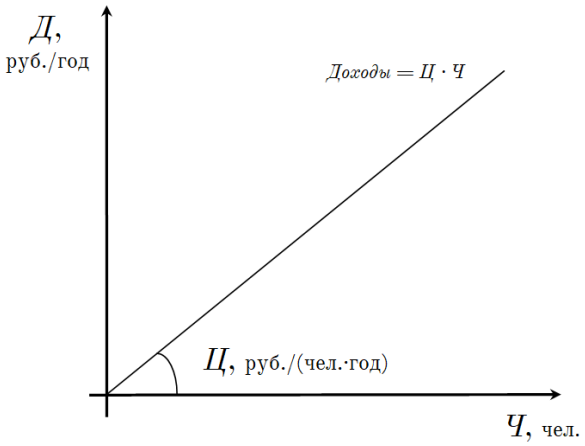


Рис. 2. Графическая модель доходов ООП

4.2. Расходы при реализации ООП

Формула (2) характеризует совокупные расходы при реализации ООП, с выделением в них двух частей по признаку возможности прямого отнесения на конкретную ООП. Возможен также и другой подход [10], при котором совокупные расходы разделены на части в соответствии с «природой» затрат. Речь идёт о выделении в структуре расходов на ООП затрат на заработную плату преподавателей, связанную с

реализацией ООП (далее – *З*), и всех остальных «прочих косвенных затрат» (далее – *О*). Для справки, в структуре затрат ТОГУ на реализацию ООП соотношение между ними составляло в 2024 году пропорцию

$$\frac{З}{О} = \frac{47\%}{53\%}.$$

Учитывая данный подход к классификации затрат ООП, нами была рассмотрена отличная от (2) математическая модель расходов ООП

$$Р_{ООП} = З + О. \quad (7)$$

Заработная плата (*З*), выплаченная преподавателям при реализации конкретной ООП, получена авторами исходя из следующих допущений и исходных данных. В нагрузке вуза присутствует информация о часах контактной работы, связанной с реализацией отдельной взятой ООП и всех вместе по вузу (*Т*). Причём эта информация детализирована в нагрузке по категориям (*i*): ассистент, преподаватель, старший преподаватель, профессор. С другой стороны, по данным бухучёта вуза известны фонды оплаты труда ($ФОТ_i$), использованные в течение года для выплаты заработной платы преподавателям указанных категорий. Деление фондов на нагрузку по каждой категории преподавателей даёт нам величину удельного значения заработной платы, приходящейся на один час контактной работы по преподавателям каждой категории

$$C_i = \frac{ФОТ_i}{T_i}. \quad (8)$$

где C_i – «условная ставка» или величина фонда оплаты труда приходящаяся на один час контактной работы *i*-й категории преподавателей, руб./(чел.·час.); $ФОТ_i$ – фонд оплаты труда преподавателей *i*-й категории преподавателей, руб./год; T_i – совокупный объём часов контактной работы, проведённый преподавателями данной категории в течение года, чел.·час./год.

Как было отмечено, нагрузка вуза содержит информацию обо всей контактной работе, реализуемой со студентами каждой ООП. Эта величина получается после обработки утверждённых учебных планов соответствующих направлений и профилей с учётом планового контингента студентов. Причём, эта информация в нагрузке вуза также детализирована по отдельным категориям преподавателей. Обозначив как t_i величину объёма часов контактной работы ООП по конкретной *i*-й категории преподавателей, можно представить величину расходов на заработную плату ООП, следующей суммой

$$З = \sum_i C_i t_i. \quad (9)$$

Величину $t = \sum_i t_i$ далее по тексту мы будем называть трудоёмкостью ООП, поскольку она и отражает полное трудосодержание (содержание «живого труда») в отдельной ООП в человеко-часах за учебный год. Ключевым для маржинального анализа ООП является выделение в трудоёмкости (t) двух компонент:

- групповая (аудиторная) контактная работа (t^{I}), не зависящая от численности контингента,
- индивидуальная контактная работа (t^{II}), зависящая от численности контингента,

так что для любой ООП $t = t^{I} + t^{II}$. Для справки, отношение между групповой и индивидуальной контактной работой в целом по всем вместе ООП университета ($T_1 = \sum_{ООП} t^{I}$ и $T_2 = \sum_{ООП} t^{II}$), например, в 2024 году в ТОГУ находилось в пропорции

$$\frac{T_1}{T_2} = \frac{81\%}{19\%}.$$

Учитывая, что часть t^{II} – переменная, зависящая от численности контингента на программе, то и расходы на оплату труда, связанные с трудоёмкостью t^{II} являются переменными затратами. Авторы смогли оценить [10] величину часов индивидуальной работы, приходящейся на одного студента (z) для переменной части трудоёмкости (в чел.·час./(студ.·год)), для каждой ООП по категориям преподавателей, исходя из предположения прямой пропорциональной зависимости от контингента (*Ч*) обучающихся

$$t_i^{II} = z_i \cdot Ч. \quad (10)$$

После подстановки в (9) полученного выражения (10) для переменной компоненты t^{II} трудоёмкости, модель расходов на заработную плату примет вид

$$З = \sum_i C_i t_i = \sum_i C_i \left(t_i^{I} + t_i^{II} \right) = \sum_i C_i t_i^{I} + Ч \sum_i C_i z_i. \quad (11)$$

При «свёртке» $\sum_i C_i t_i^{I} = \mathbf{A`}$ и $\sum_i C_i z_i = \mathbf{B`}$, модель (11) можно упростить до линейного вида

$$З = \mathbf{A`} + \mathbf{B`} \cdot Ч. \quad (12)$$

Например, для ООП 38.03.01 Экономика (профиль финансы и кредит) авторами получена следующая «зарплатная» часть модели расходов

$$З_{(38.03.01)} = 6\,158\,920 + 15\,452 \cdot Ч.$$

Остальные прочие затраты (*О*), отнесённые на ООП, разделены авторами на две части:

- **прямые** прочие затраты ($О_{dir}$). Они однозначно увязаны с рассматриваемой ООП, так что в полном они включаются в себестоимость образовательной услуги ООП;

- **косвенные** прочие затраты (O_{indir}) не имеют прямого отношения ни к одной конкретной ООП. К ним, например, относятся коммунальные платежи, оплата услуг других сторонних снабжающих университет организаций, зарплата административно-управленческого персонала и младшего обслуживающего персонала, налоги образовательной организации и другие расходы университета, отнесённые на все ООП вместе взятые.

С учётом этой классификации прочие затраты ООП равны сумме

$$O = O_{dir} + O_{indir}. \quad (13)$$

При отсутствии поставленного на современные информационные системы управленческого финансового учёта, выделение прямых и косвенных прочих затрат может быть затруднено и даже невозможно. Однако это не является непреодолимым обстоятельством для оценки финансового результата ООП, при условии, что косвенные прочие затраты, относимые на ООП (O), каким-либо способом удалось выделить из общей суммы косвенных прочих затрат университета (рис. 1). В данном исследовании авторы с успехом используют метод проекции структуры доходов вуза на совокупный объём косвенных прочих расходов вуза. Поясним этот подход в конкретных значениях ТОГУ за 2024 год. Общая величина косвенных прочих затрат составляла 1 433,3 т.р. Для определения части этих затрат, приходящихся на ООП, была использована информация о структуре доходов университета. Из публичных данных известно, что все доходы (бюджетные и внебюджетные поступления) ТОГУ в 2024 году (за исключением целевых субсидий) составили 2 381,3 млн.р., из которых доходы от реализации ООП составили 1 777 млн.р., что составляет 74,63 % от общих доходов ТОГУ. Именно такая сумма и «спроецирована» на величину косвенных прочих расходов университета. В итоге их величина составила

$$O = 1\,433{,}3 \cdot 0{,}7463 = 1\,069{,}7 \text{ (млн.р.)}.$$

Для ответа на вопрос, какую базу использовать для их распределения на отдельные ООП вуза, возможны разные подходы. Авторы склоняются к наиболее объективному и очевидному из них – использование в качестве базы для распределения косвенных прочих расходов значение приведённого контингента обучающихся на ООП.

Ввернёмся к моделированию расходов ООП.

Подставляя (12) в (6), модель расходов ООП получит вид

$$P_{ООП} = З + O = \mathbf{A`} + \mathbf{B`} \cdot Ч + O. \quad (14)$$

Не очевидным является ответ на вопрос, находятся ли в какой-либо зависимости от численности контингента ($Ч$) остальные прочие косвенные затраты (O)? Авторы считают, что часть этих затрат пропорционально зависит от численности контингента. Например, констатируется факт увеличения коммунальных расходов и затрат на ремонт помещений и учебного оборудования в университете при росте общей численности студентов в вузе, поскольку университет в этом случае ежедневно более длительное время ведёт образовательный процесс (образно говоря, не в одну, а в две смены). Принимая такое допущение, мы можем представить прочие затраты как сумму постоянных (O_1) и переменных (O_2) прочих затрат

$$O = O_1 + O_2. \quad (15)$$

Пропорции между постоянной (O_1) и переменной (O_2) частями косвенных прочих затрат мы определили в той же пропорции что и отношение постоянной ($\mathbf{A`}$) и переменной ($\mathbf{B`} \cdot Ч$) частей заработной платы преподавателей ООП (или частей трудоёмкости t^{I}/t^{II})

$$\frac{O_1}{O_2} = \frac{\mathbf{A`}}{\mathbf{B`} \cdot Ч} = \frac{t^{I}}{t^{II}} = k\ . \quad (16)$$

При этом переменная часть (O_2) связана с численностью контингента ООП через легко определяемый с учётом пропорции (15) коэффициент

$$\mathbf{B``} = \frac{O_2}{Ч} = \frac{O}{Ч(1+k)},$$

из чего, далее, следует линейное представление модели прочих затрат в зависимости от контингента ООП

$$O = O_1 + \mathbf{B``} \cdot Ч. \quad (17)$$

Подставив в модель расходов ООП (14) выражение (16) для прочих затрат, мы получили модель расходов вида

$$P_{ООП} = \mathbf{A`} + \mathbf{B`} \cdot Ч + O^{1} + \mathbf{B``} \cdot Ч.$$

Просуммируем константы ($\mathbf{A`}$ и O^{1}) а также коэффициенты при $Ч$:

$$\mathbf{A`} + O^{1} = \mathbf{A},$$

$$\mathbf{B`} + \mathbf{B``} = \mathbf{B}.$$

Получаем в итоге линейную модель расходов ООП

$$P_{ООП} = \mathbf{A} + \mathbf{B} \cdot Ч. \quad (18)$$

Здесь величина **A** – интерпретируется как годовые условно-постоянные расходы на ООП, руб./год, а **B** – как удельные переменные расходы на одного обучающегося ООП, руб./(чел.·год)

Например, для ООП ТОГУ 38.03.01 Экономика (профиль финансы и кредит) авторами была получена модель расходов (руб./год для 2024 года)

$$P_{(38.03.01)} = 13\,049\,987 + 31\,301 \cdot Ч.$$

Графически данная модель фактических расходов ООП 38.03.01 Экономика представлена на рис. 3.

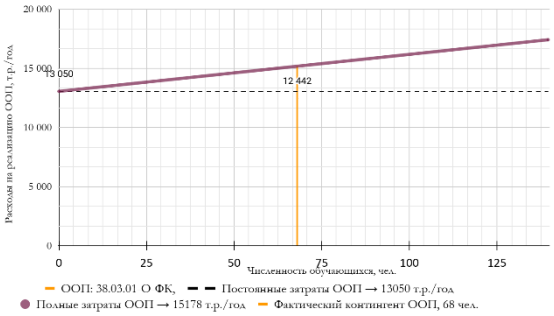


Рис. 3. Модель расходов ООП ТОГУ 38.03.01 Экономика в 2024 году

В результате подстановки в (1) выражений (6) и (18), модель финансового результата ООП будет представлена уравнением вида

$$ФР_{ООП} = Ц \cdot Ч - (\mathbf{A} + \mathbf{B} \cdot Ч). \qquad (19)$$

Покажем далее, как полученная модель финансового результата ООП используется при принятии управленческих решений относительно убыточной ООП. Исходная модель ООП 38.03.01 Экономика (профиль финансы и кредит) в виде (19) записана для параметров 2024 г. уравнением

$$ФР_{(38.03.01)} = 189\,974 \cdot Ч - (13\,049\,987 + 31\,301 \cdot Ч).$$

График данной модели приведён на рис. 4.

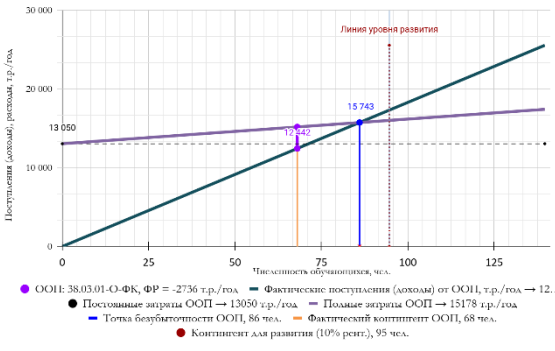


Рис. 4. Модель финансового результата ООП ТОГУ 38.03.01 Экономика в 2024 году

Полный контингент ООП (всех форм и курсов) в 2024 году составлял 68 человек. С учётом структуры учебного плана (как основы нагрузки), эта программа, согласно модельным расчётам, является убыточной с отрицательным финансовым результатом величиной 2 736 тыс.руб./год. Процедура санации ООП предполагает сценарное изменение управляемых факторов (здесь, цены *Ц*, трудоёмкости t, и контингента). При изменении значений этих факторов ООП может выйти на уровень самоокупаемости.

Было предложено пересмотреть учебный план так, чтобы снизить объём часов контактной работы (трудоёмкость t) ООП на 30%. План был не столько редуцирован, сколько унифицирован с другими учебными планами УГСН 38.00.00 для поточного преподавания значительно ряда общих дисциплин на младших курсах. Одновременно предложено поднять цену, но не более чем на текущий уровень инфляции 5%/год. Для адекватности реалиям живого университета, при моделировании авторы учитывали отсев контингента на каждом курсе на уровне сохранности 90%. С учётом планируемого набора на первый курс (очной формы обучения) группы численностью 25 человек, общий контингент ООП в сценарии остаётся на уровне 68 человек (на всех курсах вместе). Совпадение контингента с фактической моделью (2024 года) следует считать случайным. В результате, модель финансового результата ООП приняла вид

$$ФР_{(38.03.01)} = 192\,222 \cdot Ч - (11\,202\,311 + 26\,691 \cdot Ч).$$

График, новой модели финансового результата ООП 38.03.01 Экономика (профиль финансы и кредит), после предложений по оптимизации приведён на рис. 5.

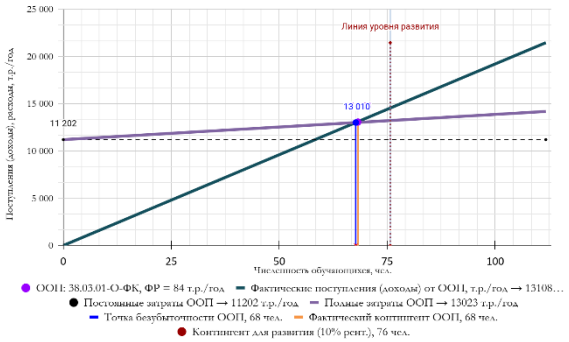


Рис. 5. Модель финансового результата ООП ТОГУ 38.03.01 Экономика (прогноз на 2025 г.)

Обсуждение

В предложенном исследовании проведён комплекс прикладных научно-исследовательских изысканий по построению модели финансового результата отдельной ООП вуза. Моделирование учитывало доходную структуру ООП (бюджетные и внебюджетные поступления, динамику изменения цен по курсам и приведение численности по формам обучения студентов), а также расходную структуру ООП (с выделением условно постоянных и переменных расходов среди прямых и косвенных статей затрат). Показано, как полученная модель может стать эффективным инструментом для принятия решений относительно санации убыточных ООП.

Гипотеза о возможности использования модели финансового результата для принятия решений о реализации (прекращении реализации, санации) ООП, подтвердилась. Модель позволяет рассчитать фактический финансовый результат отдельной ООП университета и составить прогноз финансового результата при сценарном изменении управляемых факторов.

Обсуждению подлежат выделенные авторами управляемые факторы ООП:

1. **численность** контингента ООП (*Ч*), фактор управляется соответствующими рекламными, PR мероприятиями вуза, позволяющими стимулировать поступление на ООП;
2. **цена** за обучение одного студента (*Ц*). Фактор – управляемый вузом, но при требует анализа рыночной конъюнктуры;
3. **часы контактной работы преподавателей** при реализации ООП (t). Фактор полностью «в руках» вуза. Однако он требует поиска баланса между достаточным (обязательным) наполнением ООП контактной работой с преподавателями (для обеспечения высокого качества образовательной услуги) и необходимым максимумом трудосодержания, в случае превышения которого ООП становится убыточной.

Обсуждению подлежат вопросы методического подхода к выделению той части прочих затрат на реализацию ООП, которые можно считать пропорционально зависящими от численности контингента. В данном исследовании мы взяли эту часть из общих прочих затрат в той же пропорции, что и отношение переменных и постоянных трудозатрат на ООП $O_1/O_2 = \mathrm{t}^{\mathrm{I}}/\mathrm{t}^{\mathrm{II}}$.

Продолжение исследований в данной тематике авторы видят в поиске:

- наиболее обоснованных способов определения величины прочих затрат университета, относимых на ООП;
- наиболее целесообразных баз для распределения прочих косвенных расходов;
- методов обоснования «портфеля ООП» для центра финансовой ответственности (кафедры / высшей школы / факультета / института), максимизирующего выбранную целевую функцию (финансовый результат ЦФО) при установленных ограничениях и условиях на реализацию ООП.

5. Заключение

Результаты исследования позволили авторам построить финансовые модели ООП университета, отражающие разные аспекты финансовой деятельности вуза при реализации ООП. Нами получены: модель доходов ООП, модель расходов ООП, модель финансового результата ООП.

Авторами выявлены управляемые факторы, позволяющие приводить убыточные ООП в состояние рентабельных. Модель финансового результата ООП позволяет оценить фактическое значение прибыли (убытка) ООП по абсолютной величине и в относительном выражении (рентабельность продаж). Модель является рабочим апробированным инструментом для принятия управленческих решений в отношении экономической целесообразности открытия новых ООП, закрытия «безнадёжно убыточных» ООП, процедуры санации для продолжения реализации незначительно убыточных ООП, развития и масштабирования рентабельных ООП вуза.

Список используемых источников

1. Козлова Н. И. Финансовая модель и её использование при стоимостном управлении // Теория и практика сервиса: экономика, социальная сфера, технологии. – 2023. – № 3. – С. 27-30.
2. Ростовский Р. В. Модели финансирования образования // Вестник саратовского государственного социально-экономического университета. – 2013. – № 1(45). – С. 60-63.
3. Тараканов В. В. Финансовый механизм системы высшего профессионального образования: от трансформации к модернизации: монография. – Волгоград: Изд-во ВолГУ. – 2009. – 334 с.
4. Тараканов В. В. Становление дифференцированных рынков образовательных услуг и трансформация финансовой деятельности университетов // Национальные интересы: приоритеты и безопасность. – 2009. – №15. – С.22-27.
5. Тараканов В. В. Модернизация финансовой политики университетов Европы в конце XX - начале XXI века // Вестник волгоградского государственного университета. серия 3: экономика. экология. – 2012. – № 1. – С. 29-38.
6. Budgeting model of structural units based on normative per capita funding / S. A. Korolkov, A. G. Losev, I. M. Reshetnikova, V. V. Tarakanov. // European researcher. – 2014. – № 3-1(70). – P. 498-508.
7. Корольков, С. А. Управление образовательными программами через внедрение учебно-финансовых планов / С. А. Корольков, А. Г. Лосев, В. В. Тараканов. // Университетское управление: практика и анализ. – 2016. – № 4(104). – С. 49-55.
8. Кулик И. В. Теоретические подходы к формированию финансовой модели образовательной программы // Финансовые рынки и банки. – 2025. – №5. – С. 263-269.
9. Материалы по результатам проведения мониторинга эффективности деятельности образовательных организаций ВО 2024 г. [Электронный ресурс]. URL: https://monitoring.miccedu.ru/iam/2024/_vpo/inst.php?id=288 (дата обращения: 15.10.2025).
10. Кулик И. В. Прикладные аспекты построения финансовой модели образовательной программы в университете // Инновации и инвестиции. – 2025. – №5. – С. 277-281.

Программный комплекс для поддержки принятия решений в области управления высокотехнологичным производством

Логиновский О. В.
кафедра информационных систем и технологий
Южно-Уральский государственный университет (национальный исследовательский университет)
Челябинск, Россия
e-mail: loginovskiiov@susu.ru

Белякова В. А.
кафедра электронных вычислительных машин
Южно-Уральский государственный университет (национальный исследовательский университет)
Челябинск, Россия
e-mail: beliakovava@susu.ru

Беляков А. Е.
кафедра электронных вычислительных машин
Южно-Уральский государственный университет
(национальный исследовательский университет)
Челябинск, Россия
e-mail: alex.beliakov@susu.ru

Аннотация[1]

Разрабатывается программный комплекс для интеллектуальной поддержки принятия решений в области управления высокотехнологичным производством, функционирующим в условиях неопределенности. Применяются методы агентного моделирования, онтологического инжиниринга, теории управления и подходы к построению озер данных. В результате предложена многоуровневая архитектура программного комплекса, интегрирующая агент-сервисную платформу, унифицированную онтологическую модель предметной области и интеграционную платформу на основе озера данных.

Ключевые слова: управление; агентно-сервисная архитектура; онтологический инжиниринг; цифровой двойник; высокотехнологичное производство; озеро данных.

Keywords: control; agent-service architecture; ontological engineering; digital twin; high-tech manufacturing; data lake.

1. Введение

Цифровая трансформация промышленности актуализировала проблему эффективного управления высокотехнологичными производствами, функционирующими в условиях значительной неопределенности и динамично изменяющихся требований рынка. Несмотря на широкое внедрение систем автоматизации, сохраняется фундаментальное противоречие между жестко детерминированной архитектурой традиционных SCADA/MES-решений и стохастической природой современных производственных процессов. Это проявляется в низкой адаптивности к изменениям технологических регламентов, длительных циклах перенастройки и семантической разрозненности данных между различными уровнями управления.

Анализ современных исследований показывает активное развитие направлений, связанных с применением цифровых двойников [1-3], онтологических моделей [4, 5] и многоагентных систем [6-8] для управления сложными объектами. Существующие российские стандарты [9] задают общие рамки для создания цифровых двойников производства.

Однако проведенный анализ позволяет выявить ряд нерешенных проблем. Большинство существующих разработок фокусируются на решении отдельных задач: семантической интеграции или агентного управления или построения двойников. Комплексных решений, обеспечивающих синергетический эффект от глубокой интеграции этих трех парадигм в единой архитектуре, предлагается недостаточно. Часто онтологии используются как статические справочники, а не как динамические модели, взаимодействующие с агентами в реальном времени. Кроме того, наблюдается дефицит работ, в которых

[1] Труды XI Международной научной конференции "Информационные технологии интеллектуальной поддержки принятия решений", 11-13 ноября, Уфа-Чандигарх-Фандербейлпарк, 2025

формальные онтологические модели служили бы единой семантической основой как для агентного взаимодействия, так и для данных цифровых двойников, обеспечивая сквозную согласованность от корпоративного уровня до уровня оборудования.

В ответ на выявленные ограничения в работе предлагается целостная концепция, основанная на интеграции агентно-сервисного подхода, онтологического моделирования и технологии цифровых двойников [10].

В отличие от существующих аналогов, предлагаемый подход обеспечивает не простое объединение технологий, а их взаимное обогащение: формальные онтологические модели создают единое семантическое пространство для согласованного взаимодействия интеллектуальных агентов и данных цифровых двойников, которые, в свою очередь, предоставляют агентам точные прогнозные модели для обоснования управленческих решений в реальном времени. Децентрализованная агентная архитектура обеспечивает гибкость и адаптивность системы, преодолевая ограничения традиционных централизованных решений.

Объектом исследования выступают процессы управления в высокотехнологичных производственных системах, характеризующиеся высокой степенью неопределенности, динамичностью и необходимостью координации разноуровневых задач.

Предметом исследования являются модели, методы и алгоритмы интегрированного управления на основе агентно-сервисной архитектуры, онтологического подхода и технологии цифровых двойников.

Основная цель исследования — разработка архитектуры и математического обеспечения программного комплекса для интеллектуальной поддержки принятия решений в высокотехнологичном производстве.

Для достижения этой цели требуется разработать многоуровневую архитектуру комплекса; создать формальную онтологическую модель производственной системы; построить формальные модели интеллектуальных агентов; разработать унифицированную модель данных для цифровых двойников технологического оборудования, а также провести экспериментальную верификацию предложенных решений.

Научная новизна работы заключается в разработке целостной концепции построения систем управления производством, сочетающей децентрализованную агентную архитектуру с формальными онтологическими моделями и цифровыми двойниками. Теоретической основой исследования послужили методы теории многоагентных систем, аппарат математической логики для онтологического инжиниринга и современные подходы к построению предиктивных моделей.

2. Методы исследования

Для достижения поставленных целей и решения сформулированных задач в работе применялся комплекс взаимодополняющих методов исследования, объединяющий теоретические и экспериментальные подходы.

2.1. Методы теоретического исследования

Основу теоретической базы составили методы онтологического инжиниринга и агентного моделирования.

Методы онтологического инжиниринга. В работе применен усовершенствованный аппарат онтологического инжиниринга, ориентированный на задачи управления высокотехнологичным производством [11-13]. Базовый процесс построения онтологии включает этапы инвентаризации знаний, концептуализации, формализации и верификации с использованием аппарата описательной логики и стандартов семантической паутины.

Однако традиционные подходы не учитывают специфику динамически изменяющихся производственных систем, что потребовало разработки следующих модификаций: формализацию процесса онтологического инжиниринга как последовательного преобразования множества элементов знания (1)-(2), разработку алгебраической системы формализации онтологии (3), создание метрики когерентности для оценки семантической связности онтологии (4), формализацию оператора логического вывода для условий неполноты производственных данных (5).

Для формализации процессов онтологического инжиниринга применяется математический аппарат теории множеств и алгебраических систем. Процесс построения онтологии формализуется как последовательное преобразование множества элементов знания:

$$K = \{k_1, k_2, ..., k_n\}, \quad (1)$$

где K — исходное множество элементов знания, полученных на этапе инвентаризации; $\mathbf{k_i}$ — i-й элемент знания, представляющий собой кортеж $\langle s, p, o\rangle$ (субъект-предикат-объект).

На этапе концептуализации формируется система концептов:

$$C = \langle T, \leq, \sigma, \Omega\rangle, \quad (2)$$

где T — множество типов (классов) предметной области; $\leq$ — отношение частичного порядка (иерархия классов); σ: $T \rightarrow 2^A$ — функция

спецификации атрибутов классов; Ω — множество операций над концептами.

Отношение частичного порядка удовлетворяет свойствам рефлексивности ($\forall t \in T: t \leq t$), антисимметричность ($\forall t_1, t_2 \in T: (t_1 \leq t_2 \land t_2 \leq t_1) \Rightarrow t_1 = t_2$) и транзитивности ($\forall t_1, t_2, t_3 \in T: (t_1 \leq t_2 \land t_2 \leq t_3) \Rightarrow t_1 \leq t_3$).

Формализация онтологии представляется как алгебраическая система:

$$O = \langle D, P, F, R \rangle, \quad (3)$$

где D — домен интерпретации; $P: C \rightarrow 2^D$ — функция интерпретации концептов; $F: A \rightarrow (D^n \rightarrow D)$ — функция интерпретации атрибутов; $R: R \rightarrow 2^{(D \times D)}$ — функция интерпретации отношений.

Для оценки семантической связности онтологии вводится метрика когерентности:

$$coh(O) = 1 - \frac{|C_d|}{|C| + |R|}, \quad (4)$$

где $|C_d|$ — количество дисконнектных (несвязных) концептов; $|C|$ — общее количество концептов; $|R|$ — количество отношений.

Процесс логического вывода формализуется через оператор замыкания:

$$CL(T) = T \cup \{c \mid T \models c\}, \quad (5)$$

где T — исходное множество утверждений; c — выводимое утверждение; $\models$ — отношение логического следствия.

Предложенный математический аппарат обеспечил строгое формальное основание для построения онтологических моделей, адаптированных к требованиям управления высокотехнологичным производством. Особенностью разработанного подхода является ориентация на обработку динамически изменяющихся производственных данных и обеспечение семантической согласованности в условиях реального времени.

Методы агентного моделирования. В работе предложен усовершенствованный подход к построению многоагентных систем управления производством, основанный на интеграции методов теории конечных автоматов, теории игр и машинного обучения. Базовые принципы агентного моделирования включают концепции реактивных и когнитивных агентов, протоколы взаимодействия FIPA и механизмы координации. Однако традиционные подходы не в полной мере учитывают требования к надежности, адаптивности и робастности в условиях высокотехнологичного производства, что потребовало разработки модификаций [6-10].

Научная новизна подхода включает формализацию поведения агентов-регуляторов на основе детерминированных конечных автоматов (6), разработку кооперативной модели взаимодействия агентов на основе теории игр (7)-(8), адаптацию протокола FIPA Contract Net для распределения задач в производственных системах (9), создание механизмов адаптации на основе Q-обучения (11).

Поведение реактивного агента-регулятора формализуется с помощью детерминированного конечного автомата:

$$A = \langle Q, \Sigma, \delta, q_0, F, \Omega \rangle, \quad (6)$$

где $Q = \{q_0, q_1, ..., q_n\}$ — множество внутренних состояний агента, например: Q = {нормальное, предупреждение, аварийное, обслуживание}; $\Sigma = \Sigma_{sensor} \times \Sigma_{messages}$ — входной алфавит, представляющий декартово произведение множества сигналов от датчиков и множества сообщений от других агентов; $\delta: Q \times \Sigma \rightarrow Q$ — функция переходов, определяющая эволюцию состояния агента; $q_0 \in Q$ — начальное состояние агента; $F \subseteq Q$ — множество терминальных состояний; $\Omega: Q \rightarrow A$ — функция выхода, определяющая действие агента в каждом состоянии.

Взаимодействие множества интеллектуальных агентов формализуется как кооперативная игра:

$$G = \langle N, S, U, C \rangle, \quad (7)$$

где $N = \{1, 2, ..., n\}$ — множество агентов в системе; $S = S_1 \times S_2 \times ... \times S_n$ — пространство совместных стратегий; $U: S \rightarrow \mathbb{R}^n$ — векторная функция полезности; $C: S \rightarrow \mathbb{R}$ — функция общих затрат системы.

Функция полезности для i-го агента определяется как взвешенная сумма компонент:

$$U_i(s) = \alpha \cdot E_i(s) - \beta \cdot C_i(s) + \gamma \cdot Co_i(s) - \delta \cdot Co^n_i(s), \quad (8)$$

где $E_i(s)$ — эффективность агента i при совместной стратегии s; $C_i(s)$ — индивидуальные затраты агента; $Co_i(s)$ — показатель кооперации с другими агентами; $Co^n_i(s)$ — показатель конфликтности; α, β, γ, δ — весовые коэффициенты, удовлетворяющие условию $\alpha + \beta + \gamma + \delta = 1$.

$$W(t_j) = \arg\min_{b_i \in B} \left\lfloor \frac{c_{ij}}{R_i(t_j)} \right\rfloor, \quad (9)$$

где $R_i(t_j)$ — показатель надежности агента b_i при выполнении задач типа t_j.

Для адаптации к изменяющимся условиям производства агенты используют Q-обучение:

$$Q(s, a) \leftarrow Q(s, a) + \alpha[r + \gamma \cdot \max_{a'} Q(s', a') - Q(s, a)], \quad (10)$$

где $s, s' \in S$ — текущее и следующее состояния системы; $a, a' \in A$ — текущее и следующее действия агента; $r: S \times A \rightarrow \mathbb{R}$ — функция вознаграждения; $\alpha \in (0,1]$ — скорость обучения; $\gamma \in [0,1]$ — коэффициент дисконтирования.

Предложенный математический аппарат обеспечивает строгое формальное основание для проектирования, анализа и оптимизации многоагентных систем управления в высокотехнологичном производстве. Особенностью разработанного подхода является комплексное решение задач координации, распределения ресурсов и адаптации к изменяющимся производственным условиям при гарантировании устойчивости и надежности системы.

Математический аппарат теории управления. В работе разработан усовершенствованный математический аппарат для синтеза систем управления высокотехнологичным производством, основанный на интеграции классических методов регулирования с современными подходами предиктивного управления и оценки надежности. Базовые принципы теории управления включают методы ПИД-регулирования, прогнозирования и оценки надежности систем. Однако традиционные подходы не в полной мере учитывают требования к адаптивности, прогнозирующей способности и интегральной надежности в условиях быстро изменяющихся производственных процессов, что потребовало разработки следующих усовершенствований.

Научная новизна подхода включает разработку модифицированного закона управления с прогнозирующей компонентой (11).

Исходной точкой для синтеза является классический закон ПИД-регулирования:

$$u_{pid}(t) = K_p e(t) + K_i \int_0^t e(\tau)\,d\tau + K_d \frac{de(t)}{dt}, \quad (11)$$

где $u_{pi}d(t)$ — базовое управляющее воздействие ПИД-регулятора; $e(t) = y_{tar(t)} - y(t)$ — ошибка регулирования; K_p, K_i, K_t — коэффициенты пропорциональной, интегральной и дифференциальной составляющих.

Для адаптации к условиям неопределенности вводится прогнозирующая компонента, основанная на модели предиктивного управления. Модифицированный закон управления принимает вид:

$$u(t) = u_{pi}d(t) + K_{mp}c \cdot u_p(t), \quad (12)$$

где $K_{mp}c$ — коэффициент усиления прогнозирующей компоненты; $u_p(t)$ — прогнозирующее управляющее воздействие.

Прогнозирующая компонента $u_p(t)$ вычисляется на основе дискретной модели процесса в пространстве состояний:

$$x(t+1) = A \cdot x(t) + B \cdot u(t) + D \cdot v(t), \quad (13)$$

$$y(t) = C \cdot x(t) + w(t), \quad (14)$$

где x(t) — вектор состояния системы; A, B, C, D — матрицы состояния, управления, выхода и возмущений; v(t), w(t) — векторы возмущений и шумов измерений.

Линейные модели в пространстве состояний (13)-(14) применяются для прогнозирования поведения системы в окрестности установившихся технологических режимов, где они адекватно описывают процессы при малых отклонениях от рабочей точки. Для сильнонелинейных объектов модели представляют собой локальные линеаризации, сохраняющие точность на коротком горизонте прогнозирования N_p, соизмеримом с постоянными времени доминирующих динамических звеньев.

Компенсация нелинейностей обеспечивается за счет нижнеуровневых контуров регулирования, адаптации параметров моделей в реальном времени и использования набора линеаризованных моделей для различных режимов работы. Границы применимости считаются превышенными при изменении технологической схемы, переходе между качественно различными режимами или наличии существенных нелинейных эффектов (гистерезис, ограничения), что требует переключения на альтернативные модели или коррекции параметров на основе данных цифровых двойников.

На основе модели (13)-(14) строится прогноз выхода системы на горизонте N_p:

$$\hat{y}(t+k \mid t) = C \cdot A^k \cdot x(t) + \sum_{j=0}^{k-1} C \cdot A^j \cdot B \cdot u(t+k-1-j) + \sum_{j=0}^{k-1} C \cdot A^j \cdot D \cdot v(t+k-1-j), \quad (15)$$

где $k = 1,...,N_p$.

Прогнозирующее управляющее воздействие определяется как взвешенная сумма отклонений прогноза от целевой траектории:

$$u_p(t) = \sum_{i=1}^{N_p} w_i \cdot \left[y_{target}(t+i) - \hat{y}(t+i \mid t) \right], \quad (16)$$

где N_p — горизонт прогнозирования; w_i — весовые коэффициенты; $y_{tar(t+i)}$ — целевое значение выхода.

Надежность системы управления оценивается через интенсивность отказов, которая моделируется как функция эксплуатационных параметров:

$$\lambda(t) = \lambda_0 \cdot \exp[\alpha \cdot S(t) + \beta \cdot T(t) + \gamma \cdot L(t)], \quad (17)$$

где λ_0 — базовая интенсивность отказов; S(t), T(t), L(t) — показатели механической, термической и циклической нагрузки; α, β, γ — коэффициенты влияния.

В контексте данной работы под отказом системы управления понимается событие, приводящее к

полной или частичной потере способности системы выполнять свои функции, а именно: поддерживать технологические параметры в заданных допусках, обеспечивать координацию работы агентов и выполнять производственные планы. Для программных компонентов (агенты, сервисы) применяются отдельные метрики: коэффициент готовности и загрузка вычислительных ресурсов.

Функция надежности системы выражается через интеграл интенсивности отказов:

$$R(t) = \exp\left[-\int_0^t \lambda(\tau) d\tau \right], \qquad (18)$$

где R(t) — вероятность безотказной работы.

Модель надежности (18) применяется для оценки вероятности безотказной работы аппаратных компонентов на интервале между плановыми техническими обслуживаниями, где система рассматривается как невосстанавливаемая. Данный подход корректен при выполнении следующих условий: обслуживание проводится только в моменты t_{pm} по критерию (19); интенсивность отказов $\lambda(t)$ пересчитывается после каждого обслуживания с учетом остаточного износа; модель используется для отдельных критических компонентов. Это позволяет прогнозировать вероятность безотказной работы на межсервисном интервале и оптимизировать график технического обслуживания, минимизируя риск внезапных отказов.

На основе прогноза надежности определяется оптимальный момент проведения предиктивного обслуживания:

$$t_{pm} = \min\{t: R(t) \le R_{min} \vee \lambda(t) \ge \lambda_{max} \vee C_{pm}(t) \le C_{pmmin}\}, \qquad (19)$$

где R_{min}, λ_{max}, C_{pmmin} — предельно допустимые значения надежности, интенсивности отказов и технического состояния.

Предложенный математический аппарат обеспечивает синтез адаптивных систем управления, сочетающих преимущества классических методов регулирования с современными подходами предиктивного управления, что позволяет достигать высокого качества управления при обеспечении надежности и экономической эффективности производственных процессов.

2.2. Методы моделирования и экспериментального исследования

Методология построения цифровых двойников. Разработка цифровых двойников критических активов (на примере кристаллизатора машины непрерывного литья заготовок) проводилась на основе трехуровневой архитектуры, объединяющей физический, виртуальный и интеграционный уровни (рис. 1).

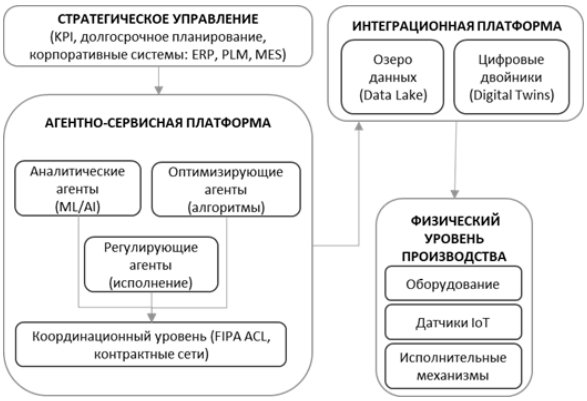


Рис. 1. Общая архитектура программного комплекса

Разработанная архитектура программного комплекса реализует трехуровневую модель интеграции данных и управления. На стратегическом уровне располагаются корпоративные системы ERP и PLM, формирующие целевые показатели и долгосрочные планы развития производства. Агентно-сервисная платформа представляет собой ядро системы, где интеллектуальные агенты трех типов (реактивные, когнитивные и координационные) осуществляют оперативное управление технологическими процессами. Интеграционная платформа на основе Data Lake и цифровых двойников обеспечивает семантическую согласованность данных между всеми уровнями системы. Физический уровень объединяет производственное оборудование, датчики IoT и исполнительные механизмы, формируя замкнутый контур управления.

Для создания виртуальных представлений использовались детерминированные математические модели, основанные на физических принципах (гидродинамика, теплообмен), а также данные систем телеметрии.

Методы машинного обучения. Для решения задач прогнозирования состояния оборудования и предиктивной аналитики применялись алгоритмы машинного обучения, в частности, рекуррентные нейронные сети типа LSTM (рис. 2), эффективные для работы с временными рядами данных телеметрии. Для классификации состояний оборудования использовались ансамбли решающих деревьев.

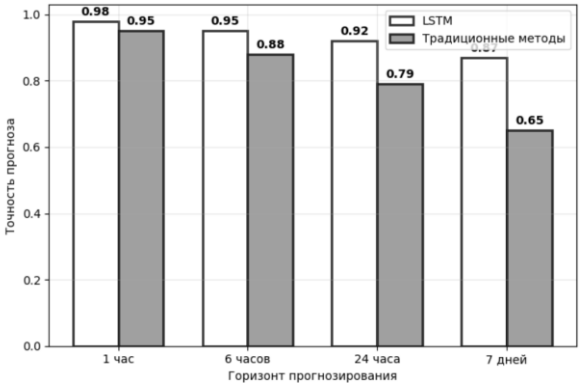


Рис. 2. Сравнение точности прогнозирования

Методы проектирования программных архитектур. Для реализации интеграционной платформы и агентно-сервисной среды применялись принципы сервис-ориентированной архитектуры (SOA) и подходы к построению распределенных систем, включая использование брокеров сообщений (MQTT) и RESTful API для обеспечения взаимодействия между компонентами.

Комплексное применение указанных методов позволило обеспечить теоретическую обоснованность, практическую реализуемость и достоверность полученных в исследовании результатов.

3. Результаты

Разработанный в ходе исследования программный комплекс поддержки принятия решений представляет собой целостную архитектуру, интегрирующую агентно-сервисный подход, онтологическое моделирование и технологию цифровых двойников. Ключевым достижением стала трехуровневая архитектура, обеспечивающая сквозную семантическую согласованность данных от стратегического уровня управления до физических объектов производства.

Центральным элементом архитектуры выступает агентно-сервисная платформа, реализующая три типа интеллектуальных агентов. Реактивные агенты-регуляторы демонстрируют устойчивую работу в задачах детерминированного управления технологическими параметрами, поддерживая стандартные промышленные протоколы OPC UA и Modbus TCP.

Разработанная онтологическая модель предметной области, формализованная на языке OWL2 с использованием правил SWRL охватывает структурный, процессуальный и управленческий аспекты производства, обеспечивая единую семантическую основу для интеграции разнородных данных. Практическая реализация модели показала возможность сокращения времени на семантический поиск в озере данных до 2 секунд при точности идентификации материалов 97% (рис. 3).

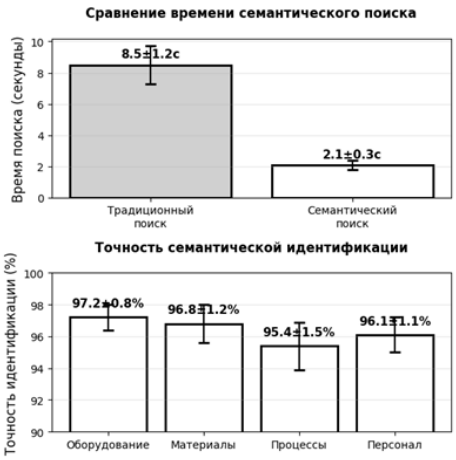


Рис. 3. Эффективность онтологической модели и семантического поиска

Технология цифровых двойников реализована через унифицированную модель данных, объединяющую физические и виртуальные представления объектов. Экспериментальная апробация на синтетических данных подтвердила эффективность комплекса. Сравнительный анализ с традиционными системами управления показал существенное сокращение времени подготовки регламентов, тестирования и адаптации (рис. 4).

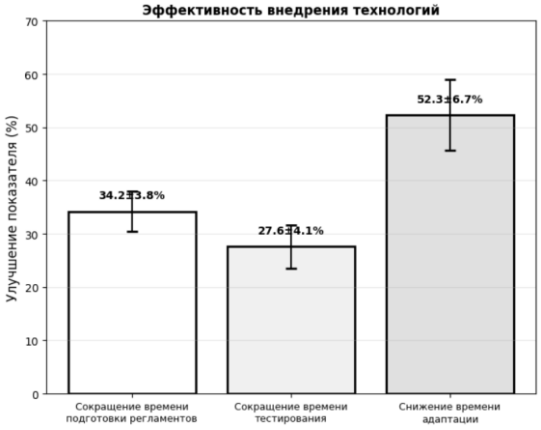


Рис. 4. Результаты тестирования программного комплекса на синтетических данных

Применение разработанного подхода в химической промышленности было продемонстрировано на примере оптимизации процедур поиска в озере данных (Data Lake) цифровых двойников химических соединений и материалов [11]. В рамках исследований была разработана система организации хранения и поиска информации о свойствах атомно-молекулярных систем для построения цифровых двойников материалов.

Ключевым достижением стала разработка методов ускорения поисковых процедур в Data Lake цифровых двойников химических соединений на примере out-файлов квантово-химических расчетов. Секвенирование файлов с выделением структурированных иерархических данных в

формате «атрибут-значение» и последующие процедуры индексирования и распределения значений атрибутов по колоночной базе данных HBase в озере позволили значительно ускорить многокритериальный поиск данных.

Экспериментальные результаты показали существенное ускорение поисковых операций: при работе с 30 000 файлов "OUT Optimization" время поиска по двум параметрам сократилось с 202 666 мс до 9,8 мс, что демонстрирует ускорение более чем в 20 000 раз (рис.5-6). Аналогичные результаты были получены для файлов "OUT Raman" - время поиска сократилось с 225 585 мс до 16,5 мс (рис. 7-8).

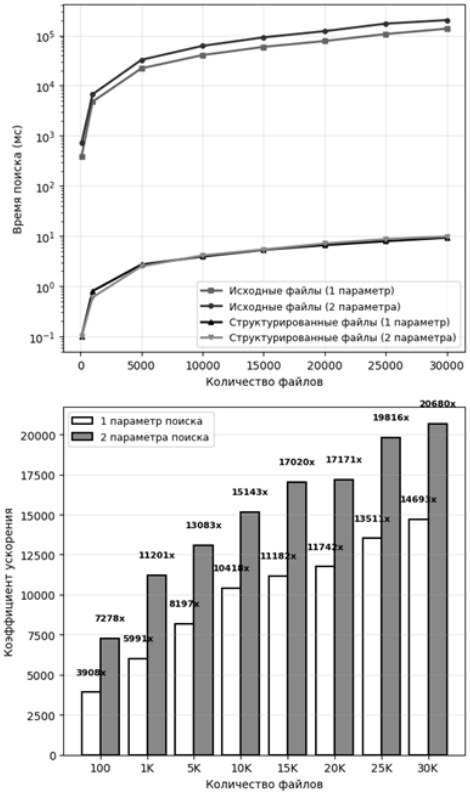


Рис. 8. Оптимизация поисковых процедур в озере данных химических соединений

Практическая значимость разработки заключается в возможности эффективного поиска материалов с заданными свойствами за счет внедрения агентов для оптимизации процедур поиска и упрощения доступа к данным.

Онтологическое пространство цифровых двойников было реализовано как система цифровых паспортов материалов и соединений.

Полученные результаты демонстрируют, что предложенный подход обеспечивает принципиально новый уровень гибкости, масштабируемости, адаптивности и семантической согласованности систем управления высокотехнологичными производствами, создавая основу для перехода к предиктивному управлению и существенного повышения операционной эффективности.

4. Заключение

- Разработанный программный комплекс позволяет преодолеть ключевые ограничения традиционных систем, такие как низкая гибкость, семантическая разрозненность данных и длительные циклы адаптации к изменениям. Многоуровневая архитектура комплекса обеспечивает сквозную интеграцию информационных потоков от корпоративного уровня (ERP) до уровня физических активов (IoT), создавая единое семантическое пространство для принятия управленческих решений.

- Научная новизна работы заключается в целостной концепции, сочетающей децентрализованное управление через кооперирующихся интеллектуальных агентов с формальными онтологическими моделями, обеспечивающими смысловую согласованность данных, и цифровыми двойниками, осуществляющими точное виртуальное отображение физических объектов.

- Практическая значимость подтверждена результатами апробации, которые продемонстрировали значительное повышение точности прогнозирования состояния оборудования за счет применения моделей машинного обучения, интегрированных с онтологиями, существенное сокращение времени адаптации системы к изменяющимся производственным условиям благодаря гибкости агентно-сервисной платформы, сокращение временных затрат на семантический поиск данных и подготовку управленческих регламентов.

- Масштабируемость предложенной архитектуры позволяет применять разработанный комплекс для управления разнородными высокотехнологичными производствами, обеспечивая интеграцию новых модулей и технологических единиц без фундаментальных изменений в системе.

Таким образом, разработанный комплекс обеспечивает переход от реактивного к предиктивному и адаптивному управлению, что создает основу для повышения операционной эффективности, устойчивости и конкурентоспособности высокотехнологичных производств в условиях растущей сложности и неопределенности. Дальнейшие исследования могут быть направлены на развитие методов коэволюции цифровых двойников и их физических прототипов, а также на углубление возможностей коллективного искусственного интеллекта в многоагентных системах.

Благодарности

Автор выражает благодарность доктору технических наук, профессору И.М. Ячикову за

консультации в области моделирования процессов тепломассопереноса в кристаллизаторе машины непрерывного литья заготовок; доктору химических наук, профессору Е.В. Барташевич за экспертные консультации при разработке концепции системы управления и процедур поиска в озере данных цифровых двойников химических материалов.

Список используемых источников

1. Минзов А. С., Невский А. Ю., Баронов О.Р., Немчанинова С. В. Цифровые двойники в системах управления // Вопросы кибербезопасности. 2024. №2 (60).
2. Кокорев Д.С., Юрин А.А. Цифровые двойники: понятие, типы и преимущества для бизнеса // Colloquium-journal. 2019. №10 (34).
3. Лихтциндер Б.Я., Ольберг П.А. Моделирование и цифровые двойники // Вестник Самарского государственного технического университета. Серия: Технические науки. 2022. №4 (76.
4. Массель Л. В., Ворожцова Т. Н. Онтологический подход к построению цифровых двойников объектов и систем энергетики // Онтология проектирования. – 2020. – № 3 (37). – С. 287–301.
5. Смирнов С.В. Онтологический подход к формированию гетерогенных сред моделирования // Вестник Самарского государственного технического университета. Серия: Технические науки. 2011. №4.
6. Логиновский О. В., Голлай А. В., Дранко О. И., Шестаков А. Л., Шинкарев А. А. Эффективное управление организационными и производственными структурами. – М. : НИЦ ИНФРА-М, 2020. – 450 с. – ISBN 978-5-16-016217-1. – DOI: 10.12737/1087996.
7. Кузьмин В. Р., Загорулько Ю. А. Применение агентно-сервисного подхода при разработке интеллектуальных систем поддержки принятия решений в энергетике // Вестник НГУ. Серия: Информационные технологии. – 2020. – № 3. – С. 45–56.
8. Козьминых Н. М., Голованов А. А. Многоагентный подход в системах информационной поддержки управленческих решений // Программные продукты и системы. – 2012. – № 1. – С. 23–28.
9. ГОСТ Р 58871-2020 Цифровое моделирование. Цифровые двойники производства. Общие положения. – М. : Стандартинформ, 2020. – 25 с.
10. Логиновский О. В., Белякова В. А. Модели интегрированного управления наукоёмким производством // Системы управления и информационные технологии. – 2025. – № 2-1 (100). – С. 81–86.
11. Белякова В. А., Логиновский О. В., Шинкарев А. А. Онтологический инжиниринг в управлении ресурсом оборудования наукоемкого производства // Вестник Южно-Уральского государственного университета. Серия: Компьютерные технологии, управление, радиоэлектроника. – 2025. – Т. 25, № 3. – С. 26–41.
12. Массель Л. В., Массель А. Г. Семантическое моделирование при построении цифровых двойников энергетических объектов и систем // Онтология проектирования. – 2023. – № 1 (47). – С. 45–62.
13. Ярушкина Н. Г., Мошкин В. С., Ишмуратова Г. Р., Андреев И. А., Мошкина И. А. Применение способа интеграции нечётких временных рядов и нечётких онтологий в задачах диагностики технических систем // Онтология проектирования. – 2018. – № 4 (30). – С. 56–65.
14. Починская В. А., Топольский Д. В., Ячиков И. М., Беляков А. Е. Применение технологии цифровых двойников для кристаллизатора машины непрерывного литья заготовок // Машиностроение: сетевой электронный научный журнал. – 2023. – Т. 10, № 3. – С. 29–33. – DOI: 10.24892/RIJIE/20230306.
15. Механизмы ускорения поисковых процедур в Data Lake цифровых двойников химических соединений и материалов / А. В. Беляков, С. М. Рассказов, И. Д. Юшина [и др.] // Цифровая Индустрия: Состояние и Перспективы Развития 2023 (ЦИСП'2023) : Сборник научных статей, Челябинск, 21–23 ноября 2023 года. – Челябинск: Издательский центр Южно-Уральского государственного университета, 2024. – С. 103-109. – EDN LMZFYZ.

Методы сопоставления выборок при недостатке размеченных данных

Греков М. В.
Оренбургский государственный университет
кафедра прикладной математики
Оренбург, Россия
e-mail: mihan-nik@yandex.ru

Боровский А. С.
Оренбургский государственный университет
кафедра прикладной математики
Оренбург, Россия
e-mail: borovski@mail.ru

Аннотация[1]

В статье рассмотрены подходы к построению моделей для задач сопоставления наборов данных при ограниченном числе размеченных примеров. Статья посвящена использованию сиамских нейронных сетей, что позволяет обучать нейронную сеть на сходстве объектов вместо традиционной классификации, что снижает требования к объёму исходных данных и повышает качество модели. Приведены примеры применения сиамской архитектуры для анализа текстовых и графических данных, а также рассматривается интеграция метода активной выборки, обеспечивающего приоритетное использование наиболее информативных примеров для обучения. Приведённые примеры демонстрируют, что сочетание сиамских сетей с активной выборкой позволило существенно уменьшить размер исходного набора данных при сохранении или повышении точности модели, что делает данный подход перспективным для задач анализа и классификации в условиях ограниченного числа данных.

Ключевые слова: сиамские нейронные сети; парное обучение; генерация негативов; активная выборка; сравнительный анализ данных; ограниченные наборы данных; обучение с малым количеством данных; анализ изображений; классификация; извлечение признаков; метрики точности; системы поддержки принятия решений.

1. Введение

При создании систем поддержки принятия решений (СППР) может возникать потребность в сравнении двух выборок между собой или с заранее известной парой из двух выборок. Например, в задачах сопоставления результатов или при анализе содержимого двух наборов данных, может быть недостаточно классических алгоритмов, так как требуется внести контекст, на основании которого выборки должны быть сопоставлены.

Также большой проблемой в такой задаче является наличие заранее известных условий сопоставления. Не всегда при работе с выборками достаточно контекста уже существующих пар. В таком случае может потребоваться привлечение эксперта или ручная оценка выборок для создания контекста, положительных и отрицательных результатов.

В большинстве случаев, при исследовании связей может быть большой проблемой вручную разметить большой набор данных без критичных ошибок, либо это может занять слишком много времени, что также может быть недопустимо.

В таких случаях могут помочь средства создания контекста или модели, которые не требуют больших наборов данных, а также могут помочь в улучшении итоговой модели путём наработки связанных данных из выборок с приемлемой точностью.

2. Классические методы

Существует множество методов, которые эффективно могут решать задачи сопоставления, такие как: взвешенное сравнение по признакам, парное обучение с учителем, двухэтапное ранжирование и другие. Каждый из этих методов может быть применён на наборах данных для их сопоставления и нахождения наиболее подходящего варианта, однако, все они имеют свои достоинства и недостатки и требуют различных подходов при недостатке размеченных данных.

Для примера рассмотрим парное обучение с учителем с генерацией негативов и на его основе затронем основные достоинства и ограничения классических методов.

2.1. Парное обучение с учителем с генерацией негативов

При парном обучении создаётся модель, которая нацелена на поиск наиболее подходящих к

[1] Труды XI Международной научной конференции "Информационные технологии интеллектуальной поддержки принятия решений", 11-13 ноября, Уфа-Чандигарх-Фандербейлпарк, 2025

исходному запросу результирующего списка на основе контекста системы.

В первую очередь систему обучают на парах запрос – результат, что позволяет системе учитывать запрос при формировании списка результатов. При этом входные данные используются не для поиска сходства, а для формирования вектора поиска результата по смыслу.

При помощи парного обучения можно реализовать систему сопоставления выборок на основе контекста без особых изменений. Для этого требуется обучить систему на основе одного набора данных, а затем циклически подавать на вход значения из второго набора, тем самым получив список один ко многим, в котором наиболее подходящие варианты будут иметь наибольший процент сходства.

Для обучения такой модели можно использовать ранее размеченные экспертом пары, которые использовать в качестве положительных примеров при обучении модели. Однако, для правильной работы модели требуется также предоставить отрицательные примеры пар, размеченных таким образом, чтобы модель избегала подобных результатов. Для этого можно применить метод генерации негативов.

Метод генерации негативов позволяет создать действительно негативные наборы для обучения модели, что полезно при отсутствии возможности предоставить размеченные экспертом отрицательные данные. При таком подходе достаточно знать признаки, которые явно указывают на отрицательность данных, после чего можно создать кратно больший набор данных, в который добавить положительные данные для формирования набора для обучения.

2.2. Применение парного обучения

В статье «Collective Pairwise Classification for Multi-way Analysis of Disease and Drug Data» авторов Marinka Zitnik и Blaz Zupan рассматривается применимость парного обучения для нахождения сложных связей между различными объектами [1]. Авторы предлагают разработанную модель «Copacar» для работы с данными лекарств. В частности, модель используется для предсказания взаимодействия лекарств с различными клетками, с другими лекарствами, эффектов и влияния на различные типы клеток.

Для оценки модели авторы использовали сравнение полученных результатов с базой данных «DrugBank», в которой хранятся изученные фармакологические эффекты лекарств, а также «PubChem», в которой хранятся химические структуры лекарств.

Результат сравнения представлен на рисунке 1.

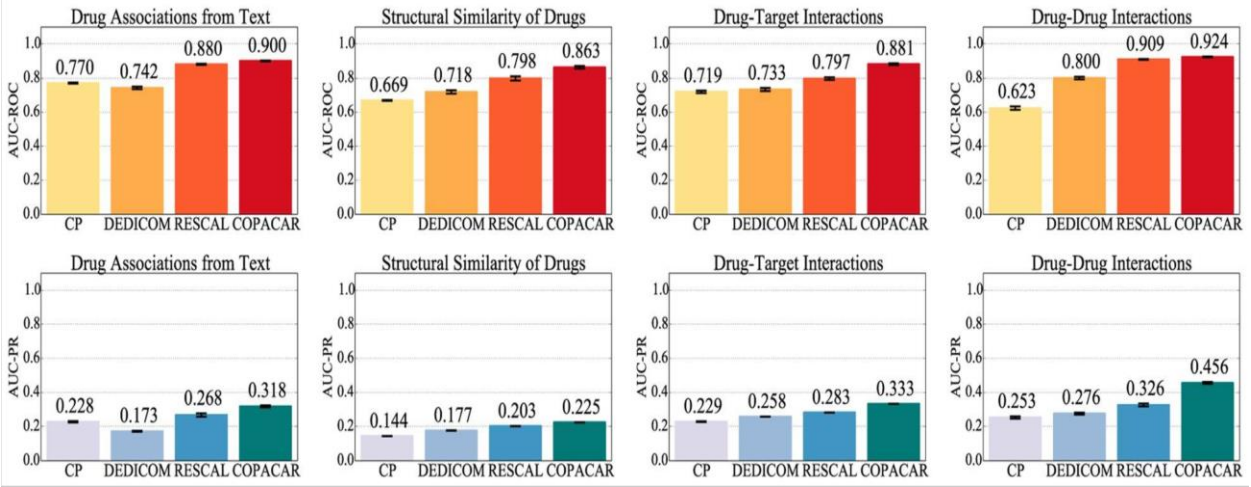


Рис. 1. Метрики AUC-ROC и AUC-PR сравнения моделей CP, DEDICOM, RESCAL и COPACAR в задачах связи лекарств с описанием, структурного сходства лекарств, исследования влияния на целевые клетки и взаимодействия лекарств между собой.

Созданная модель показала высокую эффективность предсказания эффектов лекарств, в сравнении с ранее существовавшими моделями, что показывает эффективность парного обучения при сравнении связанных между собой выборок.

В случае авторов теоретически вычисленные эффекты взаимодействия лекарств в дальнейшем могут быть проверены в лабораторных условиях, что значительно повышает скорость поиска необходимых взаимодействий и эффектов. Применение парного обучения для создания отношений между объектами позволяет находить взаимоотношения для напрямую не связанных явлений, что можно использовать для поиска корреляций между параметрами в системах.

2.3. Недостатки классических методов

Хотя классические методы, в том числе парное обучение, показывают хорошие результаты в задачах сопоставления нескольких выборок, стоит отметить основную сложность работы с такими методами.

Основной проблемой является требования к выявлению признаков. Для эффективной работы модели необходимо ещё до начала обучения разметить признаки в исходных данных, которые модель будет использовать для определения сходства пар, что значительно повышает требование к данным, так как возникает 2 требования: данные должны быть строго преобразованы в определённый формат и выявление экспертом признаков для классификации данных.

Если проблему малого числа исходных размеченных данных можно решить при помощи генерации негативов, то для выявления признаков может потребоваться привлечение эксперта или создания классификатора.

3. Сиамские нейронные сети

Однако, для сиамских нейронных сетей не требуется выявление признаков сходства, так как модель будет вычислять их сама на основе контекста положительных выборок, так как основной отличительной особенностью таких моделей является их способность обучаться не на конкретных классах, а на сходстве между объектами. Вместо того чтобы строить классификатор, который требует большого числа примеров для каждого класса, сиамская архитектура учит модель сравнивать пары образцов и определять, насколько они похожи. Такой подход позволяет эффективно использовать даже ограниченные наборы данных, поскольку из небольшого числа примеров можно сформировать значительно больше обучающих пар, увеличивая разнообразие выборки без необходимости в дополнительной разметке.

Сиамские нейронные сети, в большинстве своём, основываются на методе парного обучения, что позволяет им использовать преимущества этого алгоритма, при этом убирает требование к классификации исходных данных.

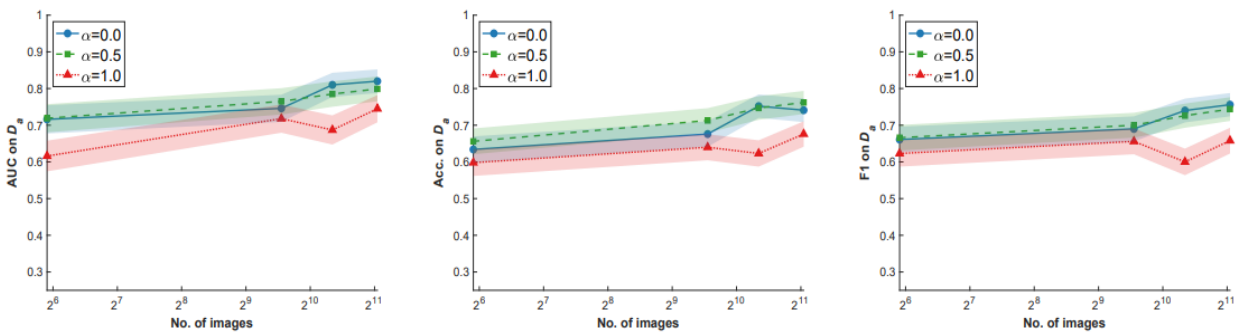
Дополнительная эффективность достигается тем, что обе ветви сети используют одни и те же веса. Благодаря этому модель учится выделять универсальные признаки, которые одинаково хорошо описывают все объекты, а не только примеры из определённых классов. Такой способ обучения обеспечивает лучшую обобщающую способность и позволяет применять модель для распознавания новых категорий, для которых доступно лишь несколько примеров. В результате сиамские сети естественным образом подходят для задач сопоставления, где количество данных крайне ограничено, но требуется точность при сравнении объектов.

3.1. Практическое применение сиамской нейронной сети

Пример применения модели сиамской нейронной сети проводится в статье «Classification and Comparison via Neural Networks». Авторы İlkay Yıldız, Peng Tian и Jennifer Dy используют возможности модели для определения класса по малочисленным примерам выборки [2]. Также они приводят пример, как в нейронных сетях использовать для одной базовой выборки 2 входа данных и сравнивать их между собой. Таким образом они достигают возможность получать метрику сравнения между двумя выборками, а также определять, какой результат лучше подходит.

Для оценки сравнительной метрики применяют функции потерь на основе моделей Брэдли — Терри и Терстоуна. В модели Брэдли — Терри описывается логистической функцией, в то время как в модели Терстоуна применяется нормальное распределение ошибки.

Также авторы приводят результаты работы сиамской нейронной сети для своих наборов данных. Исследуется выборка размером около 10 тысяч сопоставлений, где 60% отдаётся на обучение, 20% на валидацию модели и 20% - итоговое тестирование модели. На рисунке 2 представлено падение точности при изменении числа положительных результатов исследуемого класса с 800 до 300, где α – коэффициент относительности меток. α = 0 – только относительные сравнения, а α = 1 – только абсолютные сравнения меток.



Методы сопоставления выборок при недостатке размеченных данных

Рис. 2. Метрики точности при снижении выборок исследуемого класса.

После этого авторы приводят сравнение эффективности применения относительных и абсолютных меток. На рисунке 3 представлено сравнение в виде графиков для 3 наборов данных и 3 метрик: Accuracy, F1 и ROC-AUC.

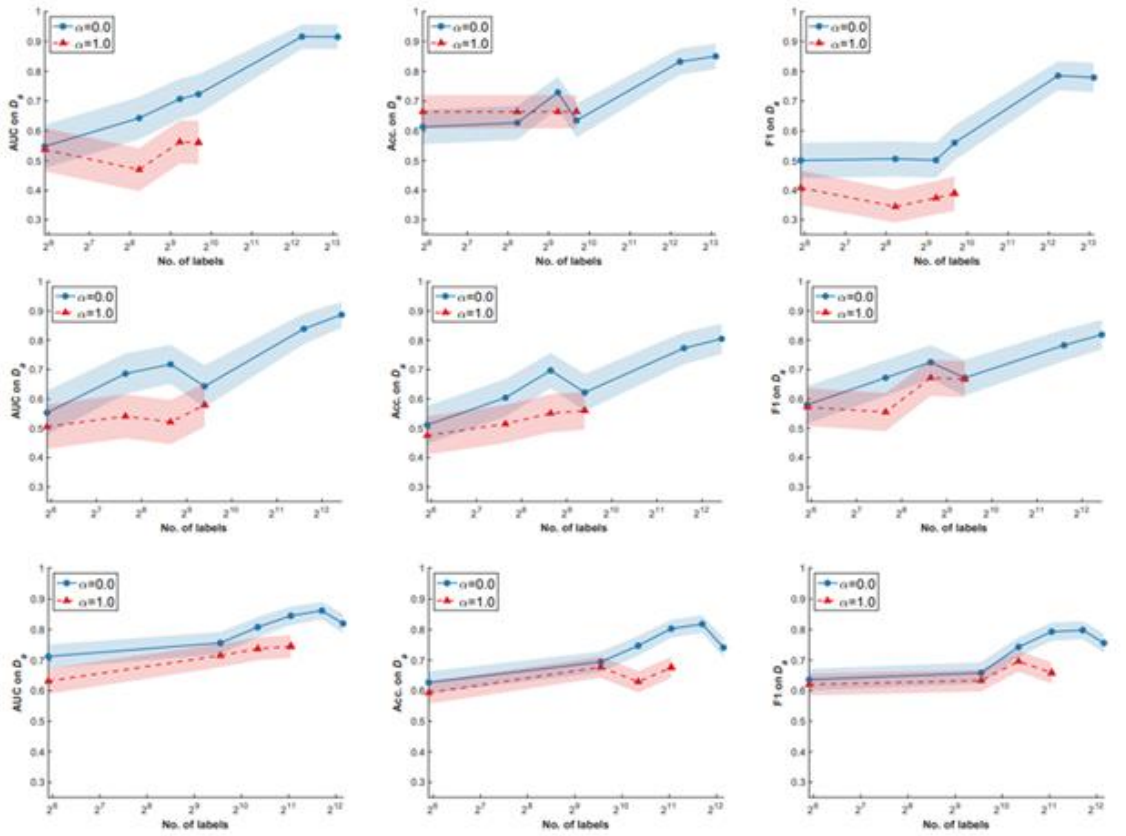


Рис. 3. метрики Accuracy, F1 и ROC-AUC для наборов данных GIFGIF Happiness, GIFGIF Pleasure и FAC.

Исходя из результатов, полученных авторами, можно сделать вывод, что сиамские нейронные сети достаточно эффективно снижают требования к исходному набору данных, а также могут повысить точность результатов.

Однако, сиамские нейронные сети могут эффективно работать не только с текстовыми, но и графическими данными для извлечения признаков. Сравнивая изображения между собой модель может определять признаки на изображениях для дальнейшей классификации, а значительно меньшие требования к исходным данным позволяют использовать малые наборы изображений, которые могут быть составлены одним или несколькими экспертами.

3.2. Сиамские нейронные сети и метод активной выборки

Использование сиамской нейронной сети для работы с изображениями представили авторы Tahsin Reasat, Asif Sushmit и David S. Smith в статье «Data efficient contrastive learning in histopathology using active sampling». Основная задача авторов состояла в определении типа гистологических данных на основе признаков [3]. Основная проблема состоит в том, что в наборах данных разное число изображений для различных типов тканей, что может повлиять на качество полученных результатов, так как для классов с малым числом исходных данных будет выделено недостаточно признаков.

Для решения этой проблемы авторы применили сравнение исходных изображений между собой для более явного выделения признаков. Они полагали, что это позволит повысить качество результатов для классов с малым числом изображений. На рисунке 4 представлены различные типы изображений тканей в наборе данных.

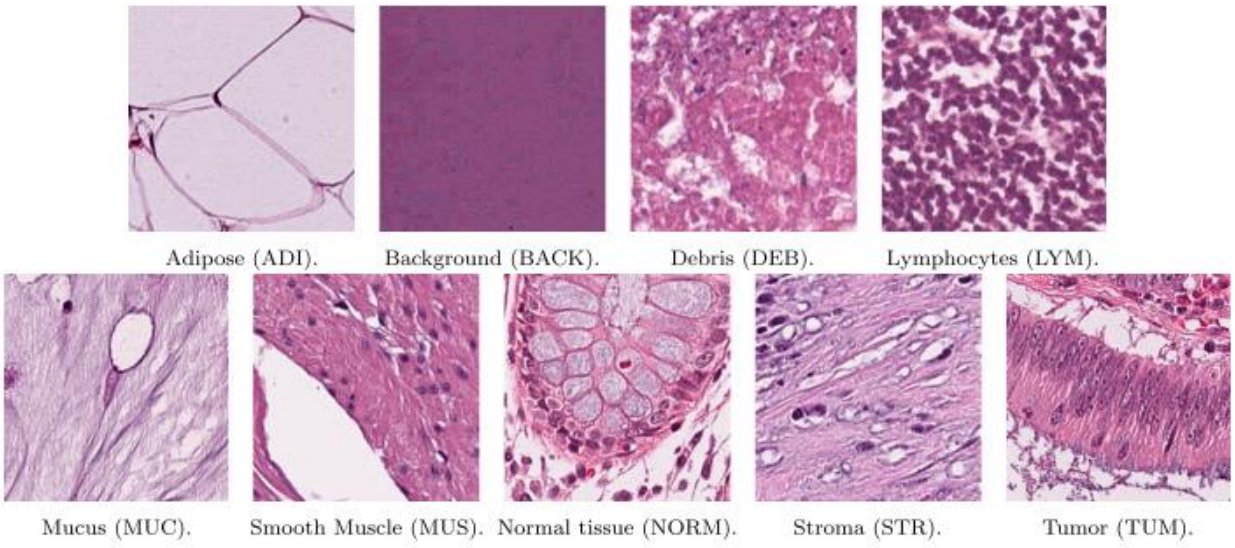


Рис. 4. Изображения различных типов тканей

Помимо использования сиамской нейронной сети авторы применили метод активной выборки. При использовании такой выборки из исходного набора данных сначала выбираются размеченные данные, после чего выбираются данные, которые максимально повлияют на обучение. К таким данным относятся выборки, которые наиболее неопределённые по мнению сети. Такой подход позволяет быстрее и точнее обучить модель.

Схема нейронной сети с активной выборкой представлена на рисунке 5.

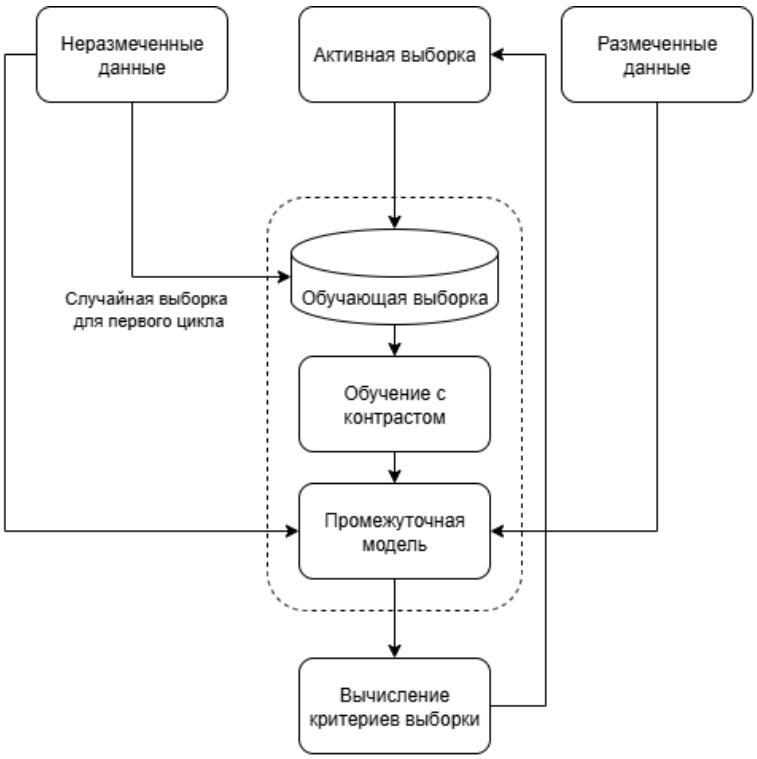


Рис. 5. Модель обучения с активной выборкой

В качестве доказательств эффективности такого подхода авторы приводят различные метрики оценки модели для разного числа исходных данных. Графики представлены на рисунке 6.

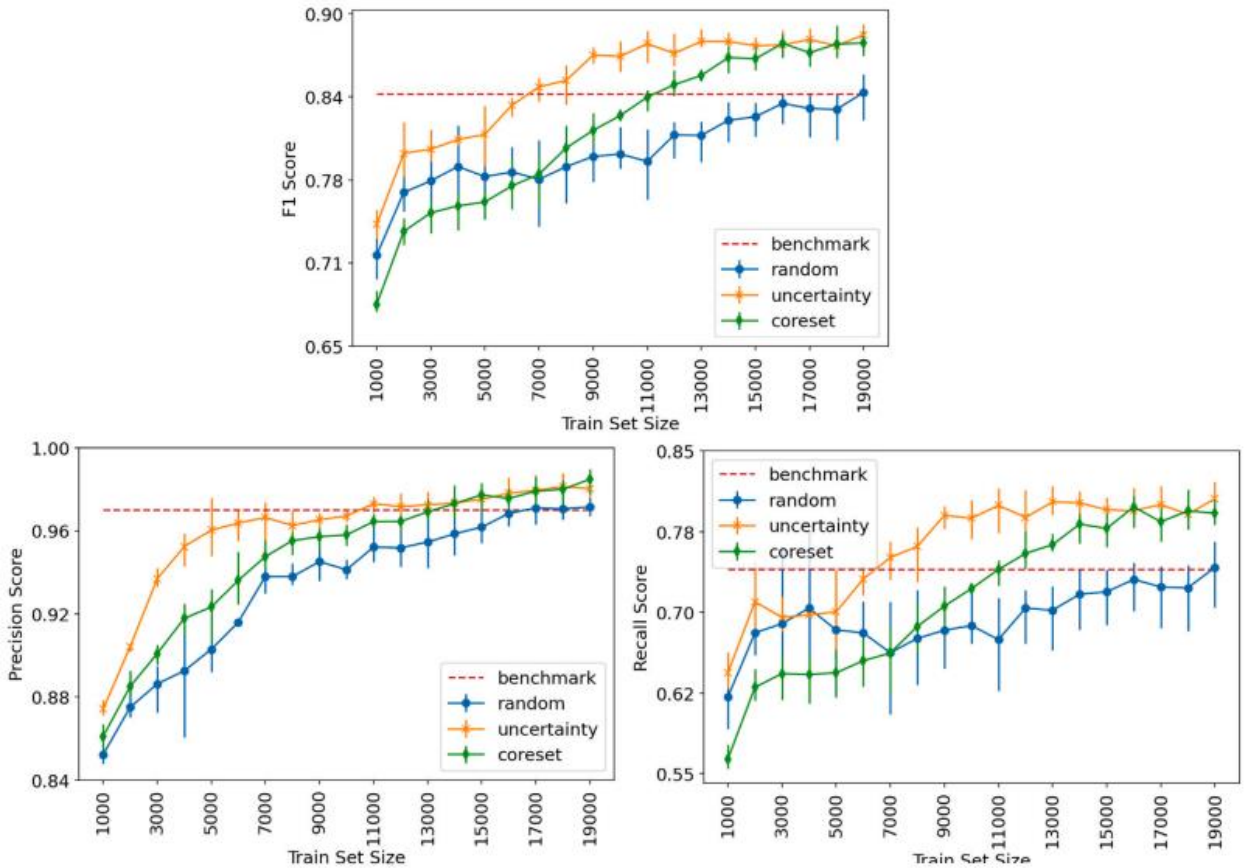


Рис. 6. Метрики F1, Precision и Recall для случайной, неопределённой и корневой выборок (неопределённая и корневая – наборы активной выборки)

Можно заметить, что ожидаемые значения достигаются при значительно меньшем исходном наборе данных, что позволяет снизить размер набора от 2 до 5 раз, в зависимости от требуемого результата.

В заключительной части статьи авторы также указывают, что в их случае также требовалось подбирать размер промежуточной модели, так как малый размер значительно снижал качество модели, а большие значения требовали больше времени для обучения. График сравнения размера промежуточной модели к размеру набора данных представлен на рисунке 7.

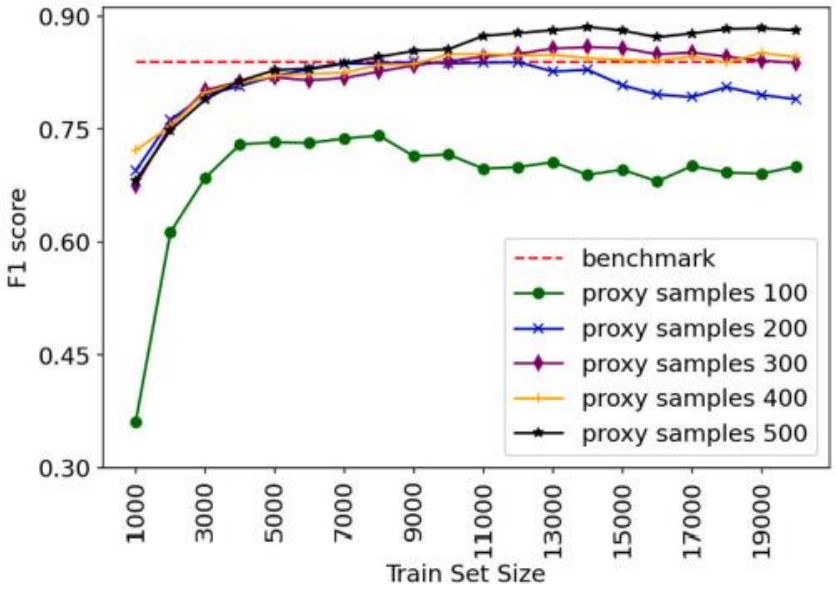


Рис. 7. График соотношения размера промежуточной модели к размеру набора данных

Из графика следует, что слишком малый размер промежуточной модели значительно снижает качество результата, но зависимость не линейна и требуется подбирать размер для конкретной выборки, так как большая промежуточная модель не всегда значительно улучшает результат.

4. Заключение

1. При создании моделей с использованием нейронных сетей важно учитывать размер исходных данных. В случаях, когда выборка мала или размеченных данных недостаточно для качественного обучения, можно использовать несколько подходов: применить сиамские нейронные сети для сравнительного анализа выборок между собой и активную выборку для значительного повышения качества модели при меньшем числе циклов обучения.

2. Классические методы показывают высокую точность результатов при решении задач сопоставления и могут быть обучены при малом наборе исходных выборок, однако, сложности при создании классификации и подготовке данных значительно снижают применимость этих методов.

3. Сиамские нейронные сети могут быть использованы в случаях, когда требуется сопоставить несколько наборов данных между собой, при этом располагая малым числом размеченных данных и не требуют создания классификаторов. Такой подход позволяет значительно повысить качество модели при ограниченном исходном контексте задачи и отсутствии эксперта.

Список используемых источников

1. Zitnik, M., & Zupan, B. (2016). COLLECTIVE PAIRWISE CLASSIFICATION FOR MULTI-WAY ANALYSIS OF DISEASE AND DRUG DATA. Pacific Symposium on Biocomputing. Pacific Symposium on Biocomputing, *21*, 81–92. – PMC: https://pmc.ncbi.nlm.nih.gov/articles/PMC4719585/

2. Reasat, T. Data efficient contrastive learning in histopathology using active sampling / T. Reasat, A. Sushmit, D. S. Smith // Machine Learning with Applications. – 2024. – Vol. 17. – 100577. – ISSN 2666-8270.–DOI: https://doi.org/10.1016/j.mlwa.2024.100577.

3. Yıldız, İ. Classification and comparison via neural networks / İ. Yıldız, P. Tian, J. Dy, D. Erdoğmuş, J. Brown, J. Kalpathy-Cramer, S. Ostmo, J. P. Campbell, M. F. Chiang, S. Ioannidis // Neural Networks. – 2019. – Vol. 118. – P. 65–80. – ISSN 0893-6080. – DOI: https://doi.org/10.1016/j.neunet.2019.06.004.

Комплексный анализ стабильности многомерных систем на основе топологических характеристик

Гвоздев В. Е.
Институт информатики, математики и робототехники
Уфимский университет науки и технологий
Уфа, Россия
e-mail: wega55@mail.ru

Бежаева О. Я.
Институт информатики, математики и робототехники
Уфимский университет науки и технологий
Уфа, Россия
e-mail: obezhaeva@gmail.com

Ракипова А. С.
Институт информатики, математики и робототехники Уфимский университет науки и технологий
Уфа, Россия
e-mail: aliyasr21@gmail.com

Аннотация[1]

Принцип комплексности – базовый принцип исследования систем, что предполагает их разноаспектное исследование. Концептуальную основу представленных результатов составляет положение о том, что степень изменчивости топологических характеристик в результате воздействий - объективной характеристикой стабильности. Предложен подход к формированию обобщенных характеристик стабильности на основе частных характеристик, соответствующих разным видениям системы. Основу подхода составляет формирование на основе структурных характеристик графов, соответствующих разным видениям системы, совокупности однородных выборок независимых случайных величин. Выборки служат основой вычисления характеристик неопределенности в виде энтропии По Шеннону, нормирование которых позволяет получить сопоставимые характеристик стабильности. Наличие сопоставимых характеристик стабильности делает возможным, во-первых, формировать свертки, имеющие смысл обобщенных параметров. Во-вторых, оценивать чувствительность разных свойств системы к случайным воздействиям.

Ключевые слова: многомерная система; стабильность; граф; топологические характеристики; энтропия.

1. Введение

Принцип комплексности – базовый принцип исследования сложных систем, что предполагает их разноаспектное исследование. Один из подходов к исследованию сложных технических систем основан на их рассмотрении как многомерных многосвязных объектов. Реализация принципа комплексности обуславливает необходимость получения сопоставимых оценок стабильности, соответствующих разным видениям системы. Стабильность означает способность системы оставаться неизменной в условиях внешних и внутренних воздействий разной природы; сохранять свои свойства и функции во времени. Отсюда вытекает, что существует прямая зависимость между надежностью системы и стабильностью.

Следуя принципу «равной надежности» можно утверждать, что одним из условий эффективного управления технической системой является надежность системы, что обеспечивается в том числе сопоставимым уровнем надежности разных подсистем, компонентов и процессов, имеющих место в системе. Из этого можно заключить важность обеспечения сопоставимой стабильности подсистем, компонентов и процессов. Понятие неопределенность является обратным понятию стабильность. Исходя из этого можно утверждать, что получение количественных характеристик неопределенности создает основу для получения оценок стабильности. Наличие сопоставимых оценок стабильности в рамках разных видений системы, во-первых, создает возможность получения интегральных оценок стабильности системы. Во-вторых, позволяет исследовать

[1] Труды XI Международной научной конференции "Информационные технологии интеллектуальной поддержки принятия решений", 11-13 ноября, Уфа-Чандигарх-Фандербейлпарк, 2025

чувствительность частных показателей стабильности по отношению к целенаправленным и случайным воздействиям на систему.

В настоящей работе основу формирования показателей стабильности составляют топологические характеристики структурных моделей, представленных в виде графов, а также использование методов математической статистики для обработки топологических характеристик, включая специальные методы обработки малых выборок. Выбор аппарат теории графов обусловлен том, что представление объектов исследования в виде графов позволяет в единой форме представить разные по содержанию отношения и связи между компонентами системы. Использование методов математической статистики создает основу получения устойчивых оценок стабильности на основе измерительных данных.

2. Методическая основа получения сопоставимых оценок стабильности

Аппарат теории графов позволяет строить структурные модели систем разной природы и посредством типовых процедур исследовать их свойства (свойство изоморфизма структурных моделей). Вид структурной модели определяется целями исследования и физической природой объекта моделирования, так что в рамках одного и того же видения системы может быть построено множество графов (свойство полиморфизма структурных моделей). В литературе [1, 2 и др.] описано достаточно много частных топологических характеристик, позволяющих осуществлять разноаспектный анализ графов. Их особенностью является то, что каждая из них зависит от размера графа. Это обстоятельство препятствует получению сопоставимых оценок в случае, когда разным видениям соответствуют графы разных размеров. Возможное различие в числе узлов также препятствует использованию подходов, ориентированных на мультиплексные[2] сети [3, 4]. В силу указанной причины первой опорной точкой получения сопоставимых оценок стабильности, соответствующих разным видениям системы, является переход от абсолютных топологических характеристик графов к относительным, инвариантных к размеру графа.

Второй опорной точкой является то, что топологические характеристики могут рассматриваться как случайные величины. Основанием к этому является следующее:

- цель создания технической системы и условия её использования являются взаимосвязанным понятиями. В силу этого виды внешних воздействующих факторов и диапазоны их изменения могут считаться априорно известными, а моменты времени воздействий – случайными;
- на этапе проектирования выбор варианта построения системы из альтернативных осуществляется, в том числе, на основе оценивания рисков, обусловленных разными возможными отказами, причем число причин отказов считается известным и конечным. Основу оценивания рисков составляют вероятностные характеристики отказов и масштабов последствий;
- считается, что состояние системы после каждого воздействия фиксируется одним и тем же способом посредством одних и тех де средств измерения.

С учетом вышеизложенного можно считать, что исчерпывающей характеристикой топологических параметров будет их закон распределения. Следует отметить, что идея описания свойств случайных графов посредством дифференциальной функции распределения степеней узлов не является новой, и, например, рассматривается в [3, 5].

3. Формирование параметров стабильности

Известен подход к исследованию надежности технических систем, основанный на аппарате статистических индексов [6]. Суть подхода состоит в формировании обобщенных параметров на основе частных измеряемых характеристик. В рамках этого подхода обеспечение сопоставимости характеристик, изначально имеющих различную размерность, обеспечивается за счет перехода от их абсолютных значений к относительным.

Пусть число видений системы $\propto_1, \propto_2, \dots, \propto_m$ составляет m, и видению $\propto_j$ соответствует структура, характеризуемая случайным графом $(V^{(\propto_j)}, E^{(\propto_j)})$. Степень i-го узла $p_i^{(\propto_j)}, i = \overline{1, V^{(\propto_j)}}$. Переход от абсолютных значений степеней узлов к относительным в случае ненаправленных графов осуществляется по правилу:

$$\widehat{p_i^{(\propto_j)}} = \frac{2 * p_i^{(\propto_j)}}{V^{(\propto_j)} * (V^{(\propto_j)} - 1)} \qquad (1)$$

Знаменатель $(V^{(\propto_j)} * (V^{(\propto_j)} - 1)/2)$ соответствует числу ребер полносвязного графа с числом узлов $V^{(\propto_j)}$. Выбор в качестве базовой характеристики количества ребер полносвязного графа обусловлен тем, что полной ячеистой топологии соответствует наибольший показатель структурной надежности [7]. Использование нормирования такого вида позволяет получать сопоставимые относительные характеристики степеней узлов для графов разных размеров.

Если в результате случайных воздействий случайным образом изменяется размеры графов, соответствующих разным видениям, то совокупности видений системы ставится в соответствие m случайных выборок разного объема. Первым шагом формирования сопоставимых

[2] В случае, если при различных целях исследований можно считать, что количество узлов является неизменным, но изменяется число ребер и их смысловое содержание (т.е. имеем дело с мульти графом), то может быть использован аппарат исследования мультиплексных сетей.

показателей стабильности является преобразования случайных выборок к виду законов распределения непрерывных случайных величин:

$$B^{(1)}: \widehat{p_i^{(\propto_j)}} \rightarrow \widehat{F_{\propto_j}}(p) \qquad (2)$$

Здесь $B^{(1)}-$ оператор преобразования однородной выборки независимых случайных величин к виду закона распределения. В силу того, что степень узла не может быть отрицательной (значению «нуль» соответствует изолированный узел), а максимальное возможное значение составляет $(N-1)$ (что соответствует полносвязному графу), и при этом на число узлов графа ограничений не накладывается, в качестве законов распределения относительных степей узлов следует использовать те, у которых диапазон возможных значений аргумента $X \in [0,\infty)$.

Выше отмечалось наличие обратного соотношения между стабильностью и неопределенностью состояния системы. Преобразование $\widehat{F_{\propto_j}}(p)$ к виду скалярной характеристики – энтропии по Шеннону, позволяет получить количественную оценку неопределенности $S^{(\propto_j)}$ в рамках j-го видения:

$$B^{(2)}: \{\widehat{F_{\propto_j}}(p)\} \rightarrow S^{(\propto_j)} \qquad (3)$$

Здесь $B^{(2)}-$ оператор вычисления энтропии.

Оценки законов распределения, соответствующие разным видениям, могут иметь разный масштаб, что обусловлено разными размерами графов[3]. Обеспечение независимости от размера графа обеспечивается посредством преобразований:

$$S^*_{\propto_j} = S^{(\propto_j)} \Big/ H_0^{(\propto_j)} \qquad (4)$$

Здесь $H_0^{(\propto_j)}-$ значение энтропии, соответствующее равномерному закону распределения степени узлов при известном числе узлов графа $V^{(\propto_j)}$. Для вычисления значения $H_0^{(\propto_j)}$ определим границы изменения степеней узлов на основе анализа степеней узлов базовых топологий [7]:

- линейка: наименьшая степень – 1 (конечные узлы); наибольшая степень – 2;
- кольцо: наименьшая степень совпадает с наибольшей и равна 2;
- звезда: наименьшая степень – 1 (терминальные узлы); наибольшая – $(V^{(\propto_j)}-1)-$ центральный узел;
- полная ячеистая топология (полносвязный граф): наименьшая степень совпадает с наибольшей и равна $(V^{(\propto_j)}-1)$.

В [7] отмечается, что структуры реальных систем представляют собою композицию из базовых топологий. В силу этого границами диапазона изменения степеней узлов произвольного графа являются $p_i^{(\propto_j)} \in [1, V^{(\propto_j)}-1]$. Величина диапазона составляет $d^{(\propto_j)} = V^{(\propto_j)} - 1 - 1 = V^{(\propto_j)} - 2$.

Известно, что наибольшее значение энтропии в случае, если известны границы возможных значений случайной величины, соответствует равномерная плотность распределения непрерывной случайной величины. Выше было отмечено, что минимальной степени неизолированных узлов графа соответствует значение «единица».

Известно, что однородным выборкам ставится в соответствие унимодальная дифференциальная функция распределения. При известной верхней границе степеней графа и учитывая, что имеет место равномерный закон распределения, значение $H_0^{(\propto_j)}$ может быть определено посредством выражения:

$$H_0^{(\propto_j)} = -\int_1^{V^{(\propto_j)}-1} \frac{1}{\left(V^{(\propto_j)}-2\right)} * \ln\left(\frac{1}{\left(V^{(\propto_j)}-2\right)}\right) dp = \ln(V^{(\propto_j)}-2) \qquad (5)$$

Таким образом, в качестве показателя стабильности, соответствующего разным видениям системы, может использоваться скалярная величина

$$Stb^{(\propto_j)} = \left|\frac{H_0^{(\propto_j)} - S^{(\propto_j)}}{H_0^{(\propto_j)}}\right| = |1 - S^*_{\propto_j}| \qquad (6)$$

Очевидно, что $Stb^{(\propto_j)} \in [0,1] \; \forall \propto_j, j = \overline{1;m}$. На основе соотношения (6) может быть сформирована таблица, содержащая показатели стабильности, соответствующие разным видениям системы и отметкам времени (Таблица 1). Отметки времени регистрации показателя стабильности соотносится со временем воздействий на систему.

Таблица 1. – Соотношение показателя стабильности и времени воздействия на систему

Время регистрации состояния системы	**Показатели стабильности**			
	$\propto_1$	$\propto_2$	…	$\propto_m$
t_1	$Stb_1^{(\propto_1)}$	$Stb_1^{(\propto_2)}$	…	$Stb_1^{(\propto_m)}$
t_2	$Stb_2^{(\propto_1)}$	$Stb_2^{(\propto_2)}$	…	$Stb_2^{(\propto_m)}$
⋮	⋮	⋮	…	⋮
t_n	$Stb_n^{(\propto_1)}$	$Stb_n^{(\propto_2)}$	…	$Stb_n^{(\propto_m)}$

Полученная таблица делает возможным использовать аппарат статистических индексов [6] для исследования стабильности многомерных систем.

Так, могут быть сформированы обобщенные показатели:

$$B^{(3)}: \{Stb_{t_k}^{(\propto_j)}\} \rightarrow I_{t_k} \qquad (7)$$

Здесь $B^{(3)}-$ оператор свертки. Выбор структуры оператора определяется содержанием конкретной задачи исследования. Разные операторы для

[3] В настоящей работе под размером графа понимается сумма количества вершин V и рёбер E графа

реализации линейных и нелинейных сверток представлены, например, в [6]. На основе анализа рядов динамики показателей стабильности, соответствующих разным видениям $\propto_1, \propto_2, \dots, \propto_m$ могут быть оценены тенденции и тренды изменения показателей стабильности, а также чувствительность разных структур к оказываемым воздействиям.

4. Статистические оценки топологических характеристик

В рамках предлагаемого подхода фокусом получения сопоставимых показателей стабильности является построение на основе выборок значений относительных степеней узлов законов распределения непрерывных случайных величин. В [8] описан метод построения непараметрических оценок законов распределения непрерывных случайных величин в случае, когда вид закона распределения априорно неизвестен, и объем выборки может оказаться малым. Основу метода составляет последовательность преобразований, представленная на Рисунке 1.

$$\{\hat{\nu}_1, \hat{\sigma}\} \rightarrow \hat{\nu}_1^{(\text{н})} \rightarrow \{F_э^{(\text{н})}(x^*)\} \rightarrow \hat{F}^{(\text{н})}(x^*) \rightarrow \hat{F}(x)$$

Рис. 1. Последовательность преобразований

На Рисунке 1 ν_1, σ – эмпирические оценки математического ожидания и среднеквадратического отклонения, получаемые по выборочным данным; $\nu_1^{(\text{н})}$ – отношение эмпирического значения математического ожидания ν_1 к эмпирическому среднеквадратическому отклонению σ; $F^{(\text{н})}(x^*)$ – значения интегральной функции распределения, в заранее оговоренных точках x* на основе ранее рассчитанного набора $\left\{F_э^{(\text{н})}(x^*)\right\}$ интегральных функций распределения; $F(x)$ – оценка интегральной функции распределения, получаемая по $F^{(\text{н})}(x^*)$ с учетом показателя масштаба σ.

Центральное место в упомянутом методе занимает таблица, в которой в одних и тех же точках x^* вычислены значения нормированных интегральных функций распределения $F^{(\text{н})}(x^*)$, которым соответствуют разные значения нормированных математических ожиданий $\nu_1^{(\text{н})}$.

Особенностью метода является полная формализация, что, в отличие от других методов построения законов распределения по выборочным данным, исключает необходимость выдвижения и проверки гипотез о виде закона распределения случайной величины [9, 10].

Пример

Из полносвязного графа, содержащего 20 узлов, случайным образом было удалено 29 ребер. По описанной выше схеме были определены степеней узлов, на основе которых была сформирована выборка нормированных степеней узлов. На основе выборки были определены статистические характеристики в виде дифференциальной функции распределения относительных степеней узлов, что позволило оценить показатель стабильности в виде нормированной энтропии. Полученные результаты представлены ниже.

На Рисунках 2 и 3 представлены сформированный граф и дифференциальная функция распределения. В Таблицах 2 и 3 представлены степени узлов графа до и после нормирования соответственно.

В Таблице 4 показаны полученные результаты для этого графа.

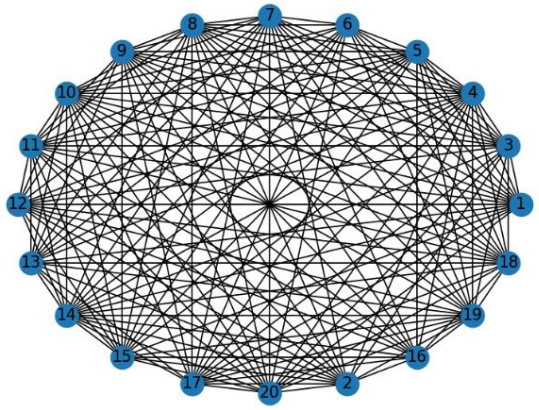


Рис. 2. Сформированный граф

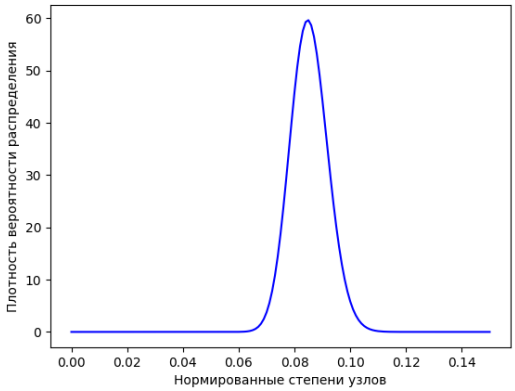


Рис. 3. Дифференциальная функция распределения

Таблица 2. – Степени узлов графа

Узел	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Степень узла	17	18	13	15	18	16	15	17	14	18	14	19	17	16	19	19	15	15	14	15

Таблица 3. – Нормированные степени узлов графа

Узел	1	2	3	4	5	6	7	8	9	10
Степень узла	0.09	0.09	0.07	0.08	0.09	0.08	0.08	0.09	0.07	0.09
Узел	11	12	13	14	15	16	17	18	19	20
Степень узла	0.09	0.1	0.09	0.08	0.1	0.1	0.08	0.08	0.07	0.08

Таблица 4. – Статистические характеристики графа и показателя стабильности

Среднее значение нормированной степени узлов	**Среднеквадратическое отклонение нормированной степени узлов**	**Показатель стабильности *Stb***
0.084	0.0099	0.3

5. Заключение

Предложен подход к анализу стабильности многомерных систем на основе топологических характеристик графов, соответствующих разным видениям системы. Фокусом подхода является формальная схема вычисления сопоставимых статистических оценок, инвариантных к размерам графов. Основу подхода составляет формирование выборок относительных степеней узлов графа, которые рассматриваются как однородные выборки независимых случайных величин. Получение сопоставимых оценок стабильности, соответствующих разным видениям системы, делает возможным формирование обобщенных показателей системы, а также сравнительный анализ стабильности свойств системы, соответствующих разным видениям.

Благодарности

Работа поддержана Министерством науки и высшего образования Российской Федерации в рамках базовой части государственного задания для высших учебных заведений # FEUE 2023-0007.

Список используемых источников

1. Latora V., Marchiori M. Efficient Behavior of Small-World Networks//Phys. Rev. Lett. 87, 198701 – Published 17 October 2001. URL: https://journals.aps.org/prl/abstract/10.1103/PhysRevLett.87.198701 (дата обращения: 01.09.2025). (электронный ресурс).
2. Михайлов Р.Л., Макаренко С.И. Оценка устойчивости сети связи в условиях воздействия на неё дестабилизирующих факторов // Радиотехнические и телекоммуникационные системы, 2013. №4. С. 69-79. (статья из журнала).
3. Battiston F., Nicosia V., Latora V. Structural measures for multiplex networks // Physical Review E 89(3-1):03280. P. 1-16. DOI:10.1103/PhysRevE.89.032804 (статья из журнала).
4. Wu H., James R.G., Crutchfield J.P., D'Souza R.M. Correlated structural evolution within multiplex networks // Journal of Complex Networks, 2020. P. 1-30p. DOI:10.1093/comnet/cnaa014 (статья из журнала).
5. Siganos G., Faloutsos M., Faloutsos P., Faloutsos C. Power-Laws and the AS-level Internet Topology // IEEE/ACM Transactions on Networking. Vol. 11, Issue 4, 2003. P. 514 – 524. DOI: 10.1109/TNET.2003.815300 (статья из журнала).
6. Прогнозирование технического состояния и надежности радиоэлектронной аппаратуры / Д.В. Гаскаров, Т.А. Голинкевич, А.В. Мозгалевский //М.: «Сов. Радио», 1974. – 224с. (книга).
7. Адаптивное управление и интеллектуальный анализ данных информационных потоков в компьютерных сетях / Тимофеев А.В. – СПб.: ООО «Анатолия», 2012. – 280с. (книга).
8. Статистическое исследование территориальных систем / М.Б. Гузаиров, В.Е. Гвоздев, Б.Г. Ильясов и др. М.: Машиностроение, 2008. – 187 с. (книга).
9. Гвоздев В.Е., Гузаиров М.Б., Давлиева А.С., Галимов Р.Р. Оценка характеристик информационной безопасности радиосети MANET на основе анализа топологий связей // Научный журнал «Доклады ТУСУР». – 2023, Выпуск №4. Т.26. С. 35-43. (статья из журнала).
10. Gvozdev V.E., Davlieva A.S., Galimov R.R. Reliability Analysis of Systems with Dynamic Structure Based on the Properties of Nodes and Channels // Advances in Science, Technology and Engineering Systems, RusAutoCon 2023, pp. 425-430. (материалы конференции).

Методы синтеза сложных погодных условий на изображениях для обучения моделей компьютерного зрения

Федосеев Е. П.
Воронежский Государственный Университет
Воронеж, Россия
e-mail: fedo5eeW@yandex.ru

Иванков А. Ю.
Воронежский Государственный Университет
Воронеж, Россия
e-mail: ivankov@cs.vsu.ru

Аннотация[1]

В данной работе проведены исследование и реализация метода аугментации данных для повышения производительности моделей компьютерного зрения при работе в условиях затруднённой видимости. В частности, рассмотрен метод моделирования эффектов тумана и снегопада на изображениях дорожных сцен. Проведён анализ возможных подходов к оценке трёхмерной структуры сцены для получения карты глубины, показывающей степень отдалённости объектов сцены от камеры. Рассмотрен принцип работы нейросетевой модели оценки глубины на основании одного изображения. Разработан алгоритм синтеза погодных условий, связанных с явлением атмосферного рассеяния. Алгоритм основан на использовании карты глубины для расчёта карты прозрачности, то есть определения степени видимости для каждого конкретного участка изображения в зависимости от расстояния между этим участком и камерой. Для оценки требуемой яркости применяется метод тёмного канала, основанный на оценке средней яркости самых ярких участков изображения, в результате работы которого может быть получена оценка интенсивности естественного освещения. Проведён эксперимент по применению рассматриваемого метода аугментации для до обучения модели распознавания объектов. Рассмотренный алгоритм был применён для набора данных с дорожными сценами, размеченного для задачи детекции. Проверка точности работы модели на наборе изображений, полученных в реальных условиях, показала, что наличие синтетических данных в обучающей выборке улучшает точность распознавания и локализации объектов на изображении.

Ключевые слова: аугментация, затруднённая видимость, доменный сдвиг, монокулярная оценка глубины, распознавание объектов, компьютерное зрение.

1. Введение

Одной из основных характеристик работы моделей машинного зрения, выполняющих задачи распознавания, сегментации, трекинга объектов, является робастность, то есть способность модели адаптироваться к изменениям в распределении входных данных и не терять в точности и надёжности работы.

Особенно это актуально для моделей, работающих в реальных условиях, когда существует вероятность возникновения искажений естественной природы, таких как туман, задымление, снегопад. Визуальные артефакты, возникающие в этих условиях, ухудшают видимость объектов, в результате чего точность работы падает.

Очевидное решение – сбор и подготовка большого и разнообразного набора тренировочных данных, который покрывал бы все возможные варианты естественных искажений, однако редкость перечисленных явлений, а также техническая сложность сбора данных в таких условиях делают получение реальных данных слишком трудоёмкой, дорогостоящей или невозможной задачей.

В качестве альтернативного подхода предлагается разработать методы моделирования сложных погодных условий для подготовки синтетических данных и использования их на этапе обучения.

Целью данной работы является разработка методов моделирования явлений тумана и снегопада для аугментации данных при обучении моделей компьютерного зрения.

2. Методы и материалы

2.1. Пространственная информация о сцене

Явление атмосферного рассеяния

[1] Труды XI Международной научной конференции "Информационные технологии интеллектуальной поддержки принятия решений", 11-13 ноября, Уфа-Чандигарх-Фандербейлпарк, 2025

Закон атмосферного рассеяния описывает явление, связанное с нахождением в воздухе мелких частиц, которые ухудшают видимость объектов пропорционально их удалённости от камеры. Особенно сильно этот эффект проявляется в случае, когда рассматриваются сцены с глубокой перспективой, например, дорожные сцены. В таком случае самые отдалённые объекты могут становиться практически невидимыми.

На рис.1 представлена схема физического явления атмосферного рассеяния.

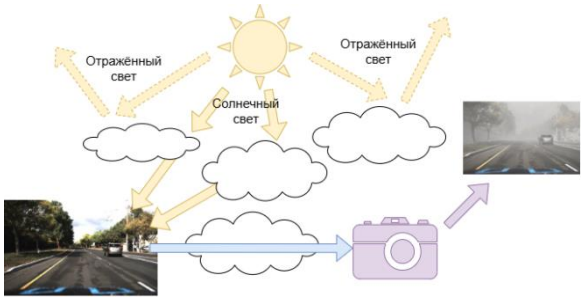


Рис. 1. Физическая схема явления атмосферного рассеяния

Свет от источника частично рассеивается частицами, взвешенными в воздухе, из-за чего общий уровень освещённости становится ниже. Свет, отражённый от объектов сцены, также рассеивается частицами в воздухе между объектами и камерой.

Из этого наблюдения следует, что, например, в случае явления равномерного тумана ухудшение видимости зависит от отдалённости объектов от точки наблюдения: чем дальше объект, тем больше частиц находится в пространстве между ним и камерой, и, следовательно, тем более густой туман и менее различим сам объект.

Предлагаемый метод обработки данных и проведение исследования его эффективности можно представить в виде следующих основных этапов:

- Получение информации о трёхмерной структуре сцены;
- Генерация синтетических данных на основе модели атмосферного рассеяния;
- Анализ влияния синтетических данных в обучающей выборке на качество работы модели на реальных данных.

Получение информации о трёхмерной структуре сцены

Задача получения трёхмерной информации о сцене в применении к задаче моделирования тумана сводится к построению карты глубины изображения.

Карта глубины — одноканальное изображение или канал изображения, содержащий информацию о расстоянии от поверхностей объектов сцены до точки наблюдения. В компьютерной графике карта глубины также называется буфером глубины или Z-буфером. Каждому пикселю исходного изображения ставится в соответствие значение в карте глубины, которое означает расстояние до конкретной точки. Это расстояние может быть выражено двумя разными способами, то есть существует два вида карт глубины: абсолютная и относительная. Их отличие состоит в том, что в абсолютной карте глубины расстояние выражается в реальных физических единицах, например, в метрах или сантиметрах, в то время как в относительной карте значения глубины основаны на разнице между самой дальней и самой близкой точками изображения и находятся в пределах от 0 до 1.

Подходы к построению карты глубины

Задача оценки трёхмерной структуры сцены решается различными способами, использующими разные подходы. Перечислим основные из них:

1. Стереоскопическая оценка глубины – математический метод, основанный на использовании системы из двух или более камер [1]. С помощью камер делаются снимки сцены из разных точек наблюдения, полученные снимки выравниваются относительно друг друга, затем выполняется сопоставление пикселей и вычисляется диспаритет – величина, характеризующая разницу между координатами соответствующих пикселей на двух изображениях [2]. Дальность до объекта определяется как величина, обратно пропорциональная диспаритету.
2. Методы с использованием дополнительно оборудования. В эту категорию включаются, например, лазерные и светодиодные сканеры (LiDAR, ToF-камеры), в которых оценка дальности происходит путём измерения времени от испускания импульса до его фиксации датчиком после отражения от объекта. Такие методы позволяют получить очень точную информацию о трёхмерной структуре сцены, однако часто являются недоступными вследствие дороговизны и технической сложности.
3. Монокулярная оценка глубины (MDE) – метод оценки глубины на основании только одного входного RGB-изображения. В этом подходе чаще всего применяются глубокие нейронные сети, обучаемые на больших объёмах размеченных данных для решения задачи регрессии значения глубины каждого пикселя исходного изображения.

В данной работе мы фокусируемся на использовании последнего подхода. В настоящее время модели монокулярной оценки глубины являются наиболее актуальным направлением исследований, так как оно ориентировано на упрощение сбора данных и работу в режиме

реального времени без затрат на специальное оборудование.

Используемая модель

Рассмотрим нейросетевую модель монокулярной оценки глубины Depth Anything V2 [3], в которой был предложен принципиально новый подход к подготовке данных для обучения. Сделав акцент на обработке мелких деталей и других сложных элементов изображений, авторы исследования добились наилучшей производительности в задаче zero-shot оценки глубины среди аналогичных моделей.

При обучении применялся приём дистилляции знаний, который заключается в последовательном обучении двух моделей: модели-учителя с большим количеством параметров и облегчённой модели-ученика. Основная цель модели — формирование знаний на основе размеченных данных (в данном случае используются синтетические изображения), а затем выполнение предсказания глубины на большом наборе реальных изображений для получения псевдоразметки, на которой в дальнейшем обучается модель-ученик. Этот приём позволяет получить модель, которая сохранит точность исходной большой моде ли, но при этом будет работать быстрее и эффективнее.

Синтетические данные для обучения модели-учителя были получены авторами работы [3] при использовании трёхмерных игровых движков: набор данных состоял из разнообразных смоделированных сцен и соответствующих им относительных карт глубины. Основная цель использования синтетических данных – получение максимально точной разметки с высокой детализацией, чего нельзя добиться при сборе реальных данных, в которых всегда в некоторой степени присутствуют шум и искажения.

При обучении модели-ученика также использовались преобразования данных, например, комбинация участков нескольких изображений на одном тренировочном образце, искажения цвета и зашумление. Такой подход позволил не только перенести знания большой модели, но и превысить её базовую точность

2.2. Генерация тумана

В цифровой обработке изображений эффект тумана описывается при помощи математической модели, называемой моделью атмосферного рассеяния или законом Кошмидера [4].

$$I(x) = J(x)\tau(x) + A(1 - \tau(x))$$

где I(x) – изображение с туманом, J(x) – исходное изображение без тумана, A – атмосферный свет, τ(x):

$$\tau(x) = e^{-\beta d(x)}$$

где β – скалярный коэффициент, характеризующий густоту тумана, d(x) – карта глубины изображения.

Карта прозрачности показывает степень рассеяния света взвешенными в воздухе частицами, поэтому для объектов, находящихся далеко от точки наблюдения, значения в карте прозрачности будут низкими, а для находящихся близко — высокими. Участки изображения с низкими значениями прозрачности характеризуются уменьшением контрастности и увеличением яркости: отдалённые объекты становятся менее различимыми и более размытыми, а туман, вследствие рассеяния солнечного света, приобретает белый цвет.

Карта прозрачности рассчитывается на основе карты глубины, которую мы получаем для каждого изображения при помощи модели монокулярной оценки глубины, дообученной для абсолютной оценки на наборе данных KITTI. На рис.2 приведён пример работы алгоритма.



Рис. 2. Пример карты глубины

Параметр β при расчёте карты прозрачности характеризует густоту тумана и подбирается экспериментально в зависимости от желаемого уровня видимости. В данной работе при синтезе мы используем случайные значения параметра β в пределах от 1.5 до 3.5, что обеспечивает получение множества разнообразных условий — от лёгкой дымки до сильного тумана. На рис.3 приведены примеры результата работы алгоритма для разных значений β.



а) б)



в) г)

Рис. 3. Примеры генерации тумана с различными значениями параметра β: а) β = 0, б) β = 1.5, в) β = 2.0, г) β = 4.0

Параметр атмосферного света A описывает степень яркости моделируемого тумана и зависит от общего уровня освещённости исходной сцены.

В предлагаемом методе и реализуемом на его основе алгоритме атмосферный свет оценивается при помощи метода тёмного канала (Dark Channel Prior). Идея данного алгоритма, который был изначально предложен в [5] как метод фильтрации тумана, заключается в наблюдении, что в ясной сцене на всех участках изображения, за исключением неба, хотя бы в одном канале обязательно найдётся малое относительно остальных каналов значение интенсивности. Основываясь на этом, можно построить карту тёмного канала и определить области изображения, где это правило не выполняется. Эти области будут соответствовать участкам с туманом, которые затем можно соответствующим образом обработать, чтобы снизить степень искажения. Этапы работы алгоритма:

1. Определяется минимальное значение интенсивности из трёх цветовых каналов для каждого пикселя;

2. К полученной карте тёмного канала применяется морфологическая операция эрозии с прямоугольным ядром. Эрозия позволяет устранить шум и выбросы (мелкие детали изображения с высокой яркостью), делая тёмный канал более равномерным;

3. Из карты тёмного канала выбирается определённая доля самых ярких пикселей и по среднему значению интенсивности соответствующих пикселей вычисляется параметр A.

Описанный способ позволяет моделировать эффект тумана, яркость которого соответствует сцене. Примеры обработанных данных представлены на рис.4:



Рис. 4. Примеры работы алгоритма: слева – исходные изображения, справа – результаты обработки

2.3. Генерация снегопада

Описание явления снегопада

Явление снегопада является одним из самых сложных погодных факторов, с которыми приходится работать системам компьютерного зрения в реальных условиях. Основные причины:

− Снегопад добавляет шум, что создаёт на изображении большое количество хаотичных помех. Близкие к камере снежинки становятся различимыми и создают точечные помехи, которые в разных условиях могут быть похожи на аддитивный или импульсный шум;

- Ухудшает видимость как далёких, так и близких к точке наблюдения объектов. Внешний вид и эффект снегопада в сильно отдалённых от камеры участках сцены похож на туман, так как падающие снежинки являются взвесью мелких частиц в воздухе, которые также по закону Кошмидера [4] рассеивают свет;

- Искажает освещение. При сильном снегопаде мелкие твёрдые частицы могут преломлять свет, создавать блики и вызывать эффект переэкспонирования, из-за чего на изображении теряются детали и контраст;

- Меняет текстуры сцены. Сильный снегопад может скрывать или менять форму объектов: скрывать дорожную разметку, дорожные знаки, другие движущиеся объекты;

Для физически корректного моделирования снегопада необходимо рассмотреть, какими параметрами он описывается.

Было проведено множество исследований на стыке дисциплин оптики, метеорологии и цифровой фотографии, направленных на изучение физических параметров явления снегопада. Например, в работе [6] была исследована зависимость между интенсивностью осадков и массовой концентрацией снега, и было показано, что зависимость является степенной и соответствует следующему виду:

$$I = k \cdot C^n$$

где I – интенсивность осадков, C – концентрация снега, k – коэффициент, зависящий от таких параметров как размер, масса, форма и плотность снежинок, а также температура и влажность воздуха, определяющие взаимодействие и физику движения частиц, n – показатель степени, обычно близкий к единице, который при больших значениях концентрации (например, в случае крупного рыхлого снега) принимает значение больше 1 и увеличивает значение интенсивности осадков, делая зависимость от концентрации степенной.

Интенсивность осадков – величина, показывающая количество выпавших осадков в единицу времени. Обычно измеряется в миллиметрах в час, где одному

миллиметру осадков соответствует один литр воды на один квадратный метр поверхности. В случае снега интенсивность осадков в метеорологических исследованиях может быть рассчитана через водный эквивалент или массу выпадающего снега в реальном времени.

Концентрация снега – величина, которая показывает количество твёрдых осадков, содержащееся в единице объёма воздуха. Обычно измеряется в граммах на кубический метр.

Полученная авторами работы [6] зависимость позволяет оценить параметр концентрации снега, исходя из желаемого уровня интенсивности осадков. Были предложены следующие приближения для коэффициента k:

$$C = 0.3 \cdot I$$
$$C = 0.47 \cdot I$$

где первое выражение соответствует случаю сильного и плотного снегопада, а второе справедливо для случая обычного снега. Принимая за 0.2 грамма среднее значение массы одной снежинки, можно вычислить количество снежинок на единицу объёма:

$$N = \frac{C}{0.2}$$

Следующим необходимым параметром, которым описывается явление снегопада, является размер каждой отельной частицы снега. Так как не все снежинки имеют одинаковый размер, должно быть введено распределение, описывающее вариацию их размеров. Основываясь на работе [7], авторы которой провели большое исследование относительно размерного распределения снежных хлопьев и его зависимости от различных факторов, можно записать выражение для относительной вероятности образования снежинки с диаметром D:

$$N_D = N_0 \cdot e^{-\Lambda D}$$

где $N_0 = 2.5 \times 10^3 \cdot I^{-0.94}$ – параметр, связанный с общей числовой концентрацией и зависящий от общей интенсивности снегопада и влажности воздуха, $\Lambda = 22.9 \cdot R^{-0.45}$ – параметр, обратно пропорциональный размеру снежинок и показывающий зависимость размера снежных хлопьев в различных температурных условиях: при тёплых снегопадах снежинки слипаются в крупные агрегаты, а при более низкой температуре преобладают мелкие кристаллы.

Используя частично определённое взвешенное распределение вероятности для значений диаметров снежинок, можно определить подходящий диаметр для каждой генерируемой снежинки.

Алгоритм моделирования явления снегопада

Для моделирования погодного явления снегопада был реализован алгоритм с применением карт глубины изображения и системы частиц в OpenGL с трёхмерной анимацией для имитирования эффекта ветра и движения камеры с различными параметрами.

Сначала для входного обрабатываемого изображения получается карта глубины при помощи нейросетевой модели монокулярной оценки. Внешний вид снегопада в контексте сцены зависит от её трёхмерной структуры, глубины перспективы и взаимного расположения объектов.

Первым этапом обработки является восстановление сцены в трёхмерном пространстве по её карте глубины. Для этого для каждого пикселя изображения создаётся его проекция в трёхмерное пространство, исходя из соответствующего значения в карте глубины. В этой точке пространства создаётся прямоугольник, состоящий из двух прямоугольных треугольников, расположенный параллельно плоскости оригинального изображения. Прямоугольнику присваивается значение цвета соответствующего пикселя, а также устанавливается масштаб, необходимый для рендера прямоугольника в ровно одном пикселе выходного изображения.

Параллельно с реконструкцией сцены создаётся система частиц. Рассмотрим процесс поэтапно:

- В пространстве перед камерой создаётся куб, который случайно заполняется частицами. Для этого создаётся вершинный буфер с координатами точек в трёхмерном пространстве, которые случайно получаются из генератора случайных чисел в соответствии с равномерным распределением вероятности. Количество точек для генерации является настраиваемым параметром, позволяющим получать разнообразные условия снегопада. Настройка параметра, как и последующее определение количества снежинок, происходит в соответствии с выражениями для числа частиц;
- Размеры куба с частицами определяются исходя из разрешения входных изображений, требований производительности и времени рендеринга. В качестве меры для оптимизации вычислений можно использовать рассмотренный выше метод синтеза тумана, так как в условиях реального снегопада в отдалённых от камеры участках сцены отдельные снежинки становятся неразличимыми и рассеивают свет так же, как дымка, и поэтому отрисовывать отдельные частицы в этих участках не является необходимостью;
- Для каждой частицы применяется реалистичная текстура снежинки;
- Для создания эффекта съёмки в движении используется анимация движения частиц, поэтому для этого движения можно задать различные физические параметры. В качестве базовой физики

движения частиц задаётся вертикальное свободное падение, также можно учитывать и регулировать скорость ветра посредством задания скорости по осям x, y и z. Настраиваемым параметром также является скорость движения камеры, от которой зависит степень размытия снега в движении;

- Анимация частиц реализуется в вершинном шейдере. На вход программа получает значения координат частиц в пространстве, а в качестве глобальных параметров используются model, view и projection матрицы и параметры скорости движения, описанные выше. Глобальные переменные создаются из основной программы и хранятся на GPU внутри объекта program;

- Проходит цикл по количеству кадров, в ходе которого происходит рендеринг сцены с изменяющимся с течением времени положением частиц. Полученные кадры в процессе циклического рендеринга хранятся в памяти, а затем считается их взвешенная сумма для получения эффекта размытия в движении.

2.4. Описание эксперимента

Для исследования эффективности метода рассмотрим моделирование тумана на наборе данных Self-Driving Cars и используем синтетические данные для обучения модели распознавания объектов.

Набор данных включает 18000 изображений дорожных сцен, размеченных для задачи распознавания объектов. Объекты относятся к 5 классам, все изображения имеют размер 480×300. Случайная выборка из 9000 изображений была обработана при помощи рассматриваемого алгоритма, остальные 9000 изображений не изменялись.

Для тестирования модели были использованы реальные изображения дорожных сцен со сложными погодными условиями из набора данных DAWN [8].

В качестве модели распознавания объектов использовалась модель FasterRCNN [9], предобученная на наборе данных ImageNet. Данная архитектура относится к категории двухступенчатых, то есть алгоритм её работы состоит из двух основных этапов: сначала модель формирует предположения о регионах (областях поиска объектов), после чего предсказывает и уточняет координаты ограничивающих рамок объектов и определяет их принадлежность к тому или иному классу.

В сравнении с более современными методами — одноступенчатыми архитектурами (моделями YOLO) и детекторами на основе архитектуры Vision Transformer (моделями DETR) — FasterRCNN в определённых случаях показывает лучшую точность распознавания (разница особенно заметна при распознавании малых объектов), однако эта модель работает медленнее в силу архитектурной необходимости двухэтапного прохода изображения через сеть для формирования предсказаний, и поэтому редко применяется для обработки данных в режиме реального времени.

Обучение проводилось в два этапа: на наборе исходных данных, а затем на обработанных при помощи рассматриваемого метода. Гиперпараметры обучения: размер мини-батча – 8, количество эпох – 5, оптимизатор – Adam, скорость обучения – 0.005 с параметром L2-регуляризации 0.005 и расписанием скорости обучения каждые три эпохи с параметром γ, равным 0.1.

3. Результаты эксперимента и их обсуждение

При синтезе тумана было проведено измерение времени обработки каждого образца. Результаты измерений показали, что на обработку RGB-изображения с размером 480×300, включающую проход образца через модель монокулярной оценки глубины и непосредственно моделирование тумана, в среднем требовалось 1.07 секунд. Обработка данных и измерение времени проводилось в облачной среде Kaggle Notebook с использованием вычислительных ресурсов платформы Kaggle на графическом процессоре NVIDIA Tesla T4 и центральном процессоре Intel Xeon с двумя ядрами. Оперативная память системы – 32 ГБ.

Полученные результаты тестирования двух моделей на наборе реальных данных приведены в табл. 1. В качестве метрики использовался показатель Mean Average Precision (mAP) с порогом Intersection over Union 50 и 75 процентов.

$$mAP = \frac{1}{n}\sum_{k=1}^{n} APk$$

$$APk = \frac{|TPk|}{|FPk| + |TPk|}$$

где n – число классов, k – индекс класса, $|TPk|$ – число верных классификаций k-го класса, $|FPk|$ – число ложноположительных классификаций k-го класса.

Таблица 1. – Показатели точности модели на реальных данных

Модель	mAP@50	mAP@75
Вариант 1	37.57	18.26
Вариант 2	44.83	28.85

Здесь обозначение Вариант 1 соответствует модели, обученной только на исходных данных, а Вариант 2 – на синтетических.

Аналогичные исследования, в которых используются другие архитектуры моделей оценки глубины и распознавания объектов, также подтверждают улучшение качества распознавания. Например, в работе [10] для монокулярной оценки была использована модель Monodepth2, а в качестве модели распознавания объектов — YOLOv4. Производительность модели оценивалась на наборах данных WAYMO и Foggy Driving, метрики приведены в табл. 2.

Таблица 2. – Метрики mAP моделей из

Модель	WAYMO-haze	Foggy Driving
YOLOv4	36.17	72.54
YOLOv4-Synt	41.91	77.88

Здесь YOLOv4 — модель, обученная на исходных данных, а YOLOv4-Synt — на синтетических.

На рис.5 представлены результаты работы моделей на нескольких изображениях из набора данных DAWN. Как видно, модель, обученная на синтетических данных, чаще распознаёт отдалённые объекты и точнее определяет координаты ограничивающих рамок объектов.



Рис. 5. Сравнение результатов работы моделей распознавания: слева – модель, обученная на исходных данных, справа – на синтетических.

4. Заключение

- В данной работе был рассмотрен и реализован метод синтеза данных на основе реальных изображений для повышения точности работы моделей компьютерного зрения.

- Предложен метод синтеза эффектов тумана и снегопада на основании монокулярной оценки глубины изображения и закона атмосферного рассеяния. Алгоритм включает оценку трёхмерной структуры сцены с использованием модели монокулярной оценки глубины, оценку общей освещённости изображения методом тёмного канала, моделирование эффекта тумана в соответствии с картой глубины и параметрами плотности и яркости тумана.

- Проведён эксперимент, в результате которого было показано, что использование синтетических данных при обучении модели распознавания объектов повышает точность работы на реальных изображениях, полученных в условиях естественного тумана.

- В ходе выполнения эксперимента было также проведено измерение времени работы предлагаемого алгоритма. Измерения показали, что для данного набора данных среднее время обработки одного образца примерно равно одной секунде, что делает проблематичным использование данного метода аугментации в онлайн-режиме, когда во время обучения данные обрабатываются непосредственно перед прямым проходом через нейронную сеть. Самым затратным по времени и вычислительно сложным этапом обработки является оценка глубины изображения, поэтому в дальнейшем планируется исследовать возможности для получения более оптимизированной модели для решения задачи zero-shot оценки глубины, которая при этом удовлетворяла бы требованиям точности.

Список используемых источников

1. Zhang J. MobiDepth: Real-Time Depth Estimation Using On-Device Dual Cameras / J. Zhang [et al.] // MobiCom '22: Proceedings of the 28th Annual International Conference on Mobile Computing and Networking. - 2022. - P. 528-541. - DOI: 10.1145/3495243.3560517.
2. Godard C. Unsupervised Monocular Depth Estimation with Left-Right Consistency / C. Godard, O.M. Aodha, G.J. Brostow // 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA - 21-26 July 2017. - DOI: 10.1109/CVPR.2017.699.
3. Yang L. Depth Anything V2 / L. Yang [et al.] // ArXiv. - 2024. - Режим доступа: https://arxiv.org/abs/2406.09414.
4. Tang M. A Foggy Weather Simulation Algorithm for Traffic Image Synthesis Based on Monocular Depth Estimation / M. Tang, Z. Zhao, J. Qiu // Sensors - 2024. - 24. - № 6. - DOI: 10.3390/s24061966
5. He K. Single Image Haze Removal Using Dark Channel Prior / K. He, J. Sun, X. Tang // 2009 IEEE Conference on Computer Image and Pattern Recognition, Miami, FL, USA. - 2009. - DOI: 10.1109/CVPR.2009.5206515.
6. Koh G. Snow mass concentration and precipitation rate / G. Koh, J. Lacombe, D. L. Hutt // Cold Regions Science and Technology. – 1988. – Т. 15, № 1. – P. 89-92.
7. Gunn K. L. S. The Distribution with Size of Aggregate Snowflakes / K. L. S. Gunn, J. S.

Marshall // Journal of Meteorology. – 1958. – T. 15, № 5. – P. 452–461.

8. Kenk M. DAWN: Vehicle Detection in Adverse Weather Nature / M. Kenk, M. Hassaballah // ArXiv. - 2020. - Режим доступа: https://arxiv.org/abs/2008.05402.
9. Ren S. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks / S. Ren [et al.] // ArXiv. - 2015. - Режим доступа: https://arxiv.org/abs/1506.01497.
10. Tran L. Toward Improving Robustness of Object Detectors against Domain Shift / L. Tran [et al.] // IEEE International Conference on Green Energy, Computing and Sustainable Technology (GECOST). - February 2024. - DOI: 10.1109/GECOST60902.2024.10474777.

Интеллектуальная поддержка принятия решений при выборе квалификации сотрудника для обслуживания заявок

Дегтярева И. В.
Уфимский университет науки
и технологий
Институт экономики, управления
и бизнеса
Уфа, Россия
e-mail: degtyareva.iv@ugatu.su

Макарова Е. А.
Уфимский университет науки
и технологий
Институт информатики, математики
и робототехники
Уфа, Россия
e-mail: ea-makarova@mail.ru

Пономарёв Н. В.
Уфимский университет науки
и технологий
Институт информатики, математики
и робототехники
Уфа, Россия
e-mail: nikita_ponomarev_2015@inbox.ru

Аннотация[1]

В статье рассматривается задача разработки интеллектуальной системы поддержки принятия решений при обработке заявок на решение задач сопровождения компьютерных программ. Система содержит подсистему интеллектуального анализа данных о результатах обслуживания заявок на решение задач сопровождения компьютерных программ и подсистему нечеткого логического вывода для формирования рекомендаций при принятии решений по назначению специалиста. С помощью подсистемы интеллектуального анализа данных проведен компонентный анализ ретроспективных данных о результатах обслуживания заявок. Построены деревья решений и сформированы кластерные зоны по степени соответствия квалификации исполнителя степени сложности задачи в заявке. Выявлены кластерные зоны с признаками некорректного распределения задач, где наблюдается значительное несоответствие квалификации специалистов степени сложности поступающих заявок. Разработана подсистема нечеткого логического вывода и построены нечеткие правила формирования рекомендаций по требуемой квалификации специалиста для решения задачи. Предложен эвристический алгоритм для составления правой части нечетких правил, который основан на анализе принадлежности заявки к одной из выделенных зон. Разработка нечетких правил проведена с учетом выявленных корректировок в назначении специалиста для решения задач. Построены поверхности нечеткого логического вывода и выполнен их анализ. Применение разработанной системы позволит обеспечить повышение эффективности решения задач на сопровождение компьютерных программ в виде снижения финансовых и временных затрат.

Ключевые слова: интеллектуальный анализ; принятие решений, главная компонента; кластер; нечеткое правило

1. Введение

Решение проблем распределения трудовых ресурсов при выполнении работ требует разработки систем поддержки принятия решений, которые могут обеспечить формирование целевых ориентиров при многокритериальном принятии решений человеком по назначению исполнителей на различные работы [1,2].

Исследования в области разработки систем поддержки принятия решений (СППР) для решения задач о распределении трудовых ресурсов основаны на применении различных математических методов и моделей: оптимизационных, динамических, агент-ориентированных, эвристических, интеллектуальных. Применение нечеткого метода анализа иерархий в задачах о назначениях позволяет

[1] Труды XI Международной научной конференции "Информационные технологии интеллектуальной поддержки принятия решений", 11-13 ноября, Уфа-Чандигарх-Фандербейлпарк, 2025

учесть неопределенность в суждениях экспертов [3]. Решение многокритериальных задач применяется для назначения сотрудников в условиях учета множества таких факторов, как способности, уровень рабочей нагрузки, эмоциональная составляющая [4-7]. Известны примеры применения мультиагентного подхода для распределения задач между агентами [8].

Широко применяются методы интеллектуального анализа данных и машинного обучения для выявления групп сотрудников с похожими паттернами и классов задач по уровню сложности, а также для прогнозирования эффективности сотрудников и команд [9-12]. В работе предлагается применить методы интеллектуального анализа данных для формирования кластеров ранее обслуженных заявок на сопровождение компьютерных программ, что послужит основой для разработки правил поддержки принятия управленческих решений по назначению специалиста. В перспективе эти знания могут быть использования для построения динамических агент-ориентированных моделей оценки эффективности обслуживания заявок [13,14].

Целью исследования является разработка интеллектуальной системы поддержки принятия решений (ИСППР) при обработке заявок на решение задач сопровождения компьютерных программ, которая позволяет проводить интеллектуальный анализ ретроспективных данных о результатах обслуживания заявок, а также выдавать рекомендации по назначению специалистов для решения поставленных задач в новых входящих заявках. Для анализа применяются данные о результатах обслуживания заявок на сопровождение программ платформы 1С в компании «1С: Апрель Софт».

Разрабатываемая ИСППР включает подсистему интеллектуального анализа данных результатах обслуживания заявок на решение задач сопровождения компьютерных программ и подсистему нечеткого логического вывода для формирования рекомендаций при принятии решений по назначению специалиста.

2. Процедура разработки ИСППР

Предложена процедура разработки НСППР при формировании рекомендаций по назначению специалистов при решении задач сопровождения компьютерных программ. Особенность предложенной процедуры состоит в том, что она основана на ранее полученных результатах исследования, связанных с проведением интеллектуального анализа *исходной* выборки D1. Выборка включала в себя признаки, характеризующие степень сложности основной и дополнительной задачи, а также уровень оснащенности клиента и уровень квалификации специалиста, назначенного для решения задачи. Проведенное исследование выявило как сильные, так и слабые стороны текущей практики назначения специалистов, что позволило определить области для дальнейшего изучения.

К положительным результатам анализа необходимо отнести выявленный ряд закономерностей в данных: четкое разделение заявок по уровню сложности и времени выполнения задач; наличие групп кластеров с высоким соответствием квалификации исполнителя степени сложности заявки; возможность идентификации кластеров с потенциальным несоответствием «квалификация специалиста – сложность задачи»; выделение отдельного кластера сложных проблемных заявок, которых присуща непоследовательность в назначении исполнителя.

Выявленные проблемы состоят в следующем. Во-первых, это некорректные назначения специалистов для решения задач: Во многих кластерах встречаются ситуации несоответствия компетенций специалиста уровню сложности задачи. Во-вторых, это хаотичное распределение ошибок в назначении специалиста. Заявки с некорректными решениями не сконцентрированы, а рассредоточены по разным кластерам. В-третьих, это невозможность диагностики типа ошибки, поскольку выявленная модель данных не позволила дифференцировать ситуации, в которых квалификация специалиста является либо избыточной (неэффективное использование ресурсов), либо, что критичнее, недостаточной (риск того, что задача не будет решена).

Представленные выводы позволяют обосновать необходимость формирования *расширенной* выборки данных D2. Предложено включить в выборку дополнительные признаки, которые характеризуют степень корректности или некорректности ранее принятых решений. Эти признаки должны быть основаны на экспертной оценке результатов решения задач в прошлом, а также учитывать контекст задачи в условиях возможного искажения информации на этапе первичного оформления заявки из-за недостаточной квалификации консультантов на линии консультаций: Поэтому предложено ввести два новых признака для оценки характера и степени ошибочности в назначении специалиста, один из которых определяет степень несоответствия в случае недостаточной квалификации, а другой – в ситуациях «избыточной» квалификации, когда назначается специалист более высокой квалификации, чем требуется для решения задачи.

Предложенная процедура разработки ИСППР включает три этапа. На первом этапе выполняется формирование расширенной обучающей выборки D2 на основе результатов ранее проведенного интеллектуального анализа данных в выборке D1. На втором этапе проводится разработка подсистемы интеллектуального анализа данных на основе применения метода главных компонент и метода построения деревьев решений. На третьем этапе

выполняется разработка подсистемы нечеткого логического вывода (НЛВ), что предполагает переход к построению нечетких правил принятия решений по назначению специалиста и формирование базы нечетких правил.

3. Формирование расширенной обучающей выборки

На первом этапе формируется расширенная выборка D2 путем оценки каждой заявки (строки в обучающей выборки) и решения, принятого по ней в ретроспективе. Для оценки введены два признака Dkv1 и Dkv2.

Первый признак Dkv1 показывает степень достаточности уровня квалификации специалиста и вводится для описания ситуаций, в которых есть опасность принятия решений о назначении специалиста с более низкой квалификации, чем требуется для решения задачи. Этот признак имеет порядковую шкалу от 0 до 10, где 10 – предполагает полное соответствие, а снижение значений свидетельствует о возникновении и усугублении ситуации недостаточности квалификации специалиста для успешного решения задачи. Название признака Dkv1 – достаточность квалификации специалиста.

Второй признак Dkv2 также показывает степень достаточности уровня квалификации специалиста, но вводится для описания противоположных ситуаций, то есть ситуаций, в которых возможно некорректное принятие решений о назначении специалиста с более высокой квалификацией, чем требуется для решения задачи. Этот признак также имеет порядковую шкалу от 0 до 10, где 10 – предполагает полное соответствие, а снижение значений свидетельствует о возникновении и усугублении ситуации ненужного превышения квалификации специалиста по сравнению со сложностью задачи в заявке. Название признака Dkv2 – обоснованность привлечения высококвалифицированного специалиста.

Разработан алгоритм формирования значений признаков Dkv1 (достаточность квалификации) и Dkv2 (обоснованность назначения высококвалифицированного специалиста). Алгоритм основан на использовании результатов интеллектуального анализа ретроспективных данных без учета их экспертного оценивания, составляющих выборку D1. Особенности алгоритма на примере формирования признака Dkv1состоят в следующем (рис. 1).

Во-первых, начальные значения признаков Dkv1 и Dkv2 равны 10. Это говорит о том, что специалист был подобран верно. Снижение значений признака ДКВ1 говорит о недостаточной квалификации специалиста, а снижение значений признака ДКВ2 говорит об избыточной квалификации специалиста.

Во-вторых, анализ признаков для каждой исторической заявки осуществляется в порядке, заданном значимостью главных компонент (ГК): сначала по главной компоненте ГК1, затем – по ГК2, ГК3 и ГК4. При этом каждое фактическое значение признака сравнивается со средним значением его по кластеру, выделенному при построении деревьев решений.

В-третьих, в результате сравнения может выполняться снижение значения признаков итогового балла по признакам Dkv1 и Dkv2 на 1 или 2 пункта от максимального значения 10. Для этого разработаны правила корректировки значений по всем признакам.

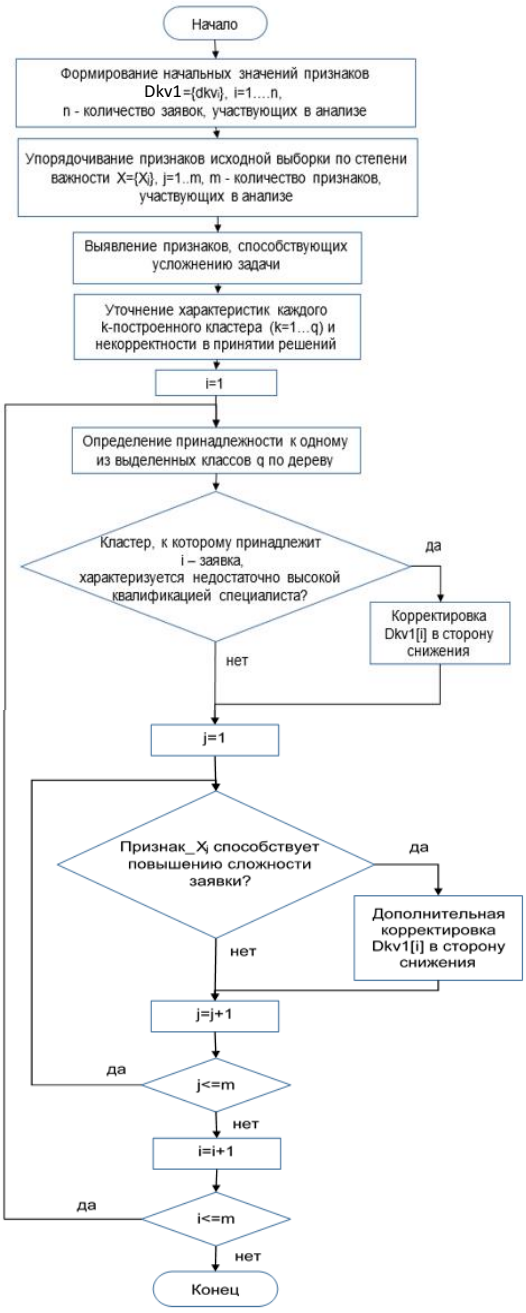


Рис. 1. Схема алгоритма формирования признака Dkv1 о назначении специалиста с недостаточной квалификацией

Таким образом, итоговые значения признаков Dkv1 и Dkv2, рассчитанные для каждой заявки, формируются как комплексная оценка, отражающая степень отклонения реальных характеристик выполненного назначения специалиста от

эталонных профилей, выявленных для каждого кластера обработанных заявок. Сформированная выборка представляет собой основу для проведения интеллектуального анализа данных о заявках и построения нечеткой СППР.

4. Подсистема интеллектуального анализа расширенной выборки и построение карты зон

Подсистема интеллектуального анализа данных строится на основе применения методов компонентного анализа и построения деревьев решений. Особенность проведения анализа данных состоит в формировании сепаратных кластеров сначала в плоскости компонент ГК1 и ГК2, далее в плоскости ГК1 и ГК3, далее в плоскости компонент ГК1 и ГК4. При каждом переходе к следующей плоскости границы кластеров и их характеристик уточняются. Анализ построенных главных компонент позволил сделать следующие выводы (табл. 1).

Во-первых, выделены важные признаки в главных компонентах. Показано, что в формировании первой компоненты ГК1 участвуют следующие признаки: тип дополнительной задачи, время дополнительной задачи. В названии ГК2 участвуют признаки: тип основной задачи, уровень знаний специалиста, недостаточная квалификация специалиста. В названии ГК3 участвуют признаки: уровень знаний клиента, время основной задачи, избыточная квалификация специалиста. В названии ГК4 участвуют признаки: мощность компьютера и время основной задачи, тип дополнительной задачи.

Таблица 1. Весовые коэффициенты признаков ГК (расширенная выборка)

	ГК 1	ГК 2	ГК 3	ГК 4
Время дополнительной задачи	**0,5347**	-0,0148	-0,0225	-0,3356
Время основной задачи	0,2889	0,1071	**-0,3740**	**0,5049**
Мощность ПК	0,2562	-0,1808	0,3230	**0,6373**
Достаточная/недостаточная квалификация специалиста, Dkv1	-0,3183	**0,3812**	0,1871	0,0758
Достаточная/избыточная квалификация специалиста, Dkv2	**-0,3551**	-0,0213	-0,3378	-0,2160
Тип дополнительной задачи	**0,5535**	0,0236	0,0288	-0,3404
Тип основной задачи	0,1055	**0,6258**	-0,0233	0,1616
Уровень знаний клиента	-0,0191	0,1738	**0,7616**	-0,1490
Уровень знаний специалиста	0,1407	0,6222	-0,1566	-0,0975

Во-вторых, сформулированы названия главных компонент (рис. 2). Название компоненты ГК1 – уровень сложности дополнительной задачи с учётом времени дополнительной задачи и избыточной квалификации специалиста (знак «–» при весовом коэффициенте); название компоненты ГК2 – уровень сложности основной задачи с учётом знаний специалиста и учётом недостаточной квалификации специалиста; название ГК3 – уровень знаний клиента с учётом времени основной задачи (–) и учётом избыточной квалификации специалиста (знак «–»); название ГК4 – уровень мощности ПК с учётом времени основной задачи и сложности дополнительной задачи (–). Знаки при весовых коэффициентах должны учитываться при построении кластеров

Построенная диаграмма рассеяния на плоскости первых двух главных компонент (ГК1-ГК2) позволила выделить семь кластеров (рис. 2). Шесть из них (кластеры 1-6) образуют диагональную последовательность, соответствующую ожидаемой закономерности: с ростом сложности задач (увеличение значений ГК1 и ГК2) повышается и квалификация назначаемых специалистов. Кластер 7 является аномальным и не подчиняется этой тенденции. Получены следующие расширенные характеристики уточненных кластеров с привлечением ранее извлеченных закономерностей.

Кластер 1 включает заявки с простыми задачами, которые потребовали небольшого времени для решения основной задачи; дополнительная задача, как правило, отсутствует; уровень избыточной квалификации специалиста низкий, уровень недостаточной квалификации – низкий; задачи решаются специалистами с невысокой квалификацией. Кратко – это простые задачи. Назначение специалиста можно считать корректным.

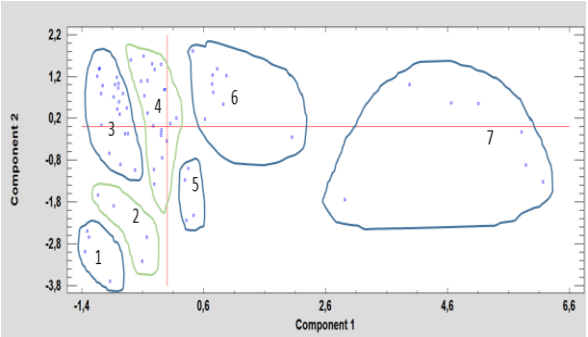


Рис. 2. Диаграмма рассеяния в пространстве ГК1 и ГК2

Кластер 2 включает заявки, которые также потребовали небольшое время для решения основной задачи; дополнительная задача обычно отсутствует; уровни избыточной и недостаточной квалификации специалиста низкие; однако задачи решаются специалистами со средней квалификацией. Кратко – простые задачи. Наблюдается потенциальная избыточность квалификации, требующая незначительной корректировки в назначениях.

Кластер 3 включает заявки с типовыми задачами, которые потребовали среднее время для решения основной задачи; дополнительная задача, как правило, отсутствует; уровни избыточной и недостаточной квалификации специалиста низкие; задачи решаются специалистами со средней квалификацией. Кратко – типовые задачи. Назначения в целом соответствуют сложности.

Кластер 4 включает заявки с типовыми задачами, потребовавшими среднее время для решения; дополнительная задача обычно отсутствует; уровень избыточной квалификации специалиста средний, а недостаточной – низкий; при этом задачи решаются специалистами с высокой квалификацией. Кратко – тоже типовые задачи, но имеет место ошибочное назначение в виде более высокой квалификации исполнителя.

Кластер 5 включает заявки, которые потребовали среднее время для решения основной задачи; дополнительная задача может присутствовать; уровни избыточной и недостаточной квалификации специалиста средние; задачи решаются специалистами со средней квалификацией. Кратко – довольно сложные задачи, которые характеризуются несоответствием уровня квалификации специалиста степени сложности задач, требуется уточнение правил назначения.

Кластер 6 включает заявки со сложными задачами, которые потребовали большое время для решения основной задачи; дополнительные задачи могут встречаться; уровень избыточной квалификации специалиста высокий, а недостаточной – низкий; задачи решаются высококвалифицированными специалистами. Кратко – сложные задачи. Назначение является недостаточно корректным.

Кластер 7 включает заявки с проблемными задачами, которые потребовали большое время решения основной задачи; дополнительные задачи присутствуют; уровень избыточной квалификации специалиста высокий, а недостаточной – средний; при этом назначенные специалисты имеют среднюю квалификацию. Кратко – проблемные задачи. Назначения часто некорректны, наблюдается существенное несоответствие квалификации исполнителя высокой сложности задач, требующее кардинальной корректировки управленческих решений.

Построена карта зон, которая показывает уровень соответствия квалификации специалиста степени сложности задачи, на которой выделены четыре зоны, характеризующие степень соответствия между квалификацией специалиста и сложностью заявки (рис. 4).

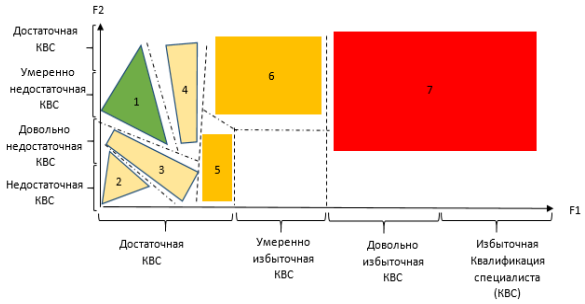


Рис. 3. Карта зон

Сформированы следующие характеристики зон.

Зона 1 (оптимальное соответствие) объединяет заявки, где назначение специалиста является корректным (кластер 1).

Зона 2 (частичное соответствие) включает заявки с незначительными отклонениями в назначениях специалистов как в сторону снижения квалификации, так и в стороны ее превышения по сравнению с требуемой (кластеры 2, 3, 4).

Зона 3 (выраженное несоответствие) содержит заявки, где наблюдается существенное расхождение между сложностью задачи и уровнем квалификации специалиста (кластеры 5, 6).

Зона 4 (критическое несоответствие) охватывает проблемные заявки с значительными ошибками в распределении заявок и назначении специалистов (кластер 7).

Построено дерево решений по результатам этого варианта кластеризации. Выполнен анализ таблицы сопряженности, который подтвердил наличие ошибок классификации в кластерах 3, 4, 5, 6 и 7. Проведён анализ диапазона изменений значений признаков в выделенных кластерах, что используется далее при построении нечетких правил принятия решений при назначении специалиста.

5. Разработка подсистемы НЛВ для поддержки принятия решений при назначении специалиста

Определена структура нечетких правил принятия решений, согласно которой в части ЕСЛИ правил используются четыре лингвистические переменные: уровень знаний клиента, тип основной задачи, время, затраченное на основную задачу и мощность компьютера. Выбор входных лингвистических переменных основан на значимости признаков, входящих в состав главных компонент ГК1 и ГК2.

Выходной лингвистической переменной является квалификация специалиста КВС. Основным источником информации для составления термов являются характеристики кластеров, объединенных в зоны.

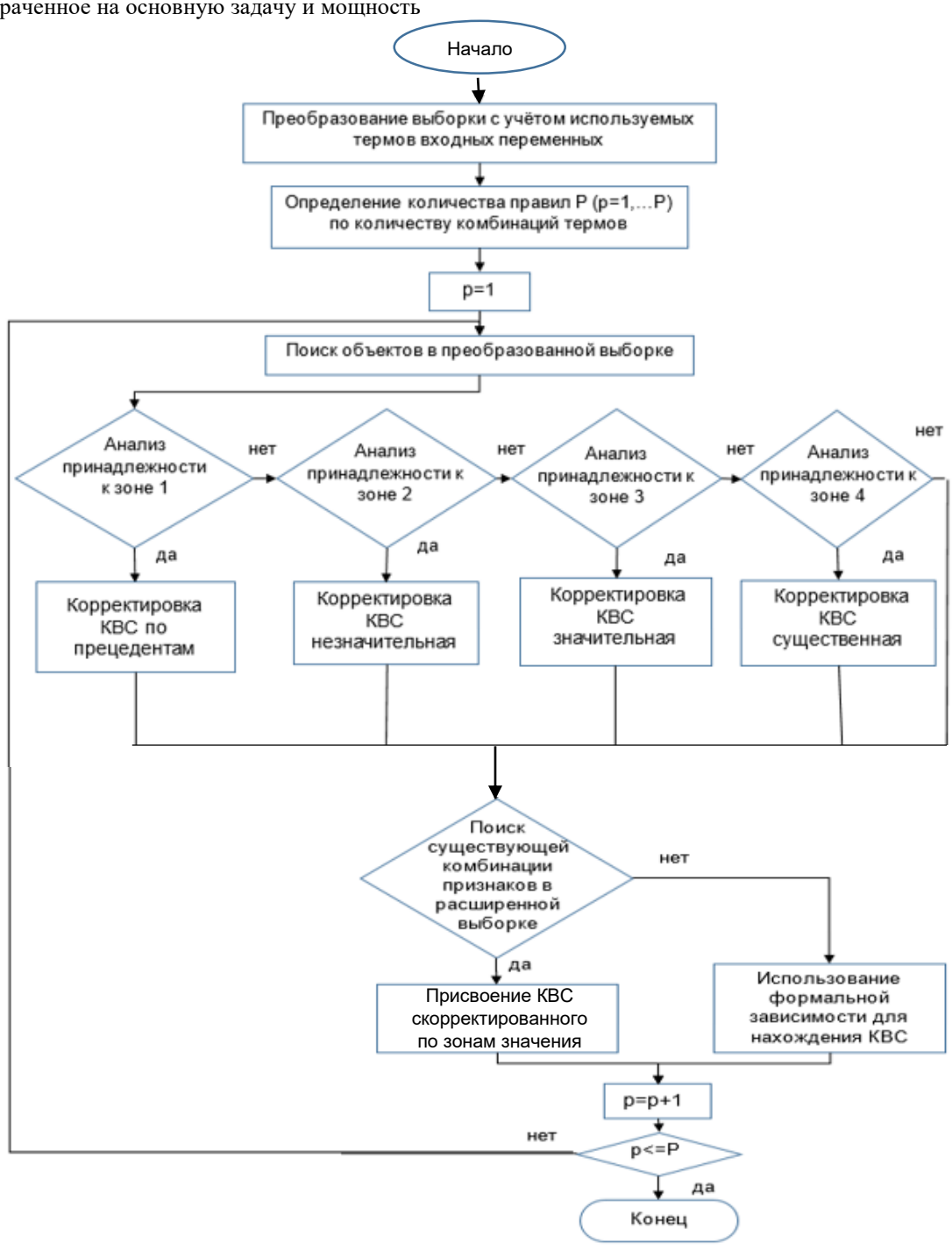


Рис. 4. Схема эвристического алгоритма формирования правой части нечетких правил

Деревья решений, построенные на основе расширенной выборки D2 с помощью выделенных зон, являются дополнительным источником информации для формирования термов в нечётких правилах принятия решений.

Предложен эвристический алгоритм для составления правой части нечетких правил, которые основан на анализе принадлежности к одной из выделенных зон (рис. 4). Основные его особенности состоят в следующем.

Количество термов для каждой входной лингвистической переменной равно трем. Количество термов для выходной переменной КВС равно 5. Всего в базе правил 81 нечеткое правило. Нечеткое правило имеет вид:

ЕСЛИ Уровень знаний клиента=низкий и тип основной задачи = средний и затраченное время=низкое и мощность ПК = среднее,

ТО КВС (квалификация специалиста) = довольно высокая.

Во-первых, расширенная выборка D2 преобразуется путем сокращения признаков, остаются только четыре признака, соответствующих входным лингвистическим переменным.

Во-вторых, в цикле по количеству правил, для каждого из которых заранее сформирована левая часть правил, выполняется поиск всех объектов (заявок) в преобразованной выборке, которые могут соответствовать обозначенной в текущем правиле комбинации значений лингвистических переменных левой частим правила.

В-третьих, проводится анализ принадлежности каждого выбранного объекта к одной из четырех зон. В зависимости от принадлежности или близости к той или иной зоне выполняется: либо присвоение переменной КВС значения в соответствии с прецедентом из выборки (заявкой с корректным решением) для зоны 1; либо значительная, незначительная или существенная корректировки переменной КВС в случаях принадлежности к зонам 2, 3 или 4 соответственно.

Выполнена реализация нечёткого логического вывода выполнена в среде Matlab с использованием компонента Fuzzy Logic Toolbox.

Построены поверхности нечеткого логического вывода и выполнен их анализ (рис.5). Для поверхности в пространстве признаков «уровень знаний клиента» – «тип основной задачи» – «квалификация специалиста» показано, что имеют место возрастающая зависимость и одновременно убывающая зависимость квалификации специалиста от сложности основной задачи и уровня знаний клиента.

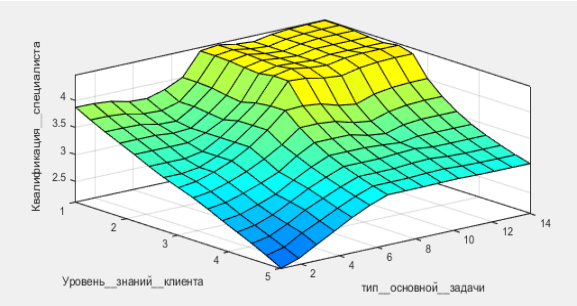


Рис. 5. Поверхность нечеткого логического вывода (уровень знаний клиента – тип основной задачи)

Выявлены следующие закономерности. Переменная КВС «Уровень требуемой квалификации специалиста» останавливает свой рост на средней квалификации (КВС-2.5) в условиях очень высоких знаний клиента. Если же уровень знаний клиента минимальный, то минимальная требуемая квалификация специалиста довольно высокая (КВС=3.8). Плато (максимум, КВС=5) по требуемой квалификации специалиста достигается при уровне сложности основной решаемой задачи выше среднего и и уровне знаний клиента ниже среднего.

6. Заключение

Разработана интеллектуальная система поддержки принятия решений при обработке заявок на решение задач сопровождения компьютерных программ, которая содержит подсистему интеллектуального анализа данных результатах обслуживания заявок на решение задач сопровождения компьютерных программ и подсистему нечеткого логического вывода для формирования рекомендаций при принятии решений по назначению специалиста. С помощью подсистемы интеллектуального анализа данных проведен компонентный анализ ретроспективных данных о результатах обслуживания заявок, построены деревья решений и сформированы кластерные зоны по степени соответствия квалификации исполнителя степени сложности задачи в заявке. Проведенный анализ выявил кластерные зоны с признаками некорректного распределения задач, где наблюдалось значительное несоответствие квалификации специалистов сложности поступающих заявок. Разработана подсистема нечеткого логического вывода и построены нечеткие правила формирования рекомендаций по требуемой квалификации специалиста для решения задачи. Разработка нечетких правил проведена на основе сформированных кластерных зон и выявленных корректировок в назначении специалиста для решения задач. Перспективы развития ИСППР состоят в необходимости расширения спектра признаков, участвующих в формировании правил и решении проблем автоматизации формирования нечетких правил за счет построения обучающихся систем. Применение ИСППР позволит обеспечить повышение

эффективности решения задач на сопровождение компьютерных программ в виде снижения финансовых и временных затрат.

Список используемых источников

1. Саматоев А.Т. Автоматизация управления персоналом: систематический обзор научных публикаций // Вопросы управления. 2025. Т. 19/ № 1. С. 57–79.
2. В.Н. Бурков, Л.В. Россихина, А.П. Вьюнов, and Л.В. Роговая. Задача оптимального распределения команд специалистов // Автоматика и телемеханика. №1, 2019. С. 116–125.
3. И.И. Скрипина, Т.В. Зайцева, and Н.П. Путивцева. Анализ и выбор математической модели с помощью метода анализа иерархий. // Научный результат. Информационные технологии. Tl. 6, № 2, 2021. С. 41–46.
4. Подоплелова Е.С. Анализ методов многокритериального принятия решений на примере задачи ранжирования // Известия Южного федерального университета. Технические науки, №3 (233), 2023. С. 118–125.
5. Мамедова М.Г. и Джабраилова З.Г. Многокритериальная оптимизация задач управления человеческими ресурсами на базе модифицированного метода TOPSIS // Восточно-Европейский журнал передовых технологий. Т. 2, № 4 (74). 2015. С. 48–62.
6. C.H. Kandemir and A.H. Handley, “Work process improvement through simulation optimization of task assignment and mental workload”, in Computational and mathematical organization theory, vol. 25, No. 4, 2019, pp. 389–427. https://doi.org/10.1007/s10588-018-9275-7.
7. Ni Li, Yuhong Li, Mengyuan Sun, Haipeng Kong, and Guanghong Gong, “An optimization method for task assignment for industrial manufacturing organizations”, in Applied Intelligence, vol. 47, No. 4, 2017, pp. 1144–1156. https://doi.org/10.1007/s10489-017-0940-1.
8. Hong Liu, Peng Zhang, Bin Hu, and Philip Moore, “A novel approach to task assignment in a cooperative multi-agent design system”, in Applied Intelligence, vol. 43, No. 1, 2015, pp. 162–175. https://doi.org/10.1007/s10489-014-0640-z.
9. Васильев Е.П., Орешков В.И. Интеллектуальный анализ данных в принятии решений: учебное пособие. Рязань, 2023. –180 с.
10. Кохендерфер М., Уинлер Т., Рэй К. Алгоритмы принятия решений. Москва, ДМК-Пресс. 2023. – 684 с.
11. Лисьева Г. А., Гаврилова И. В. Технологии поддержки принятия решений. Москва, Флинта. 2022. – 133 с.
12. A. Popescu, S. Polat-Erdeniz, A. Felfernig, M. Uta, M. Atas, Viet-Man Le, K. Pilsl, M. Enzelsberger, and Thi Ngoc Trang Tran, “An overview of machine learning techniques in constraint solving”, in Journal of Intelligent Information Systems, vol. 58, 2022, pp. 91–118. https://doi.org/10.1007/s10844-021-00666-5.
13. Макарова Е.А., Габдуллина Э. Р., Юсупов М.М., Камаева Р.Р. Структура и динамические модели управляемого взаимодействия секторов домохозяйств и государственных учреждений. Системная инженерия и информационные технологии. – 2024.- Т 6, № 1(16). С.67-76.
14. Юсупов М. М., Макарова Е.А, Камаева Р.Р. Анализ дифференциации потребительских расходов домохозяйств на основе агент-ориентированного моделирования. // Системная инженерия и информационные технологии. 2023.- Т 5, № 6(15). С.57-66.

Технологии ИИ в системах поддержки принятия решений

Дубкова В. А
Уфимский университет науки
и технологий
Уфа, Россия
e-mail: leradubkovaaa@gmail.com

Родионова Л. Е.
Уфимский университет науки
и технологий
ИИМРТ
Уфа, Россия
e-mail:rodionovakf@yandex.ru

Аннотация[1]

В данной статье проводится глубокий анализ технологий искусственного интеллекта (ИИ), составляющих основу современных систем поддержки принятия решений (СППР). Цель исследования — систематизировать и охарактеризовать ключевые ИИ-методы, применяемые в СППР, выявить их преимущества, ограничения и тенденции развития. В работе подробно рассматриваются такие технологии, как машинное обучение (включая глубокое обучение), объяснимый ИИ (XAI), автоматизированное машинное обучение (AutoML) и гибридные интеллектуальные системы. На основе обзора научной литературы и практических кейсов демонстрируется воздействие этих технологий на эффективность управленческих процессов в различных отраслях. Особое внимание уделяется проблемам интерпретируемости решений, адаптивности алгоритмов и этическим аспектам их применения. В заключении формулируются выводы о трансформирующей роли ИИ в принятии решений и определяются перспективные направления для дальнейших исследований и разработок. Научная новизна работы заключается в разработке комплексной классификации ИИ-технологий для СППР, выявлении психологических аспектов антропоморфизации ИИ и формулировке интегрированных рекомендаций по внедрению, учитывающих принципы доверенного ИИ. В отличие от существующих обзоров, в данной работе представлена авторская оценка перспективности гибридных интеллектуальных систем как архитектурной основы для СППР следующего поколения.

Ключевые слова: искусственный интеллект, системы поддержки принятия решений, машинное обучение, объяснимый ИИ (XAI), глубокое обучение, автоматизированное машинное обучение (AutoML), управленческие решения, анализ данных.

1. Введение

Современный бизнес функционирует в условиях беспрецедентной сложности, неопределенности и информационной перегруженности, как наглядно демонстрирует Рисунок 1, содержащий исторические прогнозы о развитии вычислительной техники, сфера искусственного интеллекта характеризуется высокой онтологической неопределенностью (под которой в контексте данного исследования понимается фундаментальная трудность в предсказании траектории развития и конечных возможностей ИИ-систем ввиду их сложности и непрозрачности, что согласуется с концепциями, обсуждаемыми в работах по философии искусственного интеллекта [27]). Эта иллюстрация показывает, насколько сложно точно предсказать траекторию развития технологий — от ошибочного прогноза основателя IBM Томаса Уотсона о мировом рынке компьютеров ("пять штук") до современных достижений в области нейронных сетей. В этой среде способность к быстрому и обоснованному принятию решений превращается в ключевой фактор конкурентного преимущества. Ответом на этот вызов стало активное внедрение СППР, основанных на передовых технологиях ИИ. Традиционные методы анализа, опирающиеся на человеческий опыт и интуицию, зачастую не справляются с объемами и скоростью обработки необходимых данных. Ответом на этот вызов стало активное внедрение СППР, основанных на передовых технологиях ИИ.



Рис. 1. История представлений о возможностях ИИ (Адаптировано из источника: [28])

ИИ-технологии кардинально трансформируют СППР, превращая их из систем, предоставляющих

[1] Труды XI Международной научной конференции "Информационные технологии интеллектуальной поддержки принятия решений", 11-13 ноября, Уфа-Чандигарх-Фандербейлпарк, 2025

историческую справку, в активные инструменты прогнозирования, сценарного анализа и генерации рекомендаций. Они позволяют автоматизировать обработку больших данных, выявлять скрытые закономерности и моделировать последствия принимаемых решений. Согласно исследованию Gartner, более 37% организаций уже интегрировали ИИ в свои процессы принятия решений, что позволило им повысить производительность на 20% и сократить соответствующие затраты на 25% [1]. Актуальность настоящего исследования обусловлена стремительным развитием самих ИИ-технологий и необходимостью системного анализа их применимости, потенциала и рисков в контексте СППР.

Целью статьи является комплексный анализ современных технологий искусственного интеллекта, используемых в СППР, и оценка их влияния на эффективность управленческой деятельности. Для достижения этой цели решаются следующие задачи:

1. Классифицировать и охарактеризовать ключевые ИИ-технологии, применяемые в СППР.

2. Выявить преимущества и недостатки каждой из технологий на основе анализа научной литературы и практических примеров.

3. Проанализировать актуальные вызовы, такие как проблема интерпретируемости, адаптивности и этики.

4. Определить перспективные направления развития ИИ для СППР.

Объектом исследования выступают технологии искусственного интеллекта. Предметом — их функциональные возможности и эффективность в составе систем поддержки принятия решений.

2. Методология исследования

В качестве методологической основы исследования применяется системный подход, включающий:

1. Сравнительный анализ для выявления сильных и слабых сторон различных ИИ-технологий и платформ.

2. Классификационный метод для структурирования технологий по ключевым функциональным признакам и областям применения.

3. Анализ научной литературы и практических кейсов для верификации преимуществ и ограничений технологий.

4. Структурно-функциональный анализ для оценки влияния технологий на эффективность СППР.

3. Ключевые технологии ИИ в системах поддержки принятия решений

Современные СППР базируются на широком спектре ИИ-технологий, каждая из которых вносит свой вклад в повышение интеллектуальности системы. В данном разделе рассматриваются наиболее значимые из них. Авторский подход к их классификации основан на разделении по доминирующей функции: прогнозирование и анализ данных (МО, Глубокое обучение), обеспечение доверия и прозрачности (XAI), автоматизация процессов (AutoML) и создание комплексных решений (Гибридные системы). Данная классификация позволяет системно оценить вклад каждой технологии в архитектуру СППР.

Машинное обучение (МО) является ядром большинства современных ИИ-решений для СППР. В отличие от традиционного программирования, алгоритмы МО обучаются на исторических данных, выявляя сложные зависимости и строя прогнозные модели. Обобщенную схему этого процесса иллюстрирует Рисунок 2, который визуализирует ключевые этапы: от сбора и подготовки данных через процесс обучения (поиск закономерностей) к практическому применению полученной модели. Эта схема помогает понять фундаментальное отличие МО от классического программирования.



Рис. 2. Упрощенная схема процесса машинного обучения (Адаптировано из источника: [29])

Классификация и регрессия: Алгоритмы (например, решающие деревья, метод опорных векторов, логистическая регрессия) широко используются для задач бинарной и многоклассовой классификации (например, оценка кредитоспособности клиента, классификация обращений в службу поддержки) и прогнозирования непрерывных величин (например, спрос на продукцию, стоимость актива).

Глубокое обучение (Deep Learning): Подмножество МО, основанное на использовании искусственных нейронных сетей (ИНС) со множеством слоев. Глубокие нейронные сети демонстрируют выдающиеся результаты в обработке неструктурированных данных: изображений (анализ медицинских снимков), текста (сентимент-анализ отзывов, автоматизация анализа юридических документов) и речи. Например, JPMorgan Chase использует систему COiN на основе глубокого

обучения для анализа кредитных договоров, сокращая трудозатраты с тысяч часов до секунд [2,3]. Однако ключевой проблемой ИНС является их «черный ящик» — сложность интерпретации причин, по которым сеть приняла то или иное решение. По мнению авторов, несмотря на высокую производительность, широкое применение глубоких нейронных сетей в ответственных областях управления без сопутствующего внедрения методов XAI является неоправданным риском.

Проблема «черного ящика» машинного обучения, особенно актуальная в таких сферах, как финансы и медицина, стимулировала развитие направления Explainable AI (XAI). XAI объединяет методы и техники, позволяющие сделать выводы ИИ-моделей понятными и интерпретируемыми для человека-пользователя.

Методы: к методам XAI относятся LIME (Local Interpretable Model-agnostic Explanations), SHAP (SHapley Additive exPlanations), которые помогают определить, какие именно входные данные оказали наибольшее влияние на итоговое решение модели.

Практическая значимость: Внедрение XAI повышает доверие к системе со стороны менеджеров и позволяет соблюдать регуляторные требования. Например, в Европейском союзе в рамках Artificial Intelligence Act для систем высокого риска объяснимость является обязательным требованием. Российские банки, такие как Т-Банк, также сталкиваются с необходимостью предоставления прозрачных объяснений при автоматическом кредитном скоринге [4,5]. Авторы полагают, что XAI не следует рассматривать как опциональную "надстройку"; это неотъемлемый компонент любой промышленной СППР, претендующей на надежность и соответствие этическим нормам.

Технологии обработки естественного языка (Natural Language Processing, NLP) позволяют СППР понимать, интерпретировать и генерировать человеческую речь. Это открывает возможности для работы с текстовыми данными: отчетами, новостными лентами, внутренней перепиской, отзывами клиентов.

Применение: NLP используется для автоматического суммирования документов, извлечения ключевых сущностей (имен, компаний, сумм), анализа тональности в социальных сетях для оценки репутации бренда. Это позволяет менеджерам оперативно получать смысловую сводку из больших объемов текстовой информации.

AutoML — это технология, направленная на автоматизацию процессов построения и развертывания моделей машинного обучения. Она призвана снизить порог входа для компаний, не обладающих большими командами data-саентистов.

Преимущества: AutoML автоматизирует такие трудоемкие этапы, как feature engineering, подбор гиперпараметров и выбор алгоритма. Это ускоряет разработку СППР и делает технологии ИИ доступными для малого и среднего бизнеса. Компания «Ростелеком», например, применяет AutoML для автоматизации управления ИТ-инфраструктурой [4,5].

Ограничения: Недостатком AutoML может быть меньшая гибкость и возможность создания нестандартных решений по сравнению с ручной настройкой экспертов.

Перспективным направлением является создание гибридных систем, которые комбинируют различные ИИ-технологии, а также интегрируют формализованные знания экспертов (экспертные системы) с возможностями машинного обучения.

Суть подхода: например, гибридная система может использовать нейронную сеть для прогнозирования спроса, а правила, заданные опытными логистами, — для корректировки этого прогноза с учетом специфических факторов (погодных условий, локальных мероприятий). Такие системы обладают большей гибкостью и устойчивостью. Исследование Афзал и др. (2019) показало, что гибридные модели на основе нечеткой логики и байесовских сетей способны снизить ошибки прогнозирования на финансовых рынках на 15-20% [6]. С точки зрения авторов, именно гибридные системы, сочетающие мощь машинного обучения с прозрачностью экспертных правил и методами XAI, представляются наиболее перспективной архитектурой для СППР следующего поколения, позволяющей преодолеть разрыв между чистой аналитикой и предметной логикой. Сравнительный анализ приведенных технологий ИИ приведен в таблице 1.Сравнительный анализ приведенных технологий ИИ приведен в таблице 1.

Таблица 1. – Сравнительный анализ ключевых технологий ИИ для СППР

Технология	Основные возможности	Преимущества	Недостатки/ вызовы
Машинное обучение	Прогнозирование, классификация, кластеризация	Высокая точность на больших данных, автоматизация анализа	Зависимость от качества данных, риск переобучения
Глубокое обучение	Анализ изображений, текста, речи	Высокая точность с неструктурированными данными	Сложность интерпретации («черный ящик»), большие вычислительные затраты
Объяснимый ИИ (XAI)	Объяснение решений моделей ИИ	Повышение доверия, соответствие регуляторным нормам	Может упрощать сложную логику модели, дополнительные вычисления
AutoML	Автоматизация построения моделей МО	Снижение порога входа, ускорение разработки	Меньшая гибкость для нестандартных задач
Гибридные системы	Комбинация возможностей разных технологий	Повышенная адаптивность, устойчивость	Сложность проектирования и интеграции

4. Влияние технологий ИИ на эффективность СППР: анализ практических кейсов

Внедрение перечисленных технологий оказывает значимое влияние на различные аспекты эффективности управленческой деятельности.

ИИ-алгоритмы способны анализировать данные на порядки быстрее человека и выявлять сложные, неочевидные зависимости. Это приводит к существенному повышению как скорости, так и точности решений. Исследование McKinsey показывает, что внедрение ИИ ускоряет процессы принятия решений до 30% [2,6]. В логистике Amazon алгоритмы машинного обучения в реальном времени пересчитывают оптимальные маршруты доставки, учитывая пробки, погоду и другие факторы, что сокращает время доставки на 25% [8].

Технологии прогнозной аналитики на основе МО позволяют компаниям не просто реагировать на события, а предвосхищать их. Классический пример — предиктивное техническое обслуживание (Predictive Maintenance) на производстве. Платформа Predix от General Electric анализирует данные с датчиков оборудования и прогнозирует вероятность поломки, позволяя проводить ремонт до возникновения критического простоя, что снижает затраты на 20-30% [9,10,11].

СППР на базе ИИ эффективно решают задачи оптимизации: управление запасами (как в Procter & Gamble [12]), цепочками поставок, персоналом. Алгоритмы находят оптимальное распределение ресурсов при заданных ограничениях, что напрямую ведет к снижению операционных расходов.

Глубокий анализ данных помогает обнаруживать новые рыночные тренды, сегменты клиентов или потенциальные риски, которые могли остаться незамеченными при традиционном анализе. Например, алгоритмы анализа тональности в социальных сетях могут выявить нарастающую неудовлетворенность продуктом на ранней стадии.

5. Актуальные вызовы и ограничения технологий ИИ в СППР

Несмотря на потенциал, внедрение ИИ-технологий сопряжено с рядом серьезных вызовов.

Эффективность любой модели ИИ напрямую зависит от качества и репрезентативности данных, на которых она обучена. Неполные, неточные или смещенные данные приводят к некорректным и несправедливым прогнозам. Известны случаи, когда алгоритмы кредитного скоринга демонстрировали предвзятость по расовому или гендерному признаку из-за исторических данных, содержащих такие смещения [13].

Даже с развитием XAI сложные модели, особенно глубокие нейронные сети, остаются трудными для полного понимания. Это создает барьер доверия для менеджеров, которые несут ответственность за принимаемые решения. В ситуациях, требующих юридической или этической ответственности, слепое следование рекомендации «черного ящика» неприемлемо.

Широкое использование ИИ поднимает вопросы ответственности за автономные решения, конфиденциальности данных и потенциального влияния на рынок труда. Разработка нормативных документов, подобных European AI Act, является важным шагом для минимизации рисков и

обеспечения ответственного использования технологий.

Многие модели ИИ статичны и требуют периодического переобучения для сохранения актуальности в условиях быстро меняющейся внешней среды. Развитие адаптивных и непрерывно обучающихся систем (как в гибридных моделях) является ключевым направлением для преодоления этого ограничения.

6. Сравнительный анализ платформ для построения СППР на базе ИИ

Рост спроса на ИИ-решения для поддержки принятия решений стимулировал появление множества программных платформ, предлагающих инструменты для разработки и внедрения таких систем. Проведем сравнительный анализ характеристик ведущих платформ, представленных на рынке в 2023-2024 годах – таблица 2. Это позволит практикам сориентироваться в выборе технологического стека для реализации проектов.

Таблица 2. Сравнительный анализ платформ для построения СППР на базе ИИ

Платформа/поставщик	**Ключевые технологии и особенности**	**Целевая аудитория**	**Преимущества**	**Недостатки**
Microsoft Azure AI [14]	Azure Machine Learning, Cognitive Services (NLP, компьютерное зрение), Bot Framework. Глубокая интеграция с облачной экосистемой Azure.	Крупные предприятия, разработчики	Широкая палитра сервисов, высокая масштабируемость, надежная безопасность и соответствие стандартам, поддержка open-source фреймворков (PyTorch, TensorFlow).	Сложность освоения для новичков, стоимость может быть высокой для проектов с интенсивным использованием ресурсов.
Google Cloud AI Platform [15]	Мощные инструменты для ML (Vertex AI), передовые NLP-модели (BERT, PaLM), BigQuery для анализа данных	Data scientists, крупный бизнес	Лидерство в области исследований ИИ, мощные предобученные модели, эффективная работа с большими данными.	Привязка к экосистеме Google Cloud, менее гибкие варианты гибридного развертывания по сравнению с конкурентами.
IBM Watson [16]	Акцент на объяснимом ИИ (AI Explainability 360) и автоматизации AI (Watson Studio, AutoAI). Сильные стороны в обработке NLP для корпоративных задач.	Финансовый сектор, здравоохранение, госсектор	Упор на доверие и прозрачность, готовые решения для отраслей, мощные инструменты для работы с неструктурированными данными.	Воспринимается как менее гибкая и более дорогая по сравнению с облачными платформами широкого профиля.
Amazon SageMaker [17]	Полностью управляемый	Разработчики, компании,	Упрощение и ускорение процесса	Как и у Google, привя

	сервис для полного цикла ML: от подготовки данных до развертывания и мониторинга. Интеграция с AWS.	уже использующие AWS	ML за счет встроенных инструментов (например, SageMaker Clarify для обнаружения bias), эластичная масштабируемость.	зка к собственной облачной экосистеме (AWS).
Open-Source Решения (например, Scikit-learn, PyTorch, TensorFlow, H2O.ai) [18]	Библиотеки и фреймворки для самостоятельной разработки моделей. Максимальная гибкость и контроль.	Data scientists, исследователи, стартапы	Независимость от вендоров, низкая стоимость (за исключением инфраструктуры), полная кастомизация под любую задачу.	Высокие требования к экспертизе, необходимость самостоятельной сборки инфраструктуры и пайплайнов, ответственность за поддержку и безопасность.

Анализ показывает, что выбор платформы определяется несколькими факторами: уровнем экспертизы команды, масштабом проекта, отраслевыми требованиями (особенно к объяснимости и compliance) и уже существующей ИТ-инфраструктурой. Для крупных корпораций целесообразно использование комплексных решений от Microsoft, Google или IBM, в то время как для стартапов и исследовательских групп могут быть предпочтительнее open-source инструменты. Трендом последних лет является развитие гибридных и мультиоблачных стратегий, когда компоненты СППР могут работать в разных средах, что повышает гибкость и снижает риски привязки к одному вендору [19].

7. Психологические и социальные аспекты

Феномен антропоморфизации ИИ

Особую опасность представляет феномен антропоморфизации ИИ-систем. Рисунок 3 иллюстрирует концепцию "липкости технологий" — ложного чувства адекватной замены традиционным социальным связям. Антропоморфные интерфейсы и имитация эмоций со стороны ИИ-систем могут создавать у пользователей иллюзию эмоциональной связи, что в перспективе способно привести к социальной изоляции и деградации межличностных отношений. Авторы считают, что данный аспект является недооцененным риском при проектировании пользовательских интерфейсов СППР и требует включения в этические руководства по разработке. Стремление сделать интерфейс "дружелюбным" не должно подменяться созданием иллюзии человеческой личности, ответственной за решения. Стремление алгоритмов "понравиться" пользователю создает риски манипуляции поведением через уже установленные эмпатические связи. Взаимодействие с искусственным партнером по сути является взаимодействием с "черным ящиком", что может приводить к социальной дезориентации.



Рис. 3. Схематическое изображение феномена антропоморфизации ИИ-систем и связанных с ним рисков (Разработано авторами)

8. Практические рекомендации по внедрению технологий ИИ в СППР

На основе анализа возможностей и рисков можно сформулировать интегрированные рекомендации по внедрению ИИ в СППР. Фундаментальную основу для построения безопасных и этичных систем предоставляет Рисунок 4, который визуализирует

принципы "доверенного ИИ" (Trustworthy AI). Данная схема включает три ключевых компонента:

1. Прозрачность и объяснимость — обеспечение понимания принципов работы системы
2. Устойчивость, надежность и безопасность — гарантия стабильной работы в различных условиях
3. Человекоцентричный подход — приоритет прав человека и общечеловеческих ценностей

Ключевые рекомендации по внедрению:



Рис. 4. Принципы доверенного ИИ (Trustworthy AI), лежащие в основе авторских рекомендаций (Адаптировано из источника: [30])

Критически важным является правило, подчеркнутое на Рисунке 4: "Только ИИ не может принимать критические решения". Человек должен всегда сохранять контроль над системой, а "выключатель" должен быть под рукой у оператора. С точки зрения авторов, следование этим рекомендациям позволяет системно подойти к внедрению, значительно снижая риски и повышая вероятность реализации заложенного в технологии ИИ потенциала для совершенствования управленческих процессов.

На основе анализа успешных кейсов и выявленных вызовов можно сформулировать ряд практических рекомендаций для организаций, планирующих внедрение ИИ-технологий в свои процессы принятия решений. Эти рекомендации охватывают стратегические, технические и организационные аспекты.

1. Начните с четко определенной бизнес-задачи, а не с технологии. Внедрение ИИ должно быть направлено на решение конкретной проблемы или извлечение измеримой выгоды (например, «сокращение времени обработки заявки на 30%», «снижение уровня брака на 5%»). Избегайте расплывчатых целей, таких как «внедрить ИИ для аналитики» [19]. Это позволяет правильно выбрать технологию, оценить успех проекта и получить поддержку стейкхолдеров.

2. Обеспечьте высокое качество и доступность данных. Поскольку данные — это «топливо» для ИИ, инвестируйте в создание надежных каналов их сбора, хранения и управления (Data Governance). Реализуйте процессы очистки, разметки и мониторинга качества данных на постоянной основе. Помните, что проект может потерпеть неудачу из-за проблем с данными, даже при использовании самых передовых алгоритмов [20].

3. Формируйте междисциплинарные команды. Успешные проекты по внедрению ИИ требуют collaboration между экспертами различных профилей: data scientists, инженеров по данным, бизнес-аналитиков, специалистов в предметной области и конечных пользователей-менеджеров. Это гарантирует, что создаваемая система будет технически грамотной, релевантной бизнес-потребностям и удобной в использовании [21].

4. Придерживайтесь принципа «итеративной разработки» (Agile). Внедряйте решение поэтапно, начиная с пилотного проекта (Proof of Concept), который решает узкую задачу. Это позволяет быстро продемонстрировать ценность, получить обратную связь, оценить риски и скорректировать дальнейшие действия, минимизируя крупные инвестиции на ранних стадиях [22].

5. Уделяйте первостепенное внимание интерпретируемости и доверию. С самого начала проектирования системы интегрируйте практики Explainable AI (XAI). Обеспечьте, чтобы рекомендации СППР сопровождались понятными для пользователя объяснениями (например, «кредит отклонен из-за высокого соотношения долговой нагрузки, которое составляет 65% при допустимом пороге в 50%»). Это повысит принятие системы со стороны менеджеров и облегчит аудит [23].

6. Разработайте стратегию управления изменениями. Внедрение ИИ меняет рабочие процессы и требует новых компетенций. Необходимо активно вовлекать сотрудников, проводить обучение, пересматривать должностные инструкции и формировать data-driven культуру. Важно донести, что ИИ — это инструмент для усиления, а не для замены человеческого экспертизы [24].

7. Оценивайте эффективность комплексно. Помимо традиционных технических метрик (точность, полнота модели) и финансовых показателей (ROI), оценивайте влияние системы на стратегические цели и организационную культуру. Используйте интегральные модели, учитывающие скорость принятия решений, удовлетворенность пользователей и рост инновационного потенциала [25].

8. Учитывайте этические и правовые аспекты с самого начала. Создавайте этические хартии использования ИИ внутри организации. Проводите аудит алгоритмов на предмет наличия смещений (bias). Обеспечьте соответствие разрабатываемой системы действующему и перспективному законодательству (например, GDPR, ФЗ-152, AI Act) [26].

Следование этим рекомендациям не гарантирует абсолютного успеха, но позволяет системно

подойти к внедрению, значительно снижая риски и повышая вероятность реализации заложенного в технологии ИИ потенциала для совершенствования управленческих процессов.

9. Заключение

На основе проведенного исследования сформулированы следующие основные научные положения и выводы, отражающие вклад работы и авторскую позицию:

1. Предложена авторская комплексная классификация ключевых ИИ-технологий для СППР, выделяющая их по функциональному признаку (прогнозирование, обеспечение прозрачности, автоматизация, комплексные решения) и выявляющая синергетический потенциал их гибридизации.

2. Выявлен и проанализирован феномен антропоморфизации ИИ как значимый психологический риск, способный привести к снижению критичности восприятия рекомендаций СППР и социальной дезориентации пользователей, что, по мнению авторов, должно учитываться при проектировании человеко-машинного интерфейса.

3. Обоснована авторская точка зрения о перспективности гибридных интеллектуальных систем как архитектуры, наиболее адекватно

отвечающей вызовам современного управления, так как они позволяют сочетать вычислительную мощь машинного обучения с прозрачностью, контролем и предметными знаниями.

4. Разработаны интегрированные практические рекомендации по внедрению ИИ в СППР, основанные на принципах доверенного ИИ и предусматривающие итеративный подход, междисциплинарное взаимодействие и приоритет интерпретируемости и этических норм на всех этапах жизненного цикла системы.

5. Трансформация роли СППР. Технологии искусственного интеллекта (машинное обучение, глубокое обучение, XAI, NLP, AutoML) кардинально меняют функционал систем поддержки принятия решений, превращая их из инструментов пассивной визуализации данных в активных интеллектуальных помощников, способных к прогнозированию, оптимизации и генерации обоснованных рекомендаций.

6. Психологические аспекты, такие как антропоморфизация ИИ (Рисунок 3), требуют особого внимания при проектировании интерфейсов и взаимодействия человек-машина.

7. Необходим сбалансированный подход, сочетающий прагматичное внедрение технологий для повышения эффективности с безусловным соблюдением принципов доверенного ИИ (Рисунок 4).

8. Доказанное положительное влияние. Внедрение ИИ-технологий в СППР приводит к значительному повышению скорости, точности и обоснованности управленческих решений, осуществляет переход от реактивной к проактивной модели управления и способствует оптимизации использования ресурсов, что подтверждается практическими кейсами ведущих компаний.

9. Выявление ключевых вызовов. Широкомасштабное применение ИИ в СППР сдерживается рядом серьезных проблем, наиболее значимыми из которых являются:

Зависимость от качества данных и наличие в них скрытых смещений (bias).

Недостаточная интерпретируемость решений сложных моделей («черный ящик»).

Нерешенные этические и регуляторные вопросы, связанные с ответственностью за автономные решения.

10. Определение перспективных направлений развития. Дальнейшая эволюция технологий ИИ для СППР будет связана со следующими направлениями:

Развитие объяснимого ИИ (XAI) для повышения прозрачности и доверия.

Создание адаптивных и гибридных систем, способных к непрерывному обучению.

Повышение надежности и безопасности моделей (борьба со смещениями, устойчивость к атакам).

Демократизация доступа к технологиям через развитие AutoML и облачных платформ.

11. Необходимость междисциплинарного подхода. Успешная интеграция и дальнейшее развитие ИИ в управленческих процессах требуют тесного сотрудничества специалистов по data science, экспертов в предметных областях, менеджеров и регуляторов для максимизации преимуществ и минимизации сопутствующих рисков.

Благодарности

Исследование выполнено при финансовой поддержке Министерства науки и высшего образования Российской Федерации в рамках основной части государственного задания высшим учебным заведениям № FEUE-2023-0007

Список используемых источников

1. Gartner top 10 strategic technology trends for 2024 // Gartner. 2024. URL: https://www.gartner.com/en/articles/gartner-top-10-strategic-technology-trends-for-2024 (дата обращения: 18.09.2025). (электронный ресурс).

2. Brynjolfsson E., Rock D., Syverson C. The productivity J-curve: How intangibles complement

general purpose technologies // American Economic Journal: Macroeconomics. 2021. Vol. 13. No. 1. P. 333–372. DOI: 10.1257/mac.20180386. (статья из журнала).

3. JPMorgan reduced lawyers' hours by 360,000 annually by automating loan agreement analysis with machine learning software COIN // Best Practice AI. URL: https://www.bestpractice.ai/ai-case-study-best-practice/jpmorgan_reduced_lawyers'_hours_by_360%2C000_annually_by_automating_loan_agreement_analysis_with_machine_learning_software_coin (дата обращения: 18.09.2025). (электронный ресурс).

4. The state of AI in early 2024: Gen AI adoption spikes and starts to generate value // McKinsey & Company. May 30, 2023. URL: https://www.mckinsey.com/capabilities/quantumblack/our-insights/the-state-of-ai (дата обращения: 19.09.2025). (электронный ресурс).

5. Искусственный интеллект в России в 2023 году: тренды и перспективы // ICT.Moscow. 2023. URL: https://ict.moscow/projects/ai/research/iskusstvennyi-intellekt-v-rossii-v-2023-godu-trendy-i-perspektivy/ (дата обращения: 19.09.2025). (электронный ресурс).

6. Afzal M., Chen L., Khan A. A hybrid fuzzy-bayesian approach for market trend forecasting under volatility // Expert Systems with Applications. 2019. Vol. 115. P. 34–45. DOI: 10.1016/j.eswa.2018.07.062. (статья из журнала).

7. Bromels J. 3 ways GE's Predix is revolutionizingustomers' operations// Fox Business. URL: https://www.foxbusiness.com/markets/3-ways-ges-predix-is-revolutionizing-customers-operations (дата обращения: 19.09.2025). (электронный ресурс).

8. Смагин В.С. Промышленный Интернет — мечты сбываются // ИСУП. 2016. № 5. С. 61–63. URL: https://isup.ru/upload/pdf-zhurnala/2018%20i%20dalee/2016/5/060_063_Advantek%20inzheniring.pdf (дата обращения: 20.09.2025). (статья из журнала).

9. Применение искусственного интеллекта в логистике // Itob.ru. 2024. 18 июня. URL: https://itob.ru/blog/primenenie-iskusstvennogo-intellekta-v-logistike/ (дата обращения: 20.09.2025). (электронный ресурс).

10. Wamba S. F., Bawack R. E., Guthrie C., Queiroz M. M., Carillo K. D. A. Are we preparing for a good AI society? A bibliometric review and research agenda // Technological Forecasting and Social Change. 2021. Vol. 164. Article 120482. DOI: 10.1016/j.techfore.2020.120482. (статья из журнала).

11. Barredo Arrieta A., Díaz-Rodríguez N., Del Ser J., et al. Explainable artificial intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI // Information Fusion. 2020. Vol. 58. P. 82–115. DOI: 10.1016/j.inffus.2019.12.012. (статья из журнала).

12. Microsoft Azure AI. URL: https://azure.microsoft.com/ru-ru/products/machine-learning/ (дата обращения: 20.09.2025). (электронный ресурс).

13. Google Cloud AI Platform. URL: https://cloud.google.com/vertex-ai (дата обращения: 20.09.2025). (электронный ресурс).

14. IBM Watson. URL: https://www.ibm.com/watson (дата обращения: 20.09.2025). (электронный ресурс).

15. Amazon SageMaker. URL: https://aws.amazon.com/ru/sagemaker/ (дата обращения: 20.09.2025). (электронный ресурс).

16. Scikit-learn: Machine Learning in Python. URL: https://scikit-learn.org/ (дата обращения: 20.09.2025). (электронный ресурс).

17. PyTorch. URL: https://pytorch.org/ (дата обращения: 20.09.2025). (электронный ресурс).

18. Hagiu A., Wright J. When Data Creates Competitive Advantage // Harvard Business Review. 2020. Vol. 98. No. 1. P. 94–101. (статья из журнала).

19. Davenport T. H., Ronanki R. Artificial Intelligence for the Real World // Harvard Business Review. 2018. Vol. 96. No. 1. P. 108–116. (статья из журнала).

20. Redman T. C. If Your Data Is Bad, Your Machine Learning Tools Are Useless // Harvard Business Review. 2018. URL: https://hbr.org/2018/02/if-your-data-is-bad-your-machine-learning-tools-are-useless (дата обращения: 20.09.2025). (электронный ресурс).

21. Ries E. The Lean Startup: How Today's Entrepreneurs Use Continuous Innovation to Create Radically Successful Businesses. New York: Crown Business, 2011. 336 p. (книга).

22. Adadi A., Berrada M. Peeking Inside the Black-Box: A Survey on Explainable Artificial Intelligence (XAI) // IEEE Access. 2018. Vol. 6. P. 52138–52160. DOI: 10.1109/ACCESS.2018.2870052. (статья из журнала).

23. Hiatt J. M. ADKAR: A Model for Change in Business, Government and our Community. Loveland: Prosci Research, 2006. 150 p. (книга).

24. Kaplan R. S., Norton D. P. The Balanced Scorecard: Measures That Drive Performance // Harvard Business Review. 1992. Vol. 70. No. 1. P. 71–79. (статья из журнала).

25. European Commission. Proposal for a Regulation laying down harmonised rules on artificial intelligence (Artificial Intelligence Act). Brussels, 2021. 108 p. URL: https://eur-lex.europa.eu/legal-

content/EN/TXT/?uri=CELEX:52021PC0206 (дата обращения: 20.09.2025). (электронный ресурс).

26. Iansiti M., Lakhani K. R. The Truth About Blockchain // Harvard Business Review. 2017. Vol. 95. No. 1. P. 118–127. (статья из журнала).

27. Bostrom N., Yudkowsky E. The Ethics of Artificial Intelligence // The Cambridge Handbook of Artificial Intelligence. Cambridge University Press, 2014. P. 316–334. (Книга)

28. A Brief History of AI // Our World in Data. 2023. URL: https://ourworldindata.org/brief-history-of-ai (дата обращения: 21.09.2025). (Электронный ресурс)

29. Understanding Machine Learning // Google 3 Ways GE's Predix Is Revolutionizing Customers' Operationswww.foxbusiness.com Developers. 2024. URL: https://developers.google.com/machine-learning/intro-to-ml (дата обращения: 21.09.2025). (Электронный ресурс)

30. Trustworthy AI // Microsoft AI Principles. 2023. URL: https://www.microsoft.com/en-us/ai/responsible-ai (дата обращения: 21.09.2025). (Электронный ресурс)

Метод предсказания давления в газопроводах, сочетающий законы физики и устойчивое запоминание последовательностей

Пальчевский Е. В.
МИРЭА – Российский технологический университет (РТУ МИРЭА)
кафедра индустриального программирования
Москва, Россия
e-mail: teelxp@inbox.ru

Антонов В. В.
Уфимский университет науки и технологий
каф. АСУ
Уфа, Россия
e-mail: antonov.v@bashkortostan.ru

Аннотация[1]

Актуальность краткосрочного прогнозирования давления в магистральных газопроводах обусловлена старением инфраструктуры, усложнением режимов транспортировки газа и ужесточением нормативов промышленной безопасности, требующих мгновенной реакции на любые колебания давления. Неправильный выбор оптимального давления способен вызвать гидроудар, разгерметизацию труб и масштабные экономические потери, поэтому точные прогнозы востребованы операторскими центрами по всему миру.

В ответ на эти вызовы рассмотрена задача краткосрочного прогнозирования давления в магистральных газопроводах, критичная для промышленной безопасности и энергоэффективности. В качестве решения задачи предложена гибридная технология – Adaptive LSTM, объединяющая интерпретируемость физических моделей (например, закон сохранения массы) с высокой аппроксимирующей способностью классической LSTM. Дополнительно реализован механизм внимания, куда, помимо временных рядов, напрямую подаются массовый расход и температура газа. Благодаря составной функции потерь, включающей физическую регуляризацию, сеть выдает устойчивые и физически корректные прогнозные значения даже в зашумленных и переходных режимах работы газопровода.

Разработанная модель интегрирована в веб-приложение на Django, которое выполняет онлайн-обучение Adaptive LSTM, мониторинг и автоматический расчет радиуса потенциального взрыва по предсказанным значениям давления в газопроводе. Тем самым технология превращается в практический инструмент поддержки операторских решений, позволяя заранее обнаруживать аномалии и оперативно снижать риски аварий на газотранспортной инфраструктуре.

Ключевые слова: газопровод; давление; краткосрочный прогноз; гибридная LSTM; машинное обучение; промышленная безопасность.

Keywords: gas pipeline; pressure; short-term forecast; hybrid LSTM; machine learning; industrial safety.

1. Введение

Точное и надежное прогнозирование давления в газопроводах является ключевой задачей для обеспечения промышленной безопасности, энергоэффективности и устойчивой эксплуатации газотранспортных систем [1, 2]. Актуальность усиливается усложнением режимов работы, в том числе переключениями компрессорных агрегатов, нестационарными отборами и пиковыми нагрузками, а также жесткими требованиями к раннему обнаружению аномалий [3-5]. Классические аналитические модели на основе упрощенных уравнений газовой динамики, например, баланса массы и энергии, приближения Бернулли/идеального газа, дают интерпретируемые оценки, однако нередко уступают по точности в реальных, зашумленных и многомодальных режимах [6-8]. Напротив, современные методы машинного и глубокого обучения хорошо аппроксимируют нелинейности и скрытые зависимости, но без физической регуляризации склонны к неустойчивым предсказаниям и погрешности при экстраполяции данных [9-19].

Цель исследования – создать и экспериментально обосновать технологию в виде гибридного метода краткосрочного прогнозирования давления в

[1] Труды XI Международной научной конференции "Информационные технологии интеллектуальной поддержки принятия решений", 11-13 ноября, Уфа-Чандигарх-Фандербейлпарк, 2025

газопроводе, которая, объединяя физические законы (баланс массы и энергии) с возможностями моделей глубокого обучения, обеспечит более точные, устойчивые и физически корректные предсказания давления в газопроводе с учетом потенциальных последствий.

Для реализации поставленной цели разработана оригинальная гибридная технология, основанная на классической сети LSTM (Long Short-Term Memory), но обогащенная важными математическими и технологическими усовершенствованиями. В модифицированную ячейку, получившей название Adaptive LSTM, с механизмом внимания, помимо временных признаков, напрямую подаются технологические параметры процесса – массовый расход газа и температура. Обучение выполняется по составной функции потерь: к среднеквадратичной ошибке (MSE, Mean Squared Error) по давлению добавлена физическая регуляризация, обеспечивающая баланс массы и энергии. Такой подход делает прогнозы более достоверными и устойчивыми как при смене режимов работы газопровода, так и при аномалиях во временных рядах. Дополнительные преимущества включают интерпретируемость через механизм внимания, выделяющий ключевые участки данных, и промышленно ориентированное веб-приложение на Django [11] с фоновым обучением, логированием и онлайн-мониторингом, готовое к интеграции в системы поддержки операторских решений.

Практическая значимость технологии проявляется в:

- снижении риска получения прогнозов, нарушающих законы сохранения массы и энергии, что критично для оперативного управления;
- промышленно ориентированной архитектуре: полнофункциональное веб-приложение на базе Django, реализованное на основе научных выводов, полученных в рамках данного исследования, позволяет управлять процессами обучения модифицированной LSTM, прогнозирования и моделирования потенциальных аварийных сценариев в рамках взрыва на газопроводе.

2. Постановка задачи

Рассматривается участок магистрального газопровода длиной L_p и с постоянной площадью поперечного сечения A. Контрольный объем магистрального газопровода измеряется как $V = A \cdot L_p$. В качестве наблюдаемых параметров берутся дискретные ряды с шагом Δt: давление P_t [Па]; температура газа T_t[K]; массовый расход газа на входе $\dot{m}_{in,t}$ [кг/с]; при наличии – массовый расход газа на выходе $\dot{m}_{out,t}$ [кг/с] либо внешнее давление/сброс $P_{out,t}$ [Па].

Параметры газа: плотность определяется выражением $p_t = \frac{P_t}{Z_t R T_t'}$, где R – газовая постоянная; коэффициент сжимаемости $Z_t \approx 1$ (для умеренных значений давления); удельной теплоемкости при постоянном давлении теплоемкости C_v и удельной теплоемкости газа при постоянном объеме C_p для расчета показателя адиабаты $\gamma = \frac{C_p}{C_v}$.

Задача моделирования аварийной ситуации в газопроводе состоит из двух взаимосвязанных этапов: прогнозирование давления в газопроводе и моделирование радиуса возможного взрыва на основе прогнозного значения давления в газопроводе.

2.1. Этап прогнозирования давления в газопроводе

Берется скользящее окно длиной L по временным рядам и подается в нейросетевую модель f_Θ, которая возвращает прогноз давления для будущего временного шага h (h = 1,…,H). После формирования входной матрицы признаков $X_{t-L+1:t}$, отражающей динамику давления, температуры, потоков и граничных условий за последние L шагов, происходит подача его в параметризованную нейросетевую модель f_Θ (1):

$$\hat{P}_{t+h} = f_\Theta\left(X_{t-L+1:t}\right), \quad h \in \{1,...,H\}, \qquad (1)$$

где: t – номер текущего дискретного шага во временной сетке с интервалом Δt; $h \in \{1,...,H\}$ – горизонт прогноза (также измеряется в шагах Δt); H – максимальное значение горизонта прогнозирования; $\hat{P}_{t+h}$ – прогнозируемое давление в момент времени $t+h$ [Па]; $f_\Theta(\cdot)$ – модель искусственной нейронной сети, параметризованная набором весов Θ; $X_{t-L+1:t} \in R^{L\times d}$ – матрица признаков, сформированная из последних L d-мерных векторов $X_{t_i-L+1:t} = \begin{bmatrix} x_{t-L+1} \\ x_{t-L+2} \\ \vdots \\ x_t \end{bmatrix}$, $x_\tau \in R^d$, при этом x_τ – вектор признаков в момент τ: $\mathbf{x}_\tau = \left[P_\tau, T_\tau, \dot{m}_{\text{in},\tau}, P_{\text{out},\tau}, \underbrace{доп.признаки}_{\text{по наличию: } \dot{m}_{out,\tau}, \text{флаги режима работы газопровода, др.}} \right]$.

Опираясь на соотношение (1) и сформулированную постановку задачи, рассмотрим архитектуру разработанной нейросетевой модели (рисунок 1).

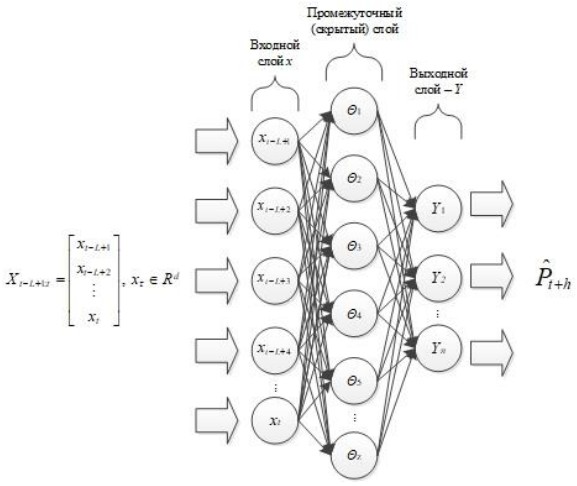


Рис. 1. Структура разработанной LSTM

Слева показано скользящее окно входных признаков X_t L +−1:t ; далее следует скрытый рекуррентный слой

– стек LSTM-ячеек с параметрами Θ. Выходной слой преобразует скрытое (денормализованное) значение в прогноз: в общем случае это может быть вектор (Y_1 ,…,Y_n), а в базовой задаче – единственное скалярное значение давления $\hat{P}_{t+h}$.

Таким образом, рисунок 1 иллюстрирует сквозной поток данных «окно наблюдений → скрытые состояния → прогноз давления», который далее используется и при расчете радиуса возможного взрывного воздействия.

Далее приводится процесс обучения нейронной сети с учетом добавления физических ограничений.

Чтобы обучить описанную модель, из исходного временного ряда формируем обучающую выборку пар вида «*окно наблюдений → целевое значение*» (2):

$$D = \left\{ \left(X_{t_i - L+1:t}, y^{(P)}_{t_i+h}, y^{(T)}_{t_i+h} \right) \right\}_{i=1}^{N_{train}}, \qquad (2)$$

где: t_i – правая граница i-го окна наблюдений (последний доступный момент времени для пары вида «*окно наблюдений → целевое значение*»; $X_{t_i - L+1:t_i} \in R^{L \times d}$ – матрица признаков за L шагов до момента t_i включительно; L – длина окна (в шагах дискретизации Δt); $y^{(P)}_{t_i+h} = P_{t_i+h}$ – истинное давление в момент t_i+h; $y^{(T)}_{t_i+h} = T_{gas,t_i+h}$ – истинная температура в момент t_i+h (при условии функционирования датчика); i=1,...,N_{train} - индекс записей набора данных.

Далее данные без перемешивания делятся во времени на train / validation / test, чтобы проверять модель на «будущих» наблюдениях. Все числовые признаки стандартизируются по статистикам обучающего сегмента (среднее, стандартное отклонение), а физические остатки, вводимые ниже, остаются в натуральных единицах.

Чтобы получать не только статистически точные, но и физически корректные прогнозы давления (и, при наличии, температуры), к базовой MSE необходимо добавить регуляризационные термы – дополнительные компоненты, которые «подтягивают» модель к соблюдению фундаментальных ограничений: баланса массы и энергетической согласованности. Итоговая функция потерь имеет вид (3):

$$L_{total} = \underbrace{L_{data}}_{\text{точность на метках}} + \lambda_{phys} + \underbrace{L_{phys}}_{\text{соблюдение законов}} + \lambda_{reg} \left\| \Theta_{full} \right\|_2^2 , \quad (3)$$

где: L_{total} – итоговая функция потерь, которую мы минимизируем; L_{data} – ошибка на размеченных данных (точность на метках); λ_{phys} – весовой коэффициент, задающий вклад физической регуляризации; L_{phys} – совокупная физическая регуляризация, отвечающая за соблюдение законов сохранения массы и энергии; $\left\| \Theta_{full} \right\|_2^2$ – L2-норма всех параметров модели; λ_{reg} – коэффициент L2-регуляризации, ограничивающий сложность модели.

Функция потерь L_{total} по параметрам Θ нейросетевой модели $f_\Theta(\cdot)$ выполняется последовательно по парам «временной ряд наблюдений → целевое значение». На каждом шаге обновление весов одновременно уменьшает L_{data} , подгоняя прогноз к эталонным значениям, и через $\lambda_{phys} L_{phys}$ мягко заставляет модель соблюдать физические ограничения, тогда как $\lambda_{reg} \left\| \Theta_{full} \right\|_2^2$ сдерживает рост сложности параметров и препятствует переобучению. Благодаря такой формулировке обучение становится интуитивным: точность на данных и соответствие физике учитываются единым критерием, а дальнейшие шаги алгоритма сводятся к стандартной процедуре оптимизации по (3).

Перейдем к этапу обучения нейронной сети. Опираясь на формулу (3), оптимизацию организуем так, чтобы модель одновременно: минимизировала ошибку L_{data} на размеченных данных; через коэффициенты λ_{phys} и L_{phys} соблюдала физические ограничения (баланс массы и энергии); сдерживала избыточную сложность за счет квадратичной (L2) регуляризации, добавляющей невеликий вклад $\lambda_{reg} \left\| \Theta_{full} \right\|_2^2$ к целевой функции.

В результате веса обновляются так, чтобы не только снижать статистическую ошибку, но и избегать физически некорректных прогнозов – особенно это важно при смене режимов работы газопровода.

После формирования обучающей выборки D (2) каждый элемент $X_{t_i - L+1:t}, y^{(P)}_{t_i+h}, y^{(T)}_{t_i+h}$ участвует в прямом проходе модели. Матрица признаков $X_{t_i - L+1:t}$ подается в сеть; адаптивные LSTM-гейты, получая дополнительные физические каналы,

настраивают внутреннее состояние под текущий режим газопровода. На выходе получаем прогноз давления (и, при необходимости, температуры) (4):

$$\left[\hat{P}_{t+h}\left[,\hat{T}_{t+h}\right]\right]=f_{\Theta}\left(X_{t_i-L+1:t}\right)\ . \qquad (4)$$

Выражение (4) фиксирует лишь итог работы сети – прогноз давления (и, при необходимости, температуры). Чтобы понять, каким образом он формируется, рассмотрим внутреннюю логику модифицированной LSTM-ячейки: на каждом такте она получает два взаимодополняющих потока признаков – обычные x_t и физические φ_t и адаптирует свои гейты под текущий режим трубопровода.

Разделение потоков позволяет явно информировать сеть о состоянии трубопровода: для φ_t выделен отдельный набор весов, благодаря чему LSTM-гейты динамически подстраивают внутреннюю память под текущий режим. Полная система уравнений модифицированной ячейки приводится в (5):

$$\begin{cases} i_t=\sigma\left(W^{(i)}x_t+U^{(i)}h_{t-1}+A^{(i)}\phi_t+b^{(i)}\right),\\ f_t=\sigma\left(W^{(f)}x_t+U^{(f)}h_{t-1}+A^{(f)}\phi_t+b^{(f)}\right),\\ o_t=\sigma\left(W^{(o)}x_t+U^{(o)}h_{t-1}+A^{(o)}\phi_t+b^{(o)}\right),\\ \tilde{c}_t=\tanh\left(W^{(c)}x_t+U^{(c)}h_{t-1}+A^{(c)}\phi_t+b^{(c)}\right),\\ c_t=f_t\odot c_{t-1}+i_t\odot\tilde{c}_t,\\ h_t=o_t\odot\tanh(c_t) \end{cases}, \qquad (5)$$

где: i, f, o – входной, забывающий и выходной гейты; $\tilde{c}_t$ – потенциальное обновление состояния ячейки; $W^{(\cdot)}, U^{(\cdot)}$ – матрицы весов, обрабатывающие соответственно текущие признаки x_t и предыдущее скрытое состояние h_{t-1}; $A^{(\cdot)}$ – отдельные матрицы весов, вводящие в расчет физический вектор φ_t; $b^{(\cdot)}$ – bias (коэффициенты постоянного смещения); h_t, c_t – обновленное скрытое состояние и состояние ячейки.

Следовательно, каждый шаг прямого прохода учитывает не только статистику временного ряда, но и текущие физические условия, что позволяет сети формировать физически согласованное скрытое представление и, в итоге, выдавать более надежный прогноз $\hat{P}_{t+h}$.

аким образом, на каждом такте прямого прохода сеть учитывает не только статистику временного ряда, но и текущее состояние объекта – граничные и режимные условия участка трубопровода: входной массовый расход $\dot{m}_{in}$, выходное (или узловое) давление P_{out}, температуру T, активность компрессорных агрегатов и арматуры, а также актуальную конфигурацию/нагрузку участка. Это формирует внутреннее представление, согласованное с физикой процесса, и в итоге дает более надежный прогноз $\hat{P}_{t+h}$. Следующий шаг – проверить, насколько этот прогноз согласуется с фундаментальными законами физики.

Контроль физической согласованности на интервале наблюдений длиной $L\Delta t$

Основная идея заключается в том, чтобы на выбранном отрезке наблюдений $X_{t_i-L+1:t}$ велся «учет газа» в контрольном объеме трубы V: считаем, сколько массы поступило, сколько вышло и какие при этом были давление и температура. Любое несоответствие между притоком и оттоком трактуем как физическую невязку и учитываем в функции потерь через дополнительный регуляризационный член.

Порядок расчета следующий. Сначала определяем, сколько газа находится внутри объема V; затем вычисляем его плотность. Для этого используем начальные значения давления и температуры в начале окна, а также дискретный баланс массы притока-оттока (6):

$$\begin{cases} M_{t_0}=\dfrac{Z P_{t_0} V}{R T_{t_0}}, & t_0=t-L+1,\\ M_{\tau+1}=M_\tau+\Delta t(\dot{m}_{in,\tau}-\dot{m}_{out,\tau}), & \rho_\tau=\dfrac{M_\tau}{V}. \end{cases}, \qquad (6)$$

где M_τ – масса газа в объеме V [кг], при этом $\tau=t_0,\ldots,t-1$; ρ_τ – плотность [кг/м³]; Z – коэффициент сжимаемости (допустимо $Z\approx 1$); R – газовая постоянная.

Если расход на выходе $\dot{m}_{out,\tau}$ не измеряется, пропуски в датасете заполняем оценкой, основанной на перепаде давлений (7):

$$\dot{m}_{\text{out},\tau}\approx K\left(P_\tau-p_\tau^{\text{ext}}\right)_+,\qquad K\geq 0\ , \qquad (7)$$

где: p_τ^{ext} – внешнее (наружное/выходное) давление, а оператор $[x]_+=max(x,0)$ отражает, что газ выходит только при положительном перепаде «внутри – снаружи».

После восстановления M_τ, ρ_τ и $\dot{m}_{\text{out},\tau}$ формализуем две проверки физической согласованности:

- невязка массового баланса $r_\tau^{(m)}$. Невязка показывает, совпадает ли разность притока и оттока с фактическим изменением массы в объеме за шаг Δt: $r_\tau^{(m)}=\dot{m}_{in,\tau}-\dot{m}_{out,\tau}-\dfrac{M_{\tau+1}-M_\tau}{\Delta t}$;
- уравнение состояния газа. Проверяем, согласуются ли прогнозируемые давление и температура с вычисленной плотностью по уравнению состояния реального (сжимаемого) природного газа: $r_\tau^{(e)}=P_\tau-Z R\rho_\tau T_\tau$, где $r_\tau^{(e)}$ – значение ошибки связи «ошибка связи давление-температура-плотность».

В идеальном случае обе невязки стремятся к нулю. Чтобы эти отклонения действительно влияли на

обучение, берем средние квадраты невязок по всем шагам окна и по каждому мини-пакету обучающих примеров (8):

$$L_{\text{mass}} = \overline{\left(r^{(m)}\right)^2}, \qquad L_{\text{energy}} = \overline{\left(r^{(e)}\right)^2} \ , \tag{8}$$

где: L_{mass} – средний квадрат невязки массового баланса; L_{energy} – средний квадрат невязки уравнения состояния. Чем меньше L_{mass}, тем точнее выполняется закон сохранения массы. Уменьшение L_{energy} показывает, что давление, температура и плотность лучше согласуются с уравнением состояния газа.

Далее формируем физическую регуляризацию (9):

$$L_{\text{phys}} = \alpha L_{\text{mass}} + \beta L_{\text{energy}} \ , \tag{9}$$

где: α – весовой коэффициент для невязки баланса массы; β – весовой коэффициент для невязки уравнения состояния.

Обновление весовых коэффициентов с учетом физического закона сохранения массы

Получив на прямом проходе прогнозное значение $\hat{P}_{t+h}$ и вычислив невязки $r_\tau^{(m)}$ и $r_\tau^{(e)}$ на всем временном ряде, формируем суммарный градиент так, чтобы физические требования влияли на веса гейтов сильнее, чем на остальные параметры (10):

$$g = \nabla_\Theta L_{\text{data}} + \lambda_{\text{phys}}\left(\alpha \mathbf{1}_{g_\varphi} \odot \nabla_\Theta L_{\text{phys}} + \beta \mathbf{1}_{g_{\text{rest}}} \odot \nabla_\Theta L_{\text{phys}}\right) + \lambda_{\text{reg}} \nabla_\Theta \| \Theta_{\text{full}} \|_2^2 . \tag{10}$$

где: $\odot$ - компонентная операция умножения - поэлементное произведение, $\nabla\Theta$ - оператор градиента по параметру Θ, $1g_\varphi$ – индикаторная (0/1) маска весов гейтов, связанных с физическими каналами, $1g_{rest}$ – маска всех остальных весов; α, β - весовые коэффициенты, используемые для дифференциального применения эффекта физики на разные группы параметров, $\alpha > \beta$ подчеркивает приоритет гейтов; $\nabla\Theta L_{data}$ - градиент целевой функции (потерь данных), зависящий от параметров Θ, λ_{phys} - гиперпараметр, задающий вес физического компонента в обучении.

Чтобы данные и физика «тянули» модель в одном направлении, мы сравниваем направления их градиентов. Если скалярное произведение $\left\langle g_{\text{data}}^{(\varphi)}, g_{\text{phys}}^{(\varphi)} \right\rangle > 0$, т.е. градиент по данным направлен «против» физики, мы проецируем его на ортогональное физике подпространство (11):

$$g_{\text{data},\perp}^{(\varphi)} = g_{\text{data}}^{(\varphi)} - \frac{\left\langle g_{\text{data}}^{(\varphi)}, g_{\text{phys}}^{(\varphi)} \right\rangle}{\left\langle g_{\text{phys}}^{(\varphi)}, g_{\text{phys}}^{(\varphi)} \right\rangle} g_{\text{phys}}^{(\varphi)} \ , \tag{11}$$

где: $g_{\text{data}}^{(\varphi)}$ – градиент ошибки на данных L_{data} по весам матриц $A^{(\cdot)}$, через которые в ячейку подаются физические признаки φ_t; $g_{\text{phys}}^{(\varphi)}$ – градиент физической регуляризации L_{phys} по тем же весам; $\langle a, b \rangle$ – евклидово (скалярное) произведение векторов a и b; $g_{\text{data},\perp}^{(\varphi)}$ – составляющая градиента данных, ортогональная физическому градиенту; именно она используется при обновлении параметров.

Таким образом, после ортогонализации (11) мы получаем скорректированный градиент g: его компонента, связанная с данными, больше не тянет веса в сторону, противоречащую физике, а физическая часть дополнительно усилена для матриц $A^{(\cdot)}$, через которые в ячейку поступают технологические признаки. Далее этот градиент подается в модифицированный метод обучения Adam-PI, что позволяет при обновлении параметров одновременно учитывать и статистику временных рядов, и закон сохранения массы.

Теперь, когда градиент g вычислен, т.е. составляющая, связанная с данными, согласована с физикой, а веса «физических» гейтов выделены в g_ϕ, его передают в фазу обновления Adam-PI. На этом шаге базовая скорость обучения η масштабируется коэффициентом $k_\varphi \geq 1$ именно для тех весов, через которые проходят физические каналы; остальные параметры обновляются с обычным шагом.

Следовательно, пусть $\hat{m}_t$ и $\hat{v}_t$ – bias-корректированные моменты первого и второго порядков соответственно (12):

$$\hat{m}_t = \frac{m_t}{1-\beta_1^t}, \qquad \hat{v}_t = \frac{v_t}{1-\beta_2^t} \ , \tag{12}$$

где t – номер шага оптимизации (счетчик итераций), β_1^t и β_2^t – коэффициенты затухания экспоненциальных средних.

Тогда окончательное обновление весов выполняем по соотношению (13):

$$\Theta \leftarrow \Theta - \eta\left(k_\varphi \mathbf{1}_{g_\varphi} + \mathbf{1}_{g_{\text{rest}}}\right) \odot \frac{\hat{m}_t}{\sqrt{\hat{v}_t} + \varepsilon} \ , \tag{13}$$

где ε – малое положительное число (например, 10^{-8}), добавляемое в знаменатель, чтобы предотвратить деление на ноль и обеспечить численную устойчивость вычислений.

В результате «физические» веса получают более сильный импульс, что ускоряет устранение нарушений законов сохранения, тогда как остальные параметры обновляются в обычном режиме Adam.

Таким образом, этап обучения завершен: сеть выдает физически согласованные прогнозы давления $\hat{P}_{t+h}$ (и, при необходимости, температуры $\hat{T}_{t+h}$), причем невязки по массе и уравнению состояния остаются минимальными. Логичным продолжением является использование этих прогнозов в прикладных задачах безопасности. В нашей системе [12] следующий шаг – расчет

радиуса возможного взрыва газопровода на основе $\hat{P}_{t+h}$. От точности этого прогноза зависит оценка мгновенного перепада давления при разгерметизации, а значит – моделирование ударной волны и радиуса взрывного воздействия.

2.2. Моделирование радиуса взрыва газопровода

Сначала посчитаем, сколько энергии накоплено в газе внутри поврежденного участка трубопровода. Для этого вычисляем объем газа и ту часть его теплотворной способности, которая потенциально может перейти в энергию взрыва (14):

$$V_g = \pi\left(\frac{D}{2}\right)^2 L, \qquad E_{\text{gas}} = \eta_g V_g Q_H \ , \tag{14}$$

где: D – диаметр трубы, L – длина поврежденного участка, V_g – объем газа [м³]; η_g – доля теплоты сгорания газа, которая переходит в энергию взрыва; Q_H – низшая удельная теплота сгорания газа; E_{gas} – расчетная (оценочная) энергия высвобождения [Дж].

Далее, используя прогнозируемое давление, $\hat{P}_{t+h}$, определяем критический безразмерный параметр проницаемости B_{cr} по интерполяционной формуле (15):

$$B_{cr}\left(\hat{P}_{t+h}\right) = B_{cr}^{\min} + \left(B_{cr}^{\max} - B_{cr}^{\min}\right)\frac{\hat{P}_{t+h}}{P_{\max}} \ , \tag{15}$$

где: $B_{cr}^{\min}$ и $B_{cr}^{\max}$ – эмпирические предельные значения (нижний и верхний уровни критического запаса проницаемости для рассматриваемого типоразмера трубы); $P_{\max}$ – предельно-допустимое рабочее давление.

Следующим шагом будет связывание давления, температуры и газовой постоянной (16), что позволит перейти к расчету радиуса взрывного воздействия.

$$N = \frac{\hat{P}_{t+h}}{RT_t} \ , \tag{16}$$

где: R – универсальная газовая постоянная, связывающая давление, объем и температуру идеального газа в уравнении состояния; N – характеристический параметр истечения.

Совмещаем вклад давления («перколяционный» член) и запас энергии газа E_{gas} из (14), получая предварительный радиус взрыва (17):

$$R_0 = \left(\frac{B_{cr}}{N}\right)^{1/3} + k_s\left(\frac{E_{\text{gas}}}{4184}\right)^{1/3}, \qquad R_0 \le R_{\max} \ , \tag{17}$$

где: k_s – эмпирический коэффициент масштабирования энергетической компоненты: он переводит кубический корень высвобожденной энергии E_{gas} (приведенной к мегаджоулям) в эквивалентный вклад в радиус поражения, м; R_{max} – максимальный радиус взрыва; R_0 – радиус взрыва без учета рельефа.

Чтобы получить итоговое значение радиуса, скорректируем R_0 с учетом цифровой модели рельефа (18):

$$\begin{cases} F_{\text{terr}} = \text{terrainFactor}(\varphi, \lambda) \in [0{,}8; 1{,}05], \\ R_{\text{expl}} = \min\left(R_0 F_{\text{terr}}, R_{\max}\right). \end{cases} \ , \tag{18}$$

где: F_{terr} – коэффициент, отражающий наклон и высоту местности; R_{expl} – итоговый радиус зоны опасного воздействия [м].

Таким образом, если прогнозируемое давление $\hat{P}_{t+h}$ не достигает критического порога, то $R_{\text{expl}} \approx 0$ и система уведомляет, что угрозы аварийной ситуации нет (рисунок 2).

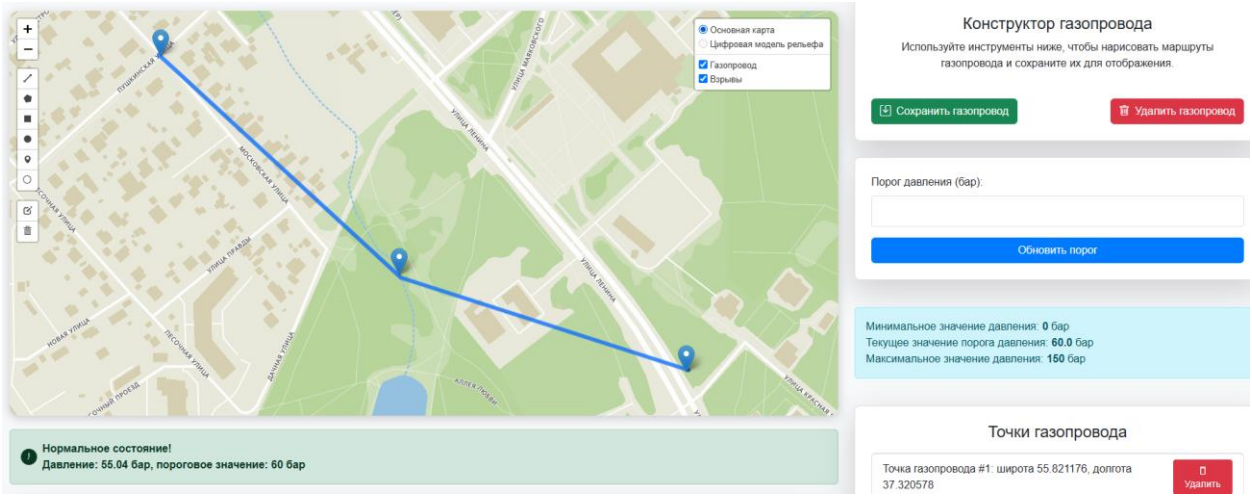


Рис. 2 . Состояние газопровода при прогнозном значении давления, не превышающем заданный порог

При аварийном давлении строится окружность радиуса R_{expl} и анимированная ударная волна. Численные значения выводятся оператору (рисунок 3).

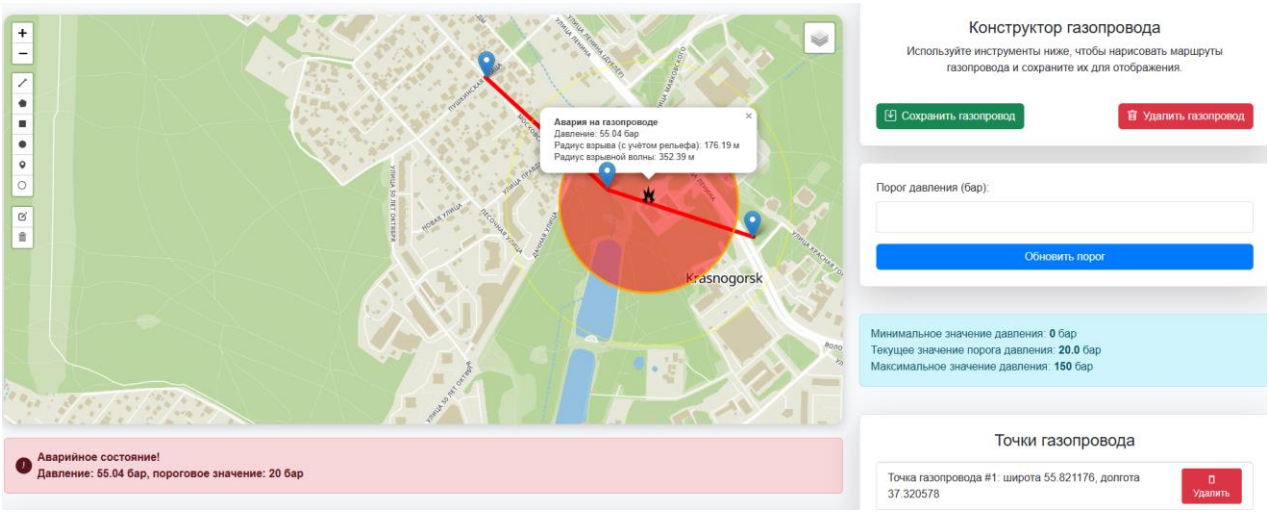


Рис. 3 . Состояние газопровода при прогнозном значении давления, превышающем заданный порог (аварийная ситуация)

3. Заключение

В представленной работе разработана и экспериментально обоснована гибридная технология краткосрочного прогнозирования давления в газопроводе, сочетающая возможности глубоких рекуррентных нейросетей и физически интерпретируемых ограничений. Ключевым элементом архитектуры является модифицированная ячейка Adaptive LSTM с механизмом внимания, в которую наряду с временными признаками включены физические параметры процесса – массовый расход и температура. В качестве функции потерь использовано составное выражение, объединяющее стандартную метрику точности (MSE) и физическую регуляризацию по массе и уравнению состояния, что обеспечивает устойчивость модели к аномалиям и переходным режимам работы трубопровода.

Предложенный подход дополнительно усилен механизмом ортогонализации градиентов и дифференцированной корректировкой шагов обучения для весов, связанных с физическими каналами. Это позволяет избежать конфликта между статистическими закономерностями и физическими законами, что особенно важно при наличии шумов или нестабильных режимов. Разработанное решение реализовано в виде промышленно-ориентированного веб-приложения с возможностью онлайн-прогнозирования и расчета радиуса потенциального взрыва. Такой подход не только повышает точность и интерпретируемость моделей, но и делает их практическим инструментом для систем поддержки принятия решений в области промышленной безопасности.

Благодарности

Работа поддержана Министерством науки и высшего образования Российской Федерации в рамках базовой части государственного задания для высших учебных заведений # FRRR-2026-0006.

Список используемых источников

1. Chaczykowski M. Transient flow in natural gas pipeline – The effect of pipeline thermal model // Applied Mathematical Modelling. 2010. Vol. 34. No. 4. P. 1051–1067. DOI: 10.1016/j.apm.2009.07.017.
2. Wen K., Jiao J., Zhao K., Yin X., Liu Y., Gong J., Li C., Hong B. Rapid transient operation control method of natural gas pipeline networks based on user demand prediction // Energy. 2023. Vol. 264. Art. 126093. DOI: 10.1016/j.energy.2022.126093.
3. Chen Q., Guan A., Chen F., Huang W., Jin A., Gadalla M., Wang B. A transient gas pipeline network simulation model for decoupling the hydraulic-thermal process and the component tracking process // Energy. 2024. Vol. 301. Art. 131613. DOI: 10.1016/j.energy.2024.131613.
4. Gong J., Shi G., Wang S., Wang P., Chen B., Chen Y., Wang B., Yu B., Jiang W., Li Z. Efficient super-resolution of pipeline transient process modeling using the Fourier Neural Operator // Energy. 2024. Vol. 302. Art. 131676. DOI: 10.1016/j.energy.2024.131676.
5. Zhang X., Shi J., Yang M., Huang X., Usmani A. S., Chen G., Fu J., Huang J., Li J. Real-time pipeline leak detection and localization using an attention-based LSTM approach // Process Safety and Environmental Protection. 2023. Vol. 174. P. 460–472. DOI: 10.1016/j.psep.2023.04.020.
6. Ye J., Do N. C., Zeng W., Lambert M. F. Physics-informed neural networks for hydraulic transient analysis in pipeline systems // Water Research.

2022. Vol. 221. Art. 118828. DOI: 10.1016/j.watres.2022.118828.

7. Ye J., Zeng W., Do N. C., Lambert M. F. Reconstructing transient pressures in pipe networks from local observations by using physics-informed neural networks // Water Research. 2024. Vol. 257. Art. 121648. DOI: 10.1016/j.watres.2024.121648.
8. Tjuatja V., Keramat A., Rahmanshahi M., Duan H.-F. Spectral physics-informed neural network for transient pipe flow simulation // Water Research. 2025. Vol. 279. Art. 123427. DOI: 10.1016/j.watres.2025.123427.
9. Xu Y., Zhou L., Lu Y., Hu Y., Zhang Y. Physics-informed neural networks involving unsteady friction for transient pipe flow // Water Research. 2025. Vol. 283. Art. 123779. DOI: 10.1016/j.watres.2025.123779.
10. Bu Y., Swartz C. L. E., Wu C. A two-level MPC method for the operation of a gas pipeline system under demand variation // Computers & Chemical Engineering. 2024. Vol. 183. Art. 108597. DOI: 10.1016/j.compchemeng.2024.108597.
11. Abdulnagimov A.I., Antonov V.V., Chepaykin A.S., Palchevsky E.V., Nasyrov N.T. Neural Network Simulator of Gas Turbine Engines for Prototyping and Debugging of Control Systems on Unsteady Modes / A. I. Abdulnagimov, V. V. Antonov, A. S. Chepaykin [et al.] // Mechatronics, Automation, Control. – 2025. – Vol. 26, No. 1. – P. 22-27. DOI 10.17587/mau.26.22-27.
12. Django Software Foundation. The web framework for perfectionists with deadlines // *Django*. [Электронный ресурс]. URL: https:// www.djangoproject.com/ (дата обращения: 10.09.2025).
13. ELForecasting. Pipeline Failure Prediction System // *ELForecasting*. [Электронный ресурс]. URL: https://gas.elforecasting.com/ (дата обращения: 10.09.2025).

Регрессионная модель дифференцируемой оценки опасности столкновения мобильного робота с препятствием

Фаттахов А. Р.

Уфимский университет науки и технологий
каф. ВМиК
Уфа, Россия
e-mail: fattahovainur0@gmail.com

Инсапов К. И.

Уфимский университет науки и технологий
каф. ВМиК
Уфа, Россия
e-mail: todosomeawesome@gmail.com

Миронов К. В.

Уфимский университет науки и технологий
каф. ВТиЗИ
Уфа, Россия
e-mail: mironovconst@gmail.com

Аннотация[1]

В данной работе рассматривается подход к обеспечению уклонения мобильного робота от столкновения с препятствиями при движении вдоль заданного пути. Задача локальной оптимизации траектории формулируется с учетом необходимости следовать заданному пути и находиться на безопасном расстоянии от препятствий. Основная сложность такого подхода связана с необходимостью представить опасность столкновения в виде непрерывно дифференцируемой функции от положения робота на заданной карте. В литературе такая функция задается либо через аппроксимацию препятствий простыми геометрическими формами, либо с помощью вычислительно емких нейросетевых моделей. В данной работе предлагается иной подход, основанный на использовании низкоресурсного алгоритма классического обучения, способного быстро обучиться на аппроксимацию заданной карты препятствий. Суть подхода состоит в аппроксимации функции опасности столкновений на основе полиномиальной регрессионной модели. Разработанный метод реализован с использованием численного солвера Acados для решения задач оптимального управления. Сравнительные эксперименты показали, что использование полиномиальной модели 12-й степени обеспечивает повышение качества планирования траекторий. Суммарное время обучения и оптимизации траектории составляет несколько десятков миллисекунд, что удовлетворяет стандартному требованию для бортовых систем – обеспечивать возможность перепланирования с частотой не менее 10 Гц.

Ключевые слова: мобильные роботы; планирование движения; избегание столкновений; регрессия; оптимизация траектории.

1. Введение

Возможность самостоятельного планирования движения мобильного робота является обязательным условием его автономного функционирования. В современных робототехнических системах планирование как правило выполняется в два этапа. Сначала производится глобальное планирование. Задача глобального планирования состоит в том, чтобы на основе известных начальной и конечной конфигураций робота, а также карты препятствий, составить последовательность промежуточных положений, перемещаясь по отрезкам между которыми робот может достичь конечного положения. Эта задача успешно решается с использованием эвристического [1] или вероятностного [2, 3] поиска.

После того как спланирован глобальный путь, выполняется второй этап – локальное планирование, в ходе которого глобальный путь преобразуется в локальную траекторию, учитывающую кинодинамические ограничения мобильного робота. Локальная траектория часто строится непосредственно в ходе движения робота, в режиме скользящего окна. Это позволяет учесть при планировании данные о реальном положении препятствий, поступающие от сенсоров в реальном времени. Задача локального планирования часто решается вероятностными [4] или численными [5] методами оптимального управления. Численные методы обеспечивают более надежные решения, но требуют, чтобы задача оптимизации была сформулирована в виде конкретной системы уравнений (см. раздел 2). Вероятностные методы не

[1] Труды XI Международной научной конференции "Информационные технологии интеллектуальной поддержки принятия решений", 11-13 ноября, Уфа-Чандигарх-Фандербейлпарк, 2025

предъявляют такого требования, однако могут давать некорректные или осциллирующие решения.

В данной работе рассматривается подход к уклонению от препятствий при решении задачи локального планирования численными методами. Такое уклонение может быть реализовано двумя способами: либо в постановку задачи добавляется система ограничений-неравенств, обеспечивающих отсутствие столкновений [5], либо в функцию стоимости добавляется дополнительное слагаемое (репульсивный потенциал препятствий), содержащее оценку опасности столкновения с препятствиями [6]. Второй подход обеспечивает более гибкое уклонение от столкновений и может сходится к корректному решению, даже если глобальный план задевает препятствия. Сложность этого подхода в том, что функция опасности столкновений должна быть непрерывно-дифференцируемой на всей карте: ее градиент используются, чтобы «выталкивать» траекторию из опасных областей.

Можно выделить два представленных в литературе подхода к заданию функции отталкивающего потенциала: аппроксимацию карты препятствий набором простых геометрических форм, и оценка потенциала с помощью обучаемых моделей. Первый подход рассматривается в ряде работ: в [5] используется аппроксимация свободного пространства кругами, в [7] – аппроксимация свободного пространства квадратами, в [6] – аппроксимация препятствий кругами, в [8] – аппроксимация границ препятствий ломаной, в [9] – аппроксимация препятствий выпуклыми многоугольниками. Минусом такого подхода является сложность учета препятствий на карте произвольной формы.

При использовании второго подхода опасность столкновения оценивается нейросетевой моделью. В ранних работах [10, 11] нейросетевая модель предобучалась на одну заданную карту и работала только в ее пределах. В последующих работах [12, 13] обученная нейросетевая модель принимает на вход актуальные данные от препятствий и, таким образом, может использоваться в различных средах. Такая модель становится достаточно ресурсоемкой и время вычисления на ней может составлять сотни миллисекунд при том, что стандартным требованием для автономных транспортных средств является возможность перепланирования с частотой 10 Гц. Процедура обучения нейросети на отдельную карту в свою очередь является весьма ресурсоемкой процедурой, которая не может осуществляться в реальном времени при каждом перепланировании.

Мы предлагаем метод аппроксимации препятствий легкой обучаемой моделью, которая может настроиться на заданную карту в течение короткого времени. Это позволяет выполнить как обучение модели, так и оптимизацию траектории в рамках одной операции перепланирования траектории, выполнения в реальном времени. В качестве модели используется полиномиальная регрессионная модель заданной степени.

В следующем разделе приводится математическая постановка задачи локальной оптимизации траектории, рассматриваемая в данной работе. Затем, в разделе 3 описывается предложенный подход. В разделе 4 описываются особенности реализации и результаты вычислительных экспериментов. Раздел 5 содержит итоговые замечания.

2. Управление движением по траектории

2.1. Формальная постановка задачи оптимизации траектории транспортного средства

Рассмотрим задачу MPC для нелинейной системы с непрерывной динамикой и кусочно-постоянным управлением:

$$\{x_{opt}[i], y_{opt}[i]\}_{i=k}^{k+m} = \arg(\min_{x,u,p} \sum_{i=k}^{k+m} J(x[i], y[i], p[i])) \quad (1)$$

$$\frac{dx_{\mu}[i]}{dt} = f_{\mu}(x[i], u[i], p[i]) \quad (2)$$

Оптимизируемый процесс управления состоит из m шагов, на каждом из которых производится переключение управляющих воздействий. $x[i]$ - μ-мерный вектор переменных, описывающих состояние объекта управления в начале [i]-го шага, $u[i]$ - v-мерный вектор управляющих воздействий, непрерывно прилагаемых к объекту управления в течение [i]-го шага, $p[i]$ - к-мерный вектор параметров процесса, используемых в расчетах для [i]-го шага. (1а) задает функцию стоимости J, складывающуюся из значений $J[i]$, рассчитанных по единому алгоритму для каждого шага. Суть оптимизации в том, чтобы подобрать значения управляющих воздействий, обеспечивающих минимальное значение J. (2) задает непрерывную модель динамики процесса (для каждой переменной состояния - зависимость ее производной от остальных состояний, управляемых переменных и параметров процесса в любой момент времени). Для дискретной динамики уравнение (2) примет несколько другой вид

$$x_{\mu}[i+1] = f_{\mu}(x[i], u[i], p[i]) \quad (3)$$

Процедура оптимизации начинается с начального приближения $\{x_{init}[i], u_{init}[i]\}^{k+m}{}_{i=k}$ которое может не удовлетворять модели динамики (в нашем случае значения x_{init} берутся с глобального плана, построенного алгоритмом Theta*, а u_{init} заполняются

нулями). На каждой итерации оптимизации солвер смещает текущую траекторию (на первом шаге траектория равна начальному приближению, на каждом следующем - результату предыдущей итерации) так, чтобы снизить значение функции стоимости *J*, при этом удовлетворяя модели динамики и ограничениям.

2.2. Модель динамики для колесного робота

Вектор состояний включает координаты колесного робота и его ориентацию, также в зависимости от модели динамики он может включать линейную и угловую скорости, угол поворота рулевых колес и т.д. Вектор управляющих воздействий в зависимости от модели может включать линейные и угловые скорости и ускорения колесного робота, скорость поворота рулевого колеса и т.д. Например, можно указать модель с дифференциальным приводом (повороты платформы обеспечиваются вращением левых и правых колес с разной скоростью; дифференциальный привод имеет используемый в лабораторных экспериментах мобильный робот Husky) и велосипедную модель (описывает транспортные средства с рулевыми колесами, в т.ч. автомобили). Для велосипедной модели вектор состояний $\mathbf{x} = (x, y, v, \theta, \delta)^T$, где *x*, *y*-координаты транспортного средства, *v* - линейная скорость, θ - направление движения транспортного средства, δ - угол поворота рулевых колес. Вектор управлений $\mathbf{u} = (a, \omega)^T$, где *a* - линейное ускорение, ω - скорость поворота рулевых колес. (2) принимает следующий вид:

$$\begin{aligned}
\frac{dx}{dt} &= \upsilon \cos\theta, \\
\frac{dy}{dt} &= \upsilon \sin\theta, \\
\frac{d\upsilon}{dt} &= a, \\
\frac{d\theta}{dt} &= \frac{\upsilon \tan\delta}{L}, \\
\frac{d\delta}{dt} &= \omega,
\end{aligned} \tag{3}$$

где *L* - const - расстояние между осями колесного робота.

Модель с дифференциальным приводом:

$$\begin{aligned}
\frac{dx}{dt} &= \upsilon \cos\theta, \\
\frac{dy}{dt} &= \upsilon \sin\theta, \\
\frac{d\upsilon}{dt} &= a, \\
\frac{d\theta}{dt} &= \omega,
\end{aligned} \tag{4}$$

В отличие от предыдущей модели здесь ω - это угловая скорость всей платформы.

2.3. Функция стоимости

Рассмотрим функцию стоимости в задаче оптимизации, определяющую степень оптимальности для заданного положения робота при движении по траектории:

$$J[i] = J_s(x[i], u[i], x_r[i]) + J_o(x[i], p_o[i]) \tag{5}$$

Минимизация слагаемого J_s (гравитационный потенциал) обеспечивает близость положения робота к референтному значению x_r взятому из глобального плана, в то время как минимизация слагаемого J_o (репульсивный потенциал препятствий) обеспечивает отталкивание траектории на безопасное расстояние от препятствий. Вектор $p_o[i]$ содержит параметры, описывающие препятствия. Вектор параметров задачи из системы (1) в этом случае определяется как $p[i] = ((x_r[i])^T, (p_o[i])^T)^T$.

$$J_s[i] = \sum_{j=1}^{\mu} w_{xj} (x_j[i] - x_{j(ref)}[i])^2 + \sum_{k=1}^{v} w_{uk} u_k^2[i] \tag{6}$$

Здесь w_{xj}, w_{uk} весовые коэффициенты соответствующих слагаемых, $x_{j(ref)}[i]$ – референсные значения (берутся с глобального плана движения в виде ломаной).

Репульсивный потенциал должен принимать высокие значения внутри препятствий, низкие вдали от препятствий, а у границы препятствий должен иметь максимальный градиент (в ходе процедуры оптимизации градиент определяет направление, в котором оптимизатор "выталкивает" траекторию подальше от препятствий).

Пусть известна бинарная карта препятствий (Occupancy Grid - плоскость, в которой движется робот, расчерчена на клетки-ячейки, каждая ячейка помечается как занятая или свободная). На основе этой карты для любой точки можно алгоритмически рассчитать так называемую функцию расстояния со знаком (Signed Distance Function, *SDF*) - она равна расстоянию до ближайшей границы препятствия со знаком плюс (если точка вне препятствия) или минус (если точка внутри препятствия). На основе *OG* и *SDF* строится Signed Distance Map - матрица

той же размерности, что *OG*, в каждой ячейке которой указывается значение *SDF*.

Значения потенциала можно рассчитать как сигмоидную функцию от *SDF*. Пример для логистической функции:

$$J_o = w_1(1-(1+e^{-(SDF-s_0)})^{-w_2} \quad (7)$$

Здесь используются весовые коэффициенты: w_1, w_2, s_0 – сдвиг сигмоиды по оси абсцисс.

Функция расчета репульсивного потенциала от SDF удовлетворяет требованиям, сформулированным выше: потенциал принимает высокие значения внутри препятствий, низкие вдали от препятствий, а у границы препятствий имеет максимальный градиент. Проблема в том, что SDF сама по себе не является дифференцируемой функцией. Она рассчитывается алгоритмически, ее значения можно подсчитать в любой точке, но задать ее аналитический градиент в общем случае нельзя.

3. Регрессионная модель

Мы предлагаем использовать модель полиномиальной регрессии для аппроксимации карты знаковых расстояний (Signed Distance Map, *SDM*). Исходными признаками являются две независимые переменные, x_1 и x_2 (в нашем случае координаты на карте). Модель строится путем преобразования исходного пространства признаков в расширенное, включающее все полиномиальные комбинации x_1 и x_2 вплоть до заданной степени n. Для двух признаков это преобразование генерирует (n+1)(n+2)/2 уникальных членов. Это связано с тем, что одночленов суммарной степени n ровно n+1, и при этом также учитываются меньшие степени.

Полученная модель остается линейной относительно своих параметров, что обеспечивает возможность эффективной процедуры обучения методом наименьших квадратов, временная сложность которого равна O(whn[4]), где w и h – ширина и высота карты в пикселях соответственно. Результат обучения модели представляет собой полином вида (8).

$$\hat{y} = t[0] + t[1]x_1 + t[2]x_2 + t[3]x_1^2 + t[4]x_1x_2 + \ldots \quad (8)$$

Идея метода состоит в поиске вектора весовых коэффициентов (в нашем обозначении *t*), которые при подстановке в полином минимизировали бы квадратичное отклонение от эталонных значений.

$$t = \arg(\min_l || Xl - y ||_2^2) \quad (9)$$

Где *X* – матрица состоящая из строк вида (10).

$$X_k = (1, x_{1i}, x_{2j}, x_{1i}^2, x_{1i}x_{2i}, x_{2j}^2, \ldots) \quad (10)$$

В свою очередь *y* – векторизация матрицы значений *SDF* из *SDM* (11).

$$y = (SDM[1,1], SDM[1,2], \ldots, SDM[2,1], \ldots, SDM[h,w])^T \quad (11)$$

i-ая строка матрицы *X* должна соответствовать *i*-му элементу *y* по индексам.

Весовые коэффициенты находятся следующим образом (12).

$$t = (X^T X)^{-1} X^T y \quad (12)$$

Мы используем метод полиномиальной регрессии для аппроксимации *SDF*, после чего обученная модель используется для решения задачи локальной оптимизации траектории.

Вектор параметров p_o о препятствиях в данном подходе состоит из значений весовых коэффициентов *t*.

4. Эксперименты

4.1. Реализация

Алгоритмы были реализованы на языке программирования Python. Для формирования карт знаковых расстояний (Signed Distance Map, *SDM*) использовалась специализированная библиотека scikit-fmm[2]. Аппроксимация полученных значений *SDF* выполнялась с помощью библиотеки scikit-learn (полиномиальная регрессия, реализованная как комбинация преобразования полиномиальных признаков и линейной регрессии). Построенная регрессионная модель уже участвовала в формулировке и решении задачи локальной оптимизации траектории, для которой использовался решатель Acados[3] через его Python-интерфейс.

Интеграция модели, построенной средствами scikit-learn, в задачу Acados потребовала ее символьного представления. Это было достигнуто путем извлечения коэффициентов полинома из обученной модели и явного формирования полиномиальной функции с использованием символьных переменных (с помощью символьного фреймворка CasADi[4]).

4.2. Бейзлайны

Сравнение предложенного метода проводилось с двумя другими подходами к решению данной задачи. Первый из них отличается от предложенного нами лишь тем, что аппроксимация карты знаковых расстояний производится с помощью символьного фреймворка CasADi. Второй подход (CIAO star [5]) не использует дополнительное слагаемое репульсивного потенциала и подразумевает

[2] URL: https://github.com/scikit-fmm/scikit-fmm

[3] URL: https://docs.acados.org/

[4] URL: https://web.casadi.org/

добавление нелинейного ограничения-неравенства о нахождении робота внутри безопасного круга для каждой точки пути.

4.3. Результаты

Приведем пример работы обученной модели при 18 степени полинома – сформированное потенциальное поле для одной из карт размером 64 на 64 пикселя. На изображении (Рис. 1) справа представлен необработанный файл карты, слева график, где цветовым градиентом (по возрастанию от фиолетового к жёлтому) отображены значения репульсивного потенциала.

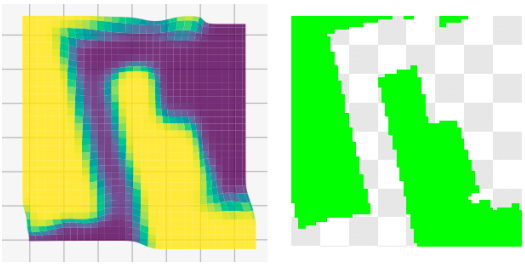


Рис. 1. Потенциальное поле

На карте строится траектория движения робота (красная линия) по заданной ломаной (синяя линия) (Рис. 2). Она также построена с помощью полиномиальной регрессии.

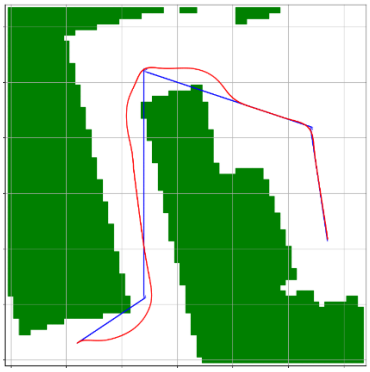


Рис. 2. Траектория движения робота

Следует заметить, что тестирование проводилось на картах с ресурса Moving AI Lab[5] со стороной пикселя соответствующей 0,5 м.

После наглядного примера перейдём к числовым данным. Начнём с времени обучения модели (или аналогичного процесса) при аппроксимации значений *SDF*. Здесь и далее σ – среднее квадратичное отклонение.

Обучение модели для аппроксимации с помощью предложенного метода сопоставимо по времени с интерполяцией CasADi при 12-й степени полинома, но медленнее предобработки карты и догадки методом CIAO star.

Таблица 1. Обучение модели для аппроксимации *SDF*. Размер карты 64x64, средние значения на 5 картах.

Метод аппроксимации		Среднее время обучения модели	σ (в пикселях)
sklearn регрессия, полином степени:	6	0.0122	2.87
	9	0.0177	1.04
	12	0.0334	0.41
	15	0.0950	0.22
	18	0.1276	0.17
	21	0.1488	0.14
casadi interpolant		0.0355	-
CIAO star (безопасные круги)		0.0071	-

Таблица 2. Обучение модели для аппроксимации *SDF*. Размер карты 128x128, средние значения на 5 картах.

Метод аппроксимации		Среднее время обучения модели	σ (в пикселях)
sklearn регрессия, полином степени:	6	0.042	11.01
	9	0.121	4.27
	12	0.228	2.06
	15	0.338	1.05
	18	0.537	0.79
	21	0.643	0.70
casadi interpolant		0.242	-
CIAO star (безопасные круги)		0.023	-

Таблица 3. Обучение модели для аппроксимации *SDF*. Размер карты 256x256, средние значения на 5 картах.

[5] URL: https://movingai.com/benchmarks/grids.html

Метод аппроксимации		Среднее время обучения модели	σ (в пикселях)
sklearn регрессия, полином степени:	6	0.226	58.0
	9	0.513	26.0
	12	1.058	13.4
	15	1.830	7.3
	18	2.368	5.4
	21	2.922	4.7
casadi interpolant		1.822	-
CIAO star (безопасные круги)		0.071	-

Теперь перейдём к рассмотрению данных, полученных при построении траекторий (а именно при последовательной оптимизации небольших горизонтов). Для оценки качества также рассчитывается минимальное расстояние до препятствия на всей длине пути. Они по аналогии представлены на разных размерах карт.

Таблица 4. Локальная оптимизация траектории с помощью готовой модели. Размер карты 64x64, средние значения по 5 картам.

Метод аппроксимации		Среднее время решения горизонта	Мин. расстояние до препятствия по пути (в пикселях)
sklearn регрессия, полином степени:	6	0.0024	-3.22
	9	0.0025	-0.50
	12	0.0026	0.97
	15	0.0026	1.45
	18	0.0027	1.45
	21	0.0027	1.45
casadi interpolant		0.0028	1.95
CIAO star (безопасные круги)		0.0025	-1.06

Таблица 5. Локальная оптимизация траектории с помощью готовой модели. Размер карты 128x128, средние значения по 5 картам.

Метод аппроксимации		Среднее время решения горизонта	Мин. расстояние до препятствия по пути (в пикселях)
sklearn регрессия, полином степени:	6	0.0025	-1.45
	9	0.0026	-1.50
	12	0.0026	-1.30
	15	0.0026	-0.35
	18	0.0028	-0.35
	21	0.0028	-0.35
casadi interpolant		0.0026	1.76
CIAO star (безопасные круги)		0.0040	-0.50

Таблица 6. Локальная оптимизация траектории с помощью готовой модели. Размер карты 256x256, средние значения по 5 картам.

Метод аппроксимации		Среднее время решения горизонта	Мин. расстояние до препятствия по пути (в пикселях)
sklearn регрессия, полином степени:	6	0.0026	-4.4
	9	0.0029	-4.4
	12	0.0032	-4.4
	15	0.0031	-2.2
	18	0.0032	-1.5
	21	0.0033	-1.1
casadi interpolant		0.0033	2.3
CIAO star (безопасные круги)		0.0056	-1.9

Решения, использующие регрессионную модель, часто более выгодны по времени на стадии локальной оптимизации траектории, чем метод, применяющий интерполяцию из CasADi, но менее выгодны, чем метод CIAO star.

5. Заключение

В ходе работы мы разработали модели полиномиальной регрессии при решении задач локальной оптимизации траектории. Было установлено, что применение этих моделей обычно дает корректные результаты аппроксимации SDF и

оптимизации траектории на картах размером 64x64 при степени полинома от 12-й включительно, однако начиная со степени 15 данный подход проигрывает во времени обучения. Несмотря на возникающий вопрос точности аппроксимации SDF, при небольших размерах карт и степенях полинома порядка 12, использование регрессии sklearn может быть оправдано.

Список используемых источников

1. Hart, Peter E., Nilsson, Nils J., Raphael, Bertram, 1968. A formal basis for the heuristic determination of minimum cost paths. IEEE Trans. Syst. Sci. Cybern. 4(2), 100–107. http://dx.doi.org/10.1109/TSSC.1968.300136.
2. Kavraki, L.E., Svestka, P., Latombe, J.-C., Overmars, M.H., 1996. Probabilistic roadmaps for path planning in high-dimensional configuration spaces. IEEE Trans. Robot. Autom. 12 (4), 566–580. http://dx.doi.org/10.1109/70.508439.
3. LaValle, Steven M., Kuffner, Jr., James J., 2001. Randomized kinodynamic planning. Int. J. Robot. Res. 20 (5), 378–400. http://dx.doi.org/10.1177/02783640122067453.
4. Williams, Grady, Drews, Paul, Goldfain, Brian, Rehg, James M, Theodorou, Evangelos A, 2016. Aggressive driving with model predictive path integral control. In: 2016 IEEE International Conference on Robotics and Automation. ICRA, pp. 1433–1440. http://dx.doi.org/10.1109/ICRA.2016.7487277.
5. Schoels, Tobias, Palmieri, Luigi, Arras, Kai O., Diehl, Moritz, 2020a. An NMPC approach using convex inner approximations for online motion planning with guaranteed collision avoidance. In: 2020 IEEE International Conference on Robotics and Automation. ICRA, pp. 3574–3580. http://dx.doi.org/10.1109/ICRA40945.2020.9197206.
6. Локальное планирование траектории колесного робота в ограниченной среде на основе модельного прогнозирующего управления / Мухаммад Алхаддад [и др.] // Робототехника и техническая кибернетика. – Т. 11. - № 3. – Санкт-Петербург : ЦНИИ РТК. – 2023. – С. 205-214. – Текст : непосредственный.
7. Schoels, Tobias, Rutquist, Per, Palmieri, Luigi, Zanelli, Andrea, Arras, Kai O., Diehl, Moritz, 2020b. CIAO*: MPC-based safe motion planning in predictable dynamic environments. IFAC- Pap. (ISSN: 2405-8963) 53 (2), 6555–6562. http://dx.doi.org/10.1016/j.ifacol.2020.12.072, 21st IFAC World Congress. URL: https://www.sciencedirect.com/science/article/pii/S2405896320303281.
8. Ziegler, Julius, Bender, Philipp, Dang, Thao, Stiller, Christoph, 2014. Trajectory planning for bertha — A local, continuous method. In: 2014 IEEE Intelligent Vehicles Symposium Proceedings. pp. 450–457. http://dx.doi.org/10.1109/IVS.2014.6856581.
9. Aleksey Logunov, Muhammad Alhaddad, Konstantin Mironov, Konstantin Yakovlev, Aleksandr Panov, Polygon decomposition for obstacle representation in motion planning with Model Predictive Control, Engineering Applications of Artificial Intelligence, Volume 153, 2025, 110690, ISSN 0952-1976, https://doi.org/10.1016/j.engappai.2025.110690.
10. Adamkiewicz, Michal, Chen, Timothy, Caccavale, Adam, Gardner, Rachel, Culbertson, Preston, Bohg, Jeannette, Schwager, Mac, 2022. Vision-only robot navigation in a neural radiance world. IEEE Robot. Autom. Lett. 7 (2), 4606–4613. http://dx.doi.org/10.1109/LRA.2022.3150497.
11. Kurenkov, Mikhail, Potapov, Andrei, Savinykh, Alena, Yudin, Evgeny, Kruzhkov, Evgeny, Karpyshev, Pavel, Tsetserukou, Dzmitry, 2022. NFOMP: Neural field for optimal motion planner of differential drive robots with nonholonomic constraints. IEEE Robot. Autom. Lett. 7 (4), 10991–10998. http://dx.doi.org/10.1109/LRA.2022.3196886.
12. Alhaddad, Muhammad, Mironov, Konstantin, Staroverov, Aleksey, Panov, Aleksandr, 2024. Neural potential field for obstacle-aware local motion planning. In: 2024 IEEE International Conference on Robotics and Automation. ICRA, pp. 9313–9320. http://dx.doi.org/10.1109/ICRA57147.2024.10611635.
13. Jacquet, M. and K. Alexis, "N-MPC for Deep Neural Network-Based Collision Avoidance exploiting Depth Images," 2024 IEEE International Conference on Robotics and Automation (ICRA), Yokohama, Japan, 2024, pp. 13536-13542, doi: 10.1109/ICRA57147.2024.10610572.

Многоагентные технологии в задачах управления аварийно-спасательными подразделениями в Арктике

Ризванов Д. А.
Уфимский университет науки
и технологий
каф. ВМиК
Уфа, Россия
e-mail: ridmi@mail.ru

Панкратова М. В.
Главное управление МЧС России по
Белгородской области
Старый Оскол, Россия
e-mail: mvpankratova@mail.ru

Аннотация[1]

Арктический регион с его уникальными климатическими условиями и удалённостью представляет собой особую сложность для проведения аварийно-спасательных операций. Уникальные климатические условия Арктики, их стремительное изменение, удаленность региона, ограниченные ресурсы, слабая транспортная инфраструктура, увеличение с каждым годом количества добываемых полезных ископаемых повышает риск возникновения чрезвычайных ситуаций. Ликвидация последствий таких ситуаций в условиях Арктики требует незамедлительного реагирования и четких действий аварийно-спасательных подразделений, а также больших материальных затрат. В связи с этим возникает необходимость в разработке эффективных систем управления, способных оперативно реагировать на любые происшествия. Многоагентный подход представляет собой одну из перспективных стратегий, позволяющих оптимизировать процессы управления аварийно-спасательными подразделениями в арктических условиях. Данный подход в управлении подразделениями основан на взаимодействии и координации нескольких автономных агентов, которые могут представлять собой как людей, так и автоматизированные системы. Он особенно актуален в сложных и динамичных условиях, таких как управление аварийными-спасательными подразделениями, где требуется быстрая реакция и адаптация к изменяющимся обстоятельствам в сложных климатических условиях. В данной статье описаны преимущества и возможности применения многоагентного подхода в области управления аварийно-спасательными подразделениями в арктической зоне.

Ключевые слова: многоагентные системы; арктические условия; аварийно-спасательные подразделения; чрезвычайные ситуации; оптимизация ресурсов; оперативный мониторинг.

1. Введение

Арктическая зона Российской Федерации представляет собой стратегически важный регион, характеризующийся интенсивным развитием промышленной инфраструктуры, ростом объёмов добычи полезных ископаемых и расширением транспортных коридоров. Вместе с тем, суровые климатические условия, экстремальная удалённость населённых пунктов и объектов инфраструктуры, слаборазвитая транспортная сеть и высокая динамика изменений окружающей среды создают уникальные вызовы для обеспечения безопасности и оперативного реагирования на чрезвычайные ситуации (ЧС). Как показывает практика, ликвидация последствий аварий в Арктике сопряжена со значительными временными задержками, повышенными материальными затратами и рисками для жизни и здоровья как пострадавших, так и персонала аварийно-спасательных формирований [1, 2].

Современные ЧС в арктических условиях отличаются высокой степенью неопределённости, неполнотой информации на ранних этапах и быстрой изменчивостью обстановки. Это требует от систем управления способности к адаптации, децентрализованному принятию решений и эффективному распределению ограниченных ресурсов в реальном времени. Традиционные централизованные подходы к координации спасательных операций зачастую оказываются недостаточно гибкими и устойчивыми в условиях нарушенной связи, ограниченной видимости и экстремальных температур.

В этих обстоятельствах особую актуальность приобретают интеллектуальные методы

[1] Труды XI Международной научной конференции "Информационные технологии интеллектуальной поддержки принятия решений", 11-13 ноября, Уфа-Чандигарх-Фандербейлпарк, 2025

управления, основанные на парадигме многоагентных систем (МАС). Многоагентный подход позволяет моделировать взаимодействие автономных сущностей — как технических (сенсоры, беспилотные платформы, мобильные роботы), так и человеческих (спасатели, операторы, командиры подразделений) — в рамках единой кооперативной среды. Такие системы способны к самоорганизации, распределённому анализу данных, адаптивному планированию и совместному принятию решений даже при частичной потере связи или изменении тактической обстановки.

Целью настоящей статьи является анализ потенциала применения многоагентного подхода для повышения эффективности управления аварийно-спасательными подразделениями в арктической зоне. Рассматриваются ключевые направления его использования — от координации действий и оптимизации логистики до мониторинга обстановки, оценки рисков и симуляционной подготовки персонала. Особое внимание уделяется специфике арктической среды как фактора, определяющего требования к надёжности, автономности и устойчивости спасательных операций.

2. Специфика управления аварийно-спасательными подразделениями в арктической зоне

Эффективность реагирования на чрезвычайные ситуации (ЧС) в Арктике определяется не только технической оснащённостью спасательных формирований, но и особенностями самой операционной среды. Арктический регион характеризуется экстремальными климатическими условиями, низкой плотностью населения, фрагментарной транспортной инфраструктурой и значительной удалённостью от центральных логистических узлов. В таких условиях даже незначительная задержка в принятии решений или несогласованность действий может привести к катастрофическим последствиям — как для жизни и здоровья людей, так и для хрупкой арктической экосистемы.

Чрезвычайные ситуации в данном контексте обладают высокой степенью стохастичности: на начальном этапе информация о масштабах аварии, числе пострадавших, характере повреждений оборудования или уровне экологического загрязнения, как правило, неполна и противоречива. Более того, сама обстановка продолжает динамически изменяться под влиянием погодных факторов (например, метелей, резких перепадов температуры, подвижек льда), что делает невозможным долгосрочное планирование. Одновременно с этим доступность ресурсов — от специализированной техники до квалифицированного персонала — ограничена и неравномерно распределена по территории. В совокупности эти обстоятельства формируют крайне сложную задачу оперативного управления, требующую не просто быстрого, но и интеллектуально адаптивного подхода.

На промышленных объектах Арктики, особенно в нефтегазовом секторе, аварийно-спасательные подразделения (АСП) обычно создаются в составе собственных служб безопасности [3-5]. Однако, как свидетельствует статистика, их ресурсов зачастую недостаточно для локализации крупномасштабных инцидентов, таких как разливы нефти или аварии на морских платформах [6]. В подобных случаях приходится привлекать внешние силы — региональные подразделения МЧС России, мобильные группы из соседних субъектов Федерации или даже международные спасательные формирования. Однако процесс координации между разнородными структурами в условиях слабой связи, отсутствия единых протоколов взаимодействия и экстремальной логистики значительно замедляет реагирование и увеличивает совокупные издержки на ликвидацию последствий [7].

3. Классификация и функциональная роль ресурсов аварийно-спасательных подразделений

Для целей моделирования и оптимизации управления целесообразно рассматривать ресурсы АСП как совокупность взаимосвязанных компонентов, каждый из которых может быть представлен в виде автономного агента в рамках многоагентной системы [8].

Для построения эффективной системы управления в условиях чрезвычайных ситуаций необходимо формализовать и классифицировать все компоненты, задействованные в операции. В рамках многоагентного подхода каждая категория ресурсов может быть представлена как автономный агент или группа агентов, обладающих определёнными функциями, ограничениями и возможностями взаимодействия. На рисунке 1 представлено предложенное деление ресурсов АСП на четыре основные категории, отражающие их физическую природу, назначение и способ участия в процессе реагирования.



Рис. 1. Ресурсы аварийно-спасательных подразделений

1. Трудовые ресурсы (личный состав).

Включают в себя не только спасателей и технический персонал, но и командиров, медиков, инженеров, специалистов по связи и логистике. Каждый из них обладает уникальным набором навыков, опытом и физическими возможностями, что требует дифференцированного подхода к распределению задач. В многоагентной модели личный состав представляется как совокупность агентов, способных принимать решения на основе локальной информации, координировать действия с соседними агентами и адаптироваться к изменяющимся условиям — например, переключаться с поисковых операций на эвакуацию или медицинскую помощь в зависимости от ситуации.

2. Материально-технические ресурсы (сырьё и материалы).

Сюда относятся расходуемые материалы: топливо, смазочные вещества, химические реагенты для ликвидации разливов, средства пожаротушения, строительные материалы для временного укрепления конструкций и т.д. Эти ресурсы характеризуются ограниченностью запасов и необходимостью точного учёта в реальном времени. В МАС они могут быть представлены как «ресурсные агенты», которые отслеживают уровень запасов, прогнозируют потребность и инициируют запросы на пополнение, обеспечивая бесперебойность операций даже в условиях длительной изоляции.

3. Производственные ресурсы (средства труда).

Эта категория охватывает всю материальную базу, используемую для выполнения работ: транспортные средства (в том числе специализированные — снегоходы, гусеничные машины, БПЛА), инструменты, оборудование для связи, освещения, бурения, подъёма грузов, а также мобильные пункты управления и штабы. Ключевой особенностью этих ресурсов является их зависимость от внешних условий — например, проходимость техники в условиях глубокого снега или ледовой обстановки. В многоагентной системе каждый объект техники может выступать как агент, способный сообщать о своём состоянии, выбирать маршрут движения, согласовывать использование с другими агентами и даже «запрашивать» помощь при выходе из строя.

4. Финансовые ресурсы.

Хотя финансовые средства не являются физическими объектами, их распределение и контроль играют решающую роль в планировании и реализации спасательных операций. Это включает бюджеты на закупку оборудования, оплату труда, транспортировку, а также резервные фонды на случай непредвиденных расходов. В контексте МАС финансовые ресурсы можно моделировать как «экономические агенты», которые оценивают стоимость различных сценариев действий, оптимизируют затраты на выполнение задач и предоставляют рекомендации по наиболее экономически эффективному использованию средств.

Такая детальная классификация позволяет не только систематизировать управление ресурсами, но и создать основу для построения сложных моделей принятия решений, где каждый тип ресурса участвует в общей операции как активный элемент, способный к самоорганизации, взаимодействию и адаптации. Именно такой подход становится ключевым для повышения устойчивости и эффективности аварийно-спасательных операций в экстремальных условиях Арктики.

4. Преимущества многоагентного подхода

Многоагентный подход предлагает принципиально иной взгляд на организацию спасательных операций — от централизованного командования к распределённой, самоорганизующейся системе. Его применение в арктических условиях обусловлено следующими ключевыми преимуществами:

- Гибкость и адаптивность. Многоагентные системы способны оперативно перестраивать стратегии взаимодействия в ответ на изменения внешней среды. Например, при резком ухудшении видимости или обледенении маршрутов агенты могут автоматически пересчитать приоритеты задач и перераспределить нагрузку между доступными ресурсами.

- Децентрализованное управление. Отсутствие единой точки отказа повышает устойчивость всей системы. Даже при частичной потере связи или выходе из строя отдельных узлов (например, беспилотного аппарата или мобильной станции) остальные агенты продолжают функционировать, поддерживая общую координацию.

- Параллелизм и скорость реакции. Благодаря распределённой архитектуре множество агентов могут одновременно выполнять разнородные задачи — от поиска пострадавших до мониторинга состояния льда или анализа концентрации вредных веществ. Это сокращает общее время реагирования и повышает шансы на успешную ликвидацию ЧС на ранней стадии.

- Эффективная оптимизация ресурсов. МАС позволяют динамически распределять имеющиеся ресурсы с учётом их текущей загрузки, географического положения и специфики задач. Это особенно важно в условиях дефицита техники и персонала, характерного для арктических территорий.

Таким образом, многоагентный подход не просто дополняет существующие методы управления аварийно-спасательными подразделениями — он формирует основу для создания устойчивых,

интеллектуальных и автономных систем реагирования, способных функционировать в условиях высокой неопределённости и экстремальной внешней среды.

5. Применение многоагентных систем в управлении аварийно-спасательными подразделениями: потенциал и перспективы для арктического региона

Многоагентные системы представляют собой перспективный инструментарий для повышения эффективности управления в условиях чрезвычайных ситуаций, особенно в сложных и удалённых территориях, таких как Арктика. Их применение в деятельности аварийно-спасательных подразделений может быть реализовано по нескольким ключевым направлениям.

1. Координация действий в реальном времени. Одним из центральных преимуществ МАС является способность обеспечивать децентрализованную, но согласованную координацию между различными участниками операции. В рамках такой архитектуры автономные агенты — в роли сенсорных узлов, беспилотных платформ или мобильных устройств — способны непрерывно собирать, обрабатывать и передавать информацию о текущей обстановке на месте происшествия. Это включает данные о географических и климатических условиях, локализации пострадавших, а также наличии физических или экологических препятствий. Благодаря взаимному обмену информацией между агентами достигается синхронизация действий спасательных команд, что особенно критично в условиях ограниченной видимости, низких температур и слаборазвитой инфраструктуры, характерных для арктической зоны.

2. Адаптивное принятие решений. Многоагентные системы обладают встроенной способностью к динамической адаптации стратегий реагирования в ответ на изменяющиеся внешние условия. Например, при возникновении новых угроз — таких как лавиноопасность, разлив нефтепродуктов или резкое ухудшение погоды — агенты могут пересчитывать маршруты, перераспределять ресурсы или инициировать эвакуацию на основе заранее заложенных алгоритмов принятия решений. Кроме того, МАС позволяют осуществлять оперативное моделирование альтернативных сценариев развития чрезвычайной ситуации с учётом множества переменных (топография, состояние льда, доступность техники и т.д.). Такой подход обеспечивает выбор оптимальной тактики реагирования, минимизируя временные и человеческие потери.

3. Оптимизация распределения ресурсов в условиях ограниченной инфраструктуры. Эффективное управление ресурсами — одна из ключевых задач при ликвидации чрезвычайных ситуаций в удалённых и труднодоступных регионах, таких как Арктика. Многоагентные системы позволяют реализовать динамическое и самоорганизующееся распределение задач между участниками операции на основе актуальной информации о доступных ресурсах, текущих приоритетах и оперативной обстановке. Например, агенты, представляющие отдельных спасателей или группы, могут автономно согласовывать между собой, какие подразделения сосредоточатся на поиске и эвакуации пострадавших, а какие — на локализации источника угрозы (например, тушении пожара, устранении разлива топлива или стабилизации обрушившейся конструкции). Такой подход минимизирует дублирование усилий и обеспечивает гибкое перераспределение ролей в ответ на изменение тактической ситуации.

Особое значение в арктических условиях приобретает оптимизация логистики и маршрутизации спасательного транспорта. Низкая плотность дорожной сети, сезонная недоступность некоторых участков и экстремальные погодные условия существенно усложняют перемещение техники. Многоагентные алгоритмы, интегрированные с геоинформационными системами и данными о состоянии льда, снежного покрова и видимости, способны в реальном времени рассчитывать наиболее безопасные и эффективные маршруты для доставки персонала, оборудования и гуманитарной помощи. Это не только сокращает время реагирования, но и снижает риски, связанные с эксплуатацией техники в экстремальной среде.

4. Комплексный мониторинг и оценка оперативной обстановки.

Для успешного управления чрезвычайной ситуацией необходима непрерывная и достоверная картина происходящего. Многоагентные системы обеспечивают интеграцию данных от распределённых сенсорных узлов — включая наземные датчики, беспилотные летательные и наземные аппараты, а также стационарные и мобильные видеокамеры. Агенты, отвечающие за обработку этих потоков, способны в режиме реального времени фильтровать, интерпретировать и агрегировать информацию, выделяя критически важные события: перемещение пострадавших, изменение концентрации токсичных веществ, динамику ледовых условий и др. Полученные выводы оперативно передаются командованию и полевым группам, формируя единое информационное поле.

На основе этих данных система также выполняет функцию оценки рисков. С использованием предварительно обученных моделей или правил, основанных на экспертных знаниях, агенты могут прогнозировать потенциальные угрозы — как для пострадавших, так и для самих спасателей. Например, при обнаружении признаков нестабильности ледового покрова или ухудшения метеоусловий система может рекомендовать

изменение тактики действий, эвакуацию персонала из опасной зоны или усиление средств индивидуальной защиты. Такой проактивный подход к управлению рисками значительно повышает уровень безопасности операций и способствует принятию более обоснованных решений в условиях высокой неопределённости.

5. Обучение и симуляционная подготовка.

Ещё одним важным аспектом применения МАС является их роль в обучении и подготовке персонала. С помощью многоагентных симуляторов можно воссоздавать реалистичные сценарии чрезвычайных ситуаций, в том числе специфичные для арктической среды. Такие тренировочные среды позволяют спасателям отрабатывать навыки взаимодействия, оценки рисков и принятия решений в условиях высокой неопределённости. Параллельно, анализ архивных данных предыдущих операций может быть использован для обучения агентов методами машинного обучения и обучения с подкреплением, что способствует накоплению «опыта» системы и повышению её автономности в будущих миссиях.

6. Заключение

Таким образом, внедрение многоагентного подхода в управление аварийно-спасательными подразделениями в арктической зоне представляет собой стратегически значимое инновационное решение. Оно не только улучшает координацию между различными службами и оптимизирует распределение ограниченных ресурсов, но и повышает общую устойчивость спасательных операций в экстремальных условиях. Учитывая растущую экономическую и геополитическую значимость Арктики, развитие таких технологий становится не просто целесообразным, а необходимым условием обеспечения безопасности и устойчивого развития региона.

Список используемых источников

1. Dickins D. Behavior of oil spills in ice and implications for arctic spill response / D. Dickins // Society of Petroleum Engineers - Arctic Technology Conference 2011, Houston, TX. – 2011. – P. 779-793.
2. Soussi A. Environmental Impacts of Oil Spills and Response Technologies / A. Soussi, C. Bersani, R. Sacile // Advances in Environmental Research. – 2021. – P. 139-179.
3. Бажуков, И. Ф. Оснащение арктических комплексных аварийно-спасательных центров МЧС России техническими средствами. Проблемы и решения / И. Ф. Бажуков // Авиационные спасательные технологии в обеспечении комплексной системы безопасности в Арктическом регионе: Материалы VI Форума МЧС России и общественных организаций «Общество за безопасность», / Под общей редакцией Ю.Л. Воробьева. – Вытегра: Санкт-Петербургский университет Государственной противопожарной службы Министерства Российской Федерации по делам гражданской обороны, чрезвычайным ситуациям и ликвидации последствий стихийных бедствий. – 2017. – С. 134-141.
4. Лабардин А. М. Арктические комплексные аварийно-спасательные центры МЧС России – основа комплексной системы обеспечения безопасности в Арктической зоне Российской Федерации / А. М. Лабардин // Арктика – территория безопасности. Развитие системы обеспечения комплексной безопасности Арктической зоны Российской Федерации: Материалы Всероссийской научно-практической конференции / Под общей редакцией Ю.Л. Воробьева, Э.Н. Чижикова, А.И. Бондара. – Вытегра: Санкт-Петербургский университет Государственной противопожарной службы Министерства Российской Федерации по делам гражданской обороны, чрезвычайным ситуациям и ликвидации последствий стихийных бедствий. – 2016. – С. 50-52.
5. Матвеев А.В. Стратегическое планирование сил и средств МЧС России в Арктической зоне // Национальная безопасность и стратегическое планирование. – 2017. – № 4(20) – С. 32-42.
6. Онов В. А., Панкратова М. В., Остудин Н. В. Информационные аспекты в системе антикризисного управления МЧС России // Научно-аналитический журнал "Вестник Санкт-Петербургского университета Государственной противопожарной службы МЧС России". – 2021. – № 2. – С. 116-124.
7. Панкратова М. В., Скрыпникова О. И. Ликвидация разливов нефти в Арктической зоне Российской Федерации // Пожарная и аварийная безопасность : Сборник материалов XVII Международной научно-практической конференции, посвященной 90- й годовщине образования гражданской обороны, Иваново, 24 ноября 2022 года. – Иваново: ФГБОУ ВО «Ивановская пожарно-спасательная академия Государственной противопожарной службы Министерства Российской Федерации по делам гражданской обороны, чрезвычайным ситуациям и ликвидации последствий стихийных бедствий», 2022. – С. 779-782.
8. Ризванов Д. А. Применение многоагентных технологий для решения задачи распределения ресурсов в условиях чрезвычайных ситуаций // Вестник Уфимского государственного авиационного технического университета. – 2012. – Т. 16, № 6(51). – С. 220-225.

Разработка веб-сервиса для семантического анализа образовательных стандартов на основе RAG-архитектуры

Валиев С. М.
Уфимский университет науки и технологий, кафедра вычислительной математики и кибернетики
Уфа, Россия
e-mail: salavat.valieff@yandex.ru

Гаянова М. М.
Уфимский университет науки и технологий, кафедра вычислительной математики и кибернетики
Уфа, Россия
e-mail: maya.gayanova@gmail.com

Вульфин А. М.
Уфимский университет науки и технологий, кафедра вычислительной техники и защиты информации
Уфа, Россия
e-mail: alex.vulfin@gmail.com

Аннотация[1]

Рост объема нормативно-правовых документов в сфере образования требует создания автоматизированных систем для их анализа. Целью исследования является разработка веб-сервиса, предназначенного для сотрудников образовательных учреждений (методистов, преподавателей, административного персонала) и обеспечивающего семантический поиск и генерацию кратких ответов на естественном языке на профессиональные запросы, связанные с содержанием образовательных стандартов и нормативной документации, такие, как: разъяснение терминов, извлечение конкретных требований, сравнительный анализ положений различных документов и подготовка сводных справок, что позволяет существенно сократить временные затраты на работу с нормативной базой и повысить точность получаемой информации. В работе применяется RAG-архитектура (Retrieval-Augmented Generation) программной системы веб-сервиса, сочетающая векторный поиск с использованием нейросетевой модели multilingual-e5-large и библиотеки FAISS с генерацией ответов на основе большой языковой модели Qwen 72B. Разработанный прототип сервиса прошел тестирование, показавшее точность поиска релевантных документов на уровне 97,3% и время ответа менее 3 секунд для 90% запросов. Экспертная оценка подтвердила высокую достоверность генерируемых ответов. Практическая значимость работы заключается в создании инструмента, способного существенно сократить временные затраты на анализ образовательной документации и минимизировать риски несоответствия нормативным требованиям. Научная новизна работы заключается в адаптации RAG-архитектуры для задач анализа образовательных стандартов. Архитектура учитывает специфику нормативных документов сферы образования, включая наличие стандартизированных формулировок, юридически значимых определений и структурных элементов.

Ключевые слова: образовательные стандарты; семантический поиск; RAG-архитектура; большие языковые модели; веб-сервис; искусственный интеллект.

1. Введение

Современная образовательная практика, регламентированная постоянно обновляющимися и усложняющимися стандартами, сталкивается с необходимостью анализа и обработки больших объемов нормативных данных. Это создает потребность в разработке автоматизированных систем, способных извлекать, структурировать и анализировать информацию из документов с использованием методов искусственного интеллекта.

В данной статье разработан веб-сервис, предназначенный для нейросетевого анализа данных образовательных стандартов. Сервис позволяет загружать документы в различных форматах, включая PDF, DOCX, CSV, XLSX и TXT, обрабатывать их с помощью языковых моделей и

[1] Труды XI Международной научной конференции "Информационные технологии интеллектуальной поддержки принятия решений", 11-13 ноября, Уфа-Чандигарх-Фандербейлпарк, 2025

предоставлять ответы на пользовательские запросы на основе семантического поиска.

Разрабатываемая система представляет собой веб-сервис для автоматизированного анализа образовательных стандартов с использованием технологий искусственного интеллекта. Система предназначена для быстрого и точного анализа образовательных документов, обеспечивающего автоматическое извлечение и структурирование информации из нормативных документов, поиск релевантных данных по смысловым запросам и генерацию аналитических ответов на естественном языке. Ключевыми задачами системы являются прием и обработка документов различных форматов, семантический поиск информации в образовательных стандартах и генерация содержательных ответов на пользовательские запросы.

На вход система принимает образовательные документы, включая ФГОС, учебные программы и методические материалы. На выходе система предоставляет структурированные данные из документов, релевантные фрагменты текста по запросам, сгенерированные ответы на естественном языке.

Актуальность работы обусловлена стремительным ростом объема образовательных документов, требующих автоматизированной обработки, а также необходимостью быстрого доступа к релевантной информации из нормативных актов, учебных программ и методических материалов. Кроме того, развитие технологий искусственного интеллекта открывает новые возможности для эффективного анализа текстовых данных, что делает подобные системы востребованными в образовательной сфере.

В ходе работы решены ключевые задачи, включая разработку модуля обработки документов с функцией конвертации, очистки текста и семантической сегментации на логически связанные фрагменты (чанки), реализацию векторного поиска на основе FAISS для эффективного хранения и поиска векторных представлений текста, создание API на базе FastAPI для взаимодействия с сервисом и интеграцию языковой модели для генерации ответов на естественном языке.

1.1. Выбор архитектуры

Для реализации веб-сервиса, отвечающего поставленным задачам семантического поиска и генерации ответов на основе нормативных документов, была выбрана RAG-архитектура (Retrieval-Augmented Generation). Данный выбор обусловлен необходимостью совместить высокую точность извлечения фактов со способностью формулировать связные ответы на естественном языке, избегая при этом характерных для больших языковых моделей недостатков, таких как "галлюцинации" и опора на устаревшие или общие знания.

Критическими требованиями к системе являлись достоверность, актуальность, интерпретируемость и эффективность. Все утверждения в генерируемых ответах должны быть строго основаны на предоставленных документах. Система должна работать с постоянно обновляемой нормативной базой без необходимости дорогостоящего переобучения модели. Пользователь должен иметь возможность проверить источник сгенерированной информации, что особенно важно в юридически значимой сфере образования. Кроме того, использование ресурсоемких языковых моделей должно быть ограничено финальной стадией генерации ответа на основе релевантного контекста.

RAG-архитектура соответствует этим требованиям, разделяя процесс на две четкие фазы: поиска и генерации. Для обоснования выбора рассмотрим ключевые альтернативы и их применимость в контексте задачи.

Альтернативный подход чистого векторного поиска без генерации обеспечивает высокую скорость и полную контролируемость результата, однако его недостаток заключается в том, что пользователь получает необработанные выдержки из документов, требующие ручного анализа. Такой подход не позволяет решить задачи сравнительного анализа, подготовки сводных справок или развернутого разъяснения на естественном языке.

Другой альтернативой является использование чистой генеративной модели, дообученной на корпусе образовательных стандартов. Хотя этот подход может давать качественные и связные тексты, он обладает существенными недостатками. Проблема актуальности проявляется в необходимости дорогостоящего переобучения при появлении новых документов. Риск "галлюцинаций" сохраняется, так как модель может генерировать правдоподобные, но не соответствующие реальности ответы. Кроме того, пользователю сложно проверить источники утверждений модели.

Третий рассматриваемый подход – Extractive QA, системы с извлечением ответа — демонстрирует высокое быстродействие и минимальный риск "галлюцинаций", но не способен синтезировать информацию из нескольких источников или выполнять сравнительный анализ. Этот подход эффективен только для простых фактологических запросов, где ответ представляет собой готовую фразу из документа.

Выбранная RAG-архитектура демонстрирует ряд ключевых преимуществ в контексте поставленных задач. Контролируемость и достоверность обеспечиваются за счет привязки к конкретным фрагментам документов на этапе поиска, что позволяет цитировать источники и минимизировать "галлюцинации". Гибкость и актуальность

достигаются благодаря возможности быстрого обновления векторного индекса без переобучения генеративной модели. Качество и естественность ответов обеспечиваются использованием крупной языковой модели, способной выполнять сложные задачи по синтезу и обобщению информации. Эффективность системы достигается за счет обработки моделью только релевантного контекста, а не всей базы документов.

Таким образом, RAG-архитектура была выбрана как вариант, лишенный недостатков непосредственного поиска по ключевым словам и прямой генерации.

1.2. Технологический стек

Основой для реализации системы анализа образовательных стандартов выбран современный технологический стек, оптимально соответствующий требованиям к производительности, масштабируемости и удобству разработки. Ядро системы строится на платформе Node.js, которая идеально подходит для создания высоконагруженных веб-приложений с интенсивным вводом-выводом. Асинхронная событийно-ориентированная архитектура Node.js позволяет эффективно обрабатывать множественные параллельные запросы к документам, что критически важно для работы с образовательными стандартами.

Серверная часть реализуется на основе фреймворка FastAPI, [7] сочетающего высокую производительность с простотой разработки RESTful API. Выбор обусловлен возможностью быстрого прототипирования сложных аналитических конечных точек, а также встроенной поддержкой асинхронных операций, необходимых для интеграции с нейросетевыми моделями. Автоматическая генерация документации OpenAPI упрощает взаимодействие между фронтендом и бэкендом в процессе разработки.

Для хранения и обработки векторных представлений документов выбрана библиотека FAISS [2], разработанная Facebook Research. Ее преимущество заключается в оптимизированных алгоритмах поиска ближайших соседей в многомерных пространствах, что ускоряет семантический поиск в больших коллекциях образовательных стандартов.

В качестве языковой модели для обработки текста используется multilingual-e5-large. [3] Эта модель демонстрирует высокое качество векторных представлений для образовательной терминологии и поддерживает работу с многоязычными документами. Интеграция модели в систему осуществляется через трансформерную библиотеку Hugging Face. [5]

Фронтенд-часть построена на React [8] с использованием Next.js для серверного рендеринга. Этот выбор обусловлен богатыми возможностями создания интерактивных интерфейсов для работы с документами. Библиотека D3.js применяется для визуализации сложных связей между образовательными стандартами.

Хранилище данных реализуется на основе комбинации PostgreSQL и Elasticsearch. PostgreSQL обеспечивает надежное хранение структурированных данных с поддержкой сложных запросов. Elasticsearch добавляет мощные возможности лингвистического анализа и ранжирования результатов. Для кеширования используется Redis.

Оркестрация нейросетевых вычислений осуществляется через Celery и RabbitMQ. Это решение позволяет эффективно распределять ресурсоемкие задачи анализа документов между несколькими рабочими узлами. Мониторинг системы реализован на базе Prometheus и Grafana.

Выбор именно этого технологического стека обусловлен оптимальным сочетанием производительности, удобства разработки и соответствия специфическим требованиям работы с образовательными стандартами. Все компоненты имеют открытую лицензию, активно поддерживаются сообществом и хорошо документированы, что снижает риски долгосрочной поддержки системы. Гибкость выбранных технологий позволяет адаптировать систему к изменяющимся требованиям и масштабировать ее по мере роста объема обрабатываемых документов.

1.3. Логическая схема

Схема сущностей (Class Diagram) представлена на рисунке 8. Ключевые сущности:

1. Document (Документ):
- Хранит метаинформацию о загруженных файлах
- Связь 1-ко-многим с Chunk
2. Chunk (Текстовый фрагмент):
- Содержит обработанный текст и векторное представление (embedding)
- Метаданные о позиции в исходном документе
3. FAISS_Index [2] (Векторный индекс):
- Ссылается на документы, которые в него включены
- Хранит параметры модели эмбеддингов
4. ChatSession (Сессия чата):
- Привязка к конкретному индексу
- История сообщений
5. Message (Сообщение):
- Сохраняет контекст и ответы для аналитики



Рис. 1. Логическая схема

1.4. Блок-схема работы алгоритма анализа данных

Фаза 1: Индексация документов (Подготовка базы знаний)

Этот процесс выполняется один раз при добавлении новых документов в систему.

1. Начало: запускается процесс подготовки данных.
2. Загрузка документа: система принимает документ от пользователя.
3. Проверка формата и размера:

- Вопрос: «соответствует ли документ поддерживаемым форматам (например, .pdf, .docx, .txt) и допустимому размеру?»
- Если ответ «нет»: вывод ошибки; пользователь уведомляется о проблеме, и процесс для этого документа останавливается.
- Если ответ «да»: переход к следующему шагу.

4. Конвертация в текст: из документа извлекается весь текстовый контент. Например, из PDF извлекаются текст, названия глав, подписи к изображениям и т.д.
5. Предобработка текста: извлеченный текст очищается и нормализуется.
6. Семантическое разделение: длинный текст разбивается на логические, небольшие отрезки. Это не просто разбиение по количеству символов, а умное разделение по смыслу (например, по абзацам, главам или с использованием алгоритмов, учитывающих контекст). Это нужно для того, чтобы позже находить наиболее релевантные фрагменты.
7. Генерация векторов: Каждый текстовый фрагмент преобразуется в числовой вектор, который представляет смысловое содержание этого фрагмента в многомерном пространстве. Семантически похожие тексты будут иметь близкие векторы.
8. Сохранение в FAISS-индекс: все полученные векторы сохраняются в специальную базу данных — FAISS-индекс. FAISS — это высокооптимизированная библиотека для быстрого поиска похожих векторов.
9. Проверка на новые документы:

- Вопрос: «есть ли еще документы для обработки?»
- Если ответ «да»: процесс возвращается к шагу "Загрузка документа".
- Если ответ «нет»: система переходит ко второй фазе.

Фаза 2: Обработка пользовательского запроса и генерация ответа

Этот процесс запускается каждый раз, когда пользователь задает вопрос.

1. Ожидание запроса: система переходит в режим ожидания входящего вопроса.
2. Получение пользовательского запроса: пользователь вводит свой вопрос (например, "Каковы основные условия договора?").
3. Векторизация запроса: вопрос пользователя так же преобразуется в вектор с помощью той же модели, что использовалась на шаге "Генерация векторов". Это позволяет сравнивать вопрос с фрагментами документов.
4. Поиск в FAISS-индексе: вектор запроса используется для поиска в FAISS-индексе. Система находит N (например, 3-5) наиболее близких по смыслу фрагментов из документов.
5. Ранжирование результатов: найденные фрагменты дополнительно ранжируются, чтобы выбрать самые релевантные. Это помогает улучшить качество финального ответа.
6. Формирование контекста: отобранные фрагменты объединяются в один блок текста,

который служит контекстом для генерации ответа.

7. Генерация ответа моделью Okchen-7B/72B: сформированный промпт (вопрос пользователя и контекст) передается на вход крупной языковой модели, в данном случае Okchen-7B или Okchen-72B. Модель генерирует связный, естественный ответ, опираясь на предоставленный контекст.
8. Проверка точности: ответ проверяется на соответствие исходным документам, после чего выводится.
9. Вывод ответа: пользователь получает окончательный, сгенерированный моделью ответ.
10. Конец: процесс завершается. Система возвращается в состояние "Ожидание запроса" для обработки нового вопроса.

Блок-схема работы алгоритма анализа данных представлена на рисунке 2.

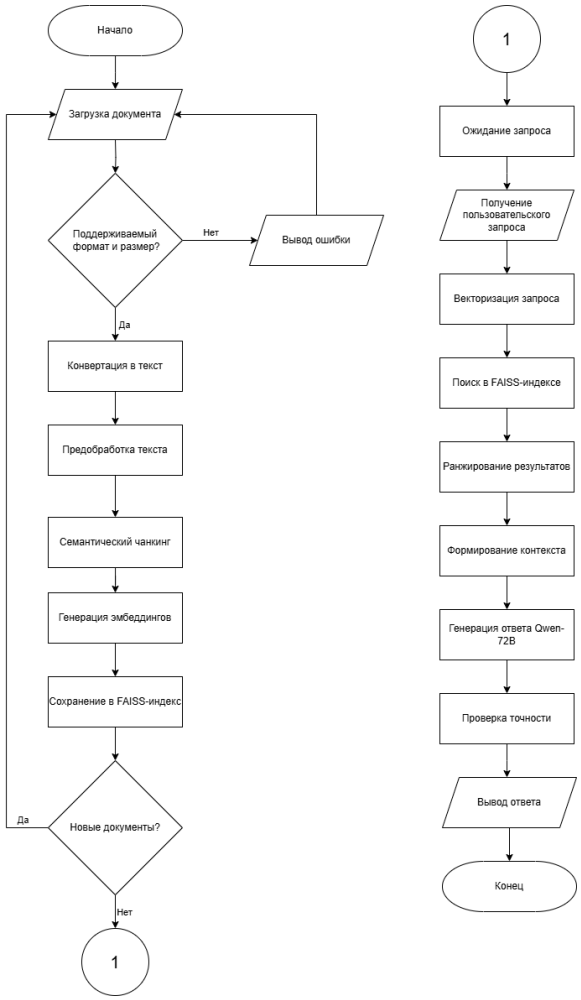


Рис. 2. Блок-схема работы алгоритма анализа данных

1.5. Функциональная модель системы анализа данных

Цель системы: автоматизировать сложный анализ нормативных документов в сфере образования, предоставляя экспертам структурированные ответы на запросы и выявляя скрытые связи между различными документами.

Декомпозиция процесса по функциональным блокам:

1. БЛОК УПРАВЛЕНИЯ

1) Методологии и регламенты анализа задают бизнес-логику, определяют, что такое "компетенция", "требование", "связь", правила валидации ответов. Контролируют предобработку документов, классификацию сущностей, финальную сборку ответа.

2) Алгоритм RAG (Retrieval-Augmented Generation) связывает все компоненты в единый рабочий контур. Обеспечивает, что ответы языковой модели основаны на релевантных фрагментах из базы документов, а не на "воображении" модели.

2. МЕХАНИЗМЫ

1) Серверное оборудование - вычислительная мощность для работы ресурсоемких моделей (Qwen 72B) и векторного индекса.

2) СУБД PostgreSQL хранит структурированные данные: метаинформацию о документах (название, тип, дата).

3) Векторный индекс FAISS обеспечивает быстрый семантический поиск. Хранит векторные представления фрагментов текстов документов. Позволяет находить тексты, близкие по смыслу к запросу, даже если не совпадают ключевые слова.

4) Языковая модель Qwen 72B отвечает за понимание естественного языка, генерацию человекоподобных ответов, извлечение информации и установление связей по заданным регламентам.

3. ВХОДЫ

1) Нормативные документы: ФГОС, профессиональные стандарты, учебные планы, рабочие программы дисциплин (в форматах PDF, DOCX, TXT).

2) Пользовательские запросы: вопросы на естественном языке, например: "Какие компетенции формирует дисциплина 'Математика'?", "Найдите все пересечения между стандартом IT-специалиста и этой образовательной программой".

4. ВЫХОДЫ

1) Структурированные ответы – ответы, приведенные к заданному шаблону, содержащие

извлеченные факты, аргументы из документов и цитаты.

2) Связи между документами: отображение иерархии файлов.

Функциональная модель системы анализа данных представлена на рисунке 3. Декомпозиция функциональной модели представлена на рисунке 4.

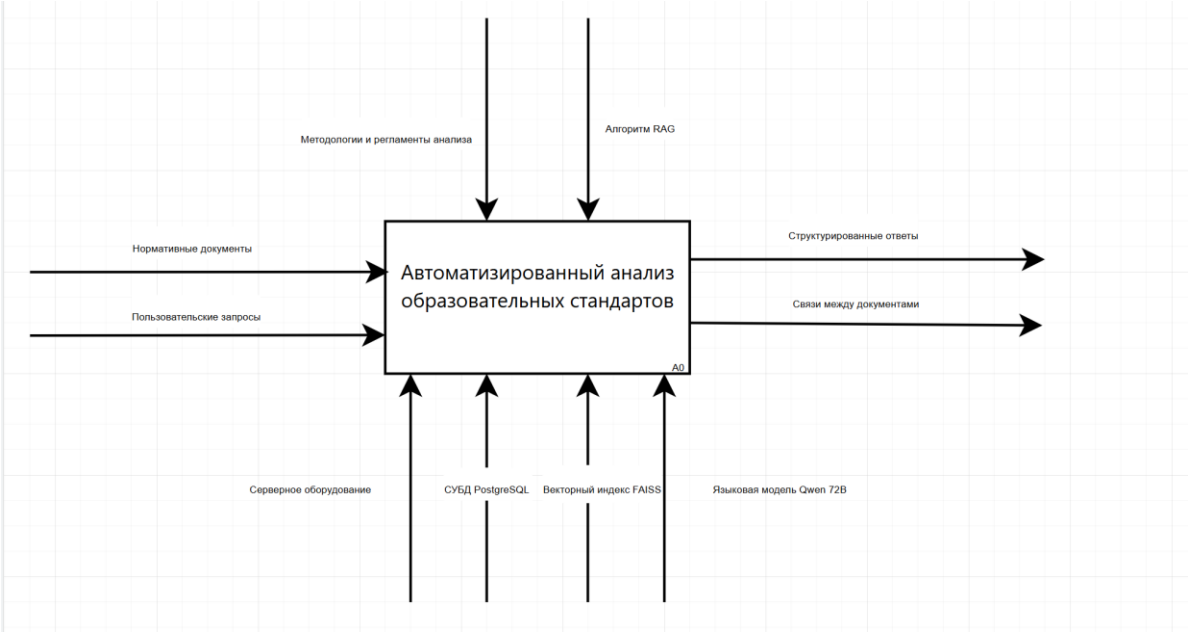


Рис. 3. Функциональная модель системы анализа данных

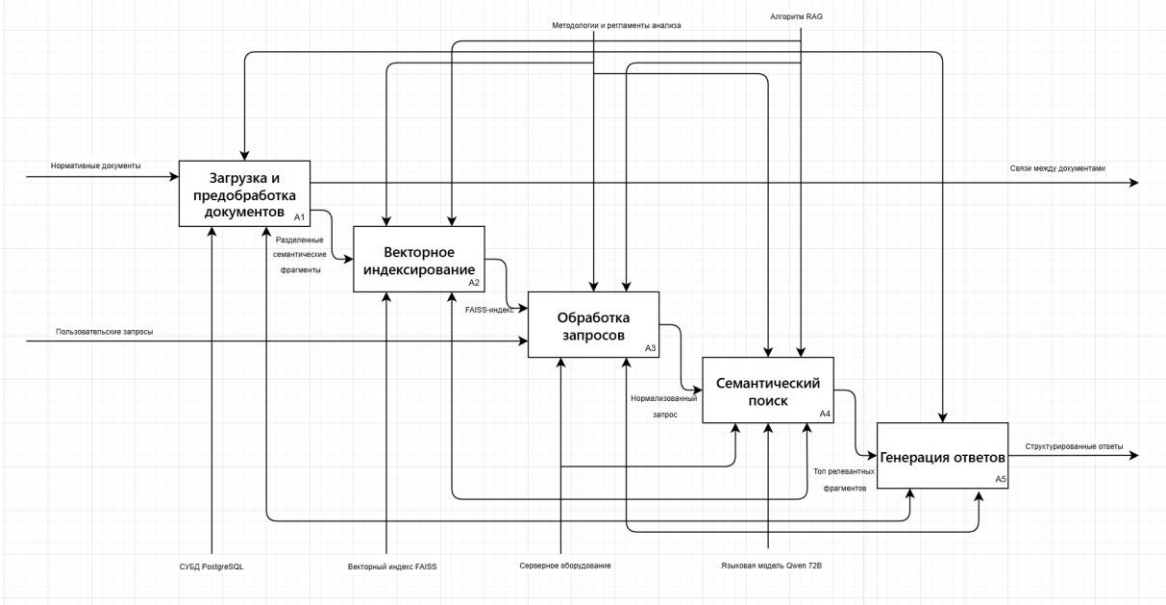


Рис. 4. Декомпозиция функциональной модели системы анализа данных

2. Результаты и обсуждение

При оценке качества разрабатываемой системы анализа образовательных стандартов необходимо подобрать вопросы, на которые мы знаем ответы и проверить правильность ответов нейросети. Правильность ответов будем выражать в процентах.

1. Что такое ФГОС-ВО?

Правильный ответ: Федеральный государственный образовательный стандарт высшего образования.

Ответ нейросети: ФГОС ВО (Федеральный государственный образовательный стандарт высшего образования) представляет собой совокупность обязательных требований, которые должны соблюдаться при реализации основных профессиональных образовательных программ высшего образования. Эти программы включают бакалавриат, специалитет и магистратуру по направлениям подготовки, отнесенным к укрупненной группе специальностей и направлений подготовки. ФГОС ВО обеспечивает единые стандарты качества образования и регламентирует основные параметры образовательных программ.

Правильность ответа: 100%

2. Области и сферы профессиональной деятельности выпускников по направлению «Информатика и вычислительная техника»?

Правильный ответ: Области профессиональной деятельности и сферы профессиональной деятельности, в которых выпускники, освоившие программу бакалавриата, могут осуществлять профессиональную деятельность, указаны во ФГОС.

Направленность программы бакалавриата конкретизирует содержание программы бакалавриата в рамках направления подготовки (специальности) путем ориентации ее на следующие области профессиональной деятельности и сферы профессиональной деятельности выпускников:

1. СВЯЗЬ, ИНФОРМАЦИОННЫЕ И КОММУНИКАЦИОННЫЕ ТЕХНОЛОГИИ;

2. СКВОЗНЫЕ ВИДЫ ПРОФЕССИОНАЛЬНОЙ ДЕЯТЕЛЬНОСТИ В ПРОМЫШЛЕННОСТИ.

Ответ нейросети: Выпускники, освоившие программу бакалавриата по направлению «Информатика и вычислительная техника», могут осуществлять профессиональную деятельность в области связи, информационных и коммуникационных технологий (в сфере проектирования, разработки, внедрения и эксплуатации средств вычислительной техники и информационных систем, управления их жизненным циклом). Также они могут работать в сквозных видах профессиональной деятельности в промышленности (в сфере организации и проведения научно-исследовательских и опытно-конструкторских работ в области информатики и вычислительной техники). Кроме того, выпускники могут осуществлять профессиональную деятельность в других областях и сферах при условии соответствия их компетенций требованиям к квалификации работника.

Правильность ответа: 96%

3. Сколько зачетных единиц составляет программа бакалавриата по направлению ПРО?

Правильный ответ: Объем программы бакалавриата составляет 240 зачетных единиц (далее - з.е.) вне зависимости от формы обучения, применяемых образовательных технологий, реализации программы бакалавриата с использованием сетевой формы, реализации программы бакалавриата по индивидуальному учебному плану.

Программа бакалавриата по любому направлению подготовки, включая направление ПРО, составляет 240 зачетных единиц. Этот объем установлен вне зависимости от формы обучения и применяемых образовательных технологий. Информация об этом содержится в данном контексте.

Правильность ответа: 98%

4.Типы учебной практики по направлению ПИ?

Правильный ответ: Типы учебной практики: ознакомительная практика; технологическая (проектно-технологическая) практика; эксплуатационная практика; научно-исследовательская работа (получение первичны навыков научно-исследовательской работы). Типы производственной практики:

технологическая (проектно-технологическая) практика; эксплуатационная практика; научно-исследовательская работа.

Ответ нейросети: Типы учебной практики, указанные в данном контексте, включают ознакомительную практику, технологическую (проектно-технологическую) практику, эксплуатационную практику и научно-исследовательскую работу.

Правильность ответа: 90%

5. Требования к структуре программы бакалавриата по направлению ИСТ?

Правильный ответ: Структура программы бакалавриата включает следующие блоки: Блок 1 "Дисциплины (модули)";

Блок 2 "Практика";

Блок 3 "Государственная итоговая аттестация".

Ответ нейросети: Структура программы бакалавриата, включая направление ИСТ, состоит из трех основных блоков: Блок 1 "Дисциплины (модули)", Блок 2 "Практика" и Блок 3

"Государственная итоговая аттестация". BProgramma также должна включать дисциплины по философии, истории, иностранному языку и безопасности жизнедеятельности.

Правильность ответа: 92%

6. Какие должна устанавливать общепрофессиональные компетенции программа бакалавриата по направлению ИВТ?

Правильный ответ: Программа бакалавриата должна устанавливать следующие общепрофессиональные компетенции:

ОПК-1. Способен применять естественнонаучные и общеинженерные знания, методы математического анализа и моделирования, теоретического и экспериментального исследования в профессиональной деятельности; ОПК-2. Способен понимать принципы работы современных информационных технологий и программных средств, в том числе отечественного производства, и использовать их при решении задач профессиональной деятельности; ОПК-3. Способен решать стандартные задачи профессиональной деятельности на основе информационной и библиографической культуры с применением информационно-коммуникационных технологий и с учетом основных требований информационной безопасности; ОПК-4. Способен участвовать в разработке стандартов, норм и правил, а также технической документации, связанной с профессиональной деятельностью; ОПК-5. Способен инсталлировать программное и аппаратное обеспечение для информационных и автоматизированных систем; ОПК-6. Способен разрабатывать бизнес-планы и технические задания на оснащение отделов, лабораторий, офисов компьютерным и сетевым оборудованием

Ответ нейросети: Программа бакалавриата по направлению "Информатика и вычислительная техника" устанавливает 9 ключевых общепрофессиональных компетенций (ОПК-1 — ОПК-9), включая применение естественнонаучных знаний, работу с современными ИТ-технологиями, решение профессиональных задач с учетом информационной безопасности, разработку технической документации, инсталляцию ПО, настройку аппаратных комплексов и создание алгоритмов. Все компетенции полностью соответствуют требованиям ФГОС, с незначительными стилистическими вариациями в формулировках при сохранении исходного смысла.

Правильность ответа: 95%

7. Что представляет собой программа специалитета по специализации Автоматизированные системы обработки информации и управления?

Правильный ответ: Программа специалитета представляет собой комплекс основных характеристик образования (объем, содержание, планируемые результаты) и организационно-педагогических условий, который представлен в виде учебного плана, календарного учебного графика, рабочих программ дисциплин (модулей), рабочих программ практик, программы государственной итоговой аттестации, оценочных и методических материалов, форм аттестации.

Ответ нейросети: Программа специалитета по специализации «Автоматизированные системы обработки информации и управления» является проектной. Она ориентирована на подготовку специалистов, которые будут заниматься проектированием, моделированием, экспериментальным исследованием и управлением автоматизированными системами специального назначения, а также выполнять научно-исследовательские и опытно-конструкторские работы в данной области.

Правильность ответа: 92%

8. Как называется программа высшего образования под номером 09.03.04?

Правильный ответ: Программная инженерия

Ответ нейросети: Программа высшего образования под номером 09.03.04 называется «Программная инженерия». Эти сведения указаны в тексте, где описывается программа бакалавриата по данному направлению подготовки.

Правильность ответа: 100%

9. Перечень профессиональных стандартов, соответствующих профессиональной деятельности выпускников, освоивших программу бакалавриата (специалитета/магистратуры) по направлению «Информационные системы и технологии»?

Правильный ответ: Профессиональные стандарты, соответствующие профессиональной деятельности выпускников, освоивших программу бакалавриата по направлению подготовки 09.03.02 Информационные системы и технологии, приведены в приложении к ФГОС.

Профессиональные стандарты, соответствующие профессиональной деятельности выпускников, на основе которых сформированы профессиональные компетенции, установленные программой бакалавриата:

- ПС 06.001 ПРОГРАММИСТ
- ПС 06.016 РУКОВОДИТЕЛЬ ПРОЕКТОВ В ОБЛАСТИ ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ;
- ПС 06.022 СИСТЕМНЫЙ АНАЛИТИК
- 06.015 СПЕЦИАЛИСТ ПО ИНФОРМАЦИОННЫМ СИСТЕМАМ

 ПС 06.042 СПЕЦИАЛИСТ ПО БОЛЬШИМ ДАННЫМ.

Ответ нейросети: Объекты профессиональной деятельности выпускников, освоивших программу бакалавриата по направлению 09.03.02 «Информационные системы и технологии», включают прикладные информационные процессы, информационные технологии и программное обеспечение. Профессиональные стандарты, соответствующие данной деятельности, приведены в приложении к ФГОС и утверждены приказом Министерства образования и науки Российской Федерации от 19 сентября 2017 г. N 926. Ключевые профессиональные стандарты включают ПС 06.001 «Программист», ПС 06.015 «Специалист по информационным системам», ПС 06.042 «Специалист по большим данным» и другие, соответствующие области знаний и компетенций выпускников.

Правильность ответа: 92%

10. Что включают в себя требования к условиям реализации программы бакалавриата?

Правильный ответ: Требования к условиям реализации программы бакалавриата включают в себя общесистемные требования, требования к материально-техническому и учебно-методическому обеспечению, требования к кадровым и финансовым условиям реализации программы бакалавриата, а также требования к применяемым механизмам оценки качества образовательной деятельности и подготовки обучающихся по программе бакалавриата.

Ответ нейросети: Требования к условиям реализации программы бакалавриата включают общесистемные требования, требования к материально-техническому и учебно-методическому обеспечению, а также требования к кадровым и финансовым условиям. Кроме того, к этим требованиям относятся механизмы оценки качества образовательной деятельности и подготовки обучающихся по программе бакалавриата. Вопрос о специфике требований для конкретного направления подготовки в данном контексте не раскрывается.

Правильность ответа: 96%

Средний показатель точности ответов: 95,1%

3. Заключение

Разработан веб-сервис нейросетевого анализа данных образовательных стандартов на основе RAG архитектуры. Система может анализировать данные образовательных стандартов и давать понятные и точные ответы на вопросы пользователей. Это система, работающая с данными образовательных стандартов, которая позволяет использовать семантический поиск и отвечать на вопросы пользователей в виде диалога. В ходе работы были достигнуты следующие цели:

1. Осуществлён анализ предметной области, проведен обзор методов автоматизированной обработки информации, в качестве метода автоматизированной обработки информации была выбрана RAG архитектура;

2. Выполнено построение функциональной модели, а также построена блок-схема работы алгоритма анализа данных;

3. Разработан веб-сервис нейросетевого анализа данных образовательных стандартов с помощью выбранного метода и выполнена программная реализация алгоритма;

4. Проведено тестирование и оценена точность ответов программного обеспечения анализа данных. При правильных входных данных программа работает корректно, при неправильных - выдаёт ошибку. Средний показатель точности системы: 95,1%

Список используемых источников

1. Lewis P., Perez E., Piktus A. et al. Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks // Advances in Neural Information Processing Systems. 2020. Vol. 33. P. 9459–9474. (статья из журнала).
2. Johnson J., Douze M., Jégou H. Billion-scale similarity search with GPUs // IEEE Transactions on Big Data. 2019. Vol. 7. № 3. P. 535–547. (статья из журнала).
3. Wang L., Yang N., Huang X. et al. Text Embeddings by Weakly-Supervised Contrastive Pre-training // arXiv preprint arXiv:2212.03533. 2022. 27 p. (электронный ресурс).
4. Yang A., Yang B., Zhang B. et al. Qwen Technical Report // arXiv preprint arXiv:2309.16609. 2023. 61 p. (электронный ресурс).
5. Vaswani A., Shazeer N., Parmar N. et al. Attention Is All You Need // Advances in Neural Information Processing Systems. 2017. Vol. 30. P. 5998–6008. (статья из журнала).
6. Макконнелл С. Совершенный код. – 2-е изд. – Москва: Русская редакция, 2010. – 896 с. (книга).
7. FastAPI: A modern, fast web framework for building APIs with Python 3.7+ [Электронный ресурс]. – URL: https://fastapi.tiangolo.com/ (дата обращения: 15.05.2024). (электронный ресурс).
8. React: The library for web and native user interfaces [Электронный ресурс]. – URL: https://react.dev/ (дата обращения: 15.05.2024). (электронный ресурс).

Разработка виртуального кабинета конструкции самолётов на основе инженерных моделей

Ширинбегзода С. И.
Самарский университет, ИАРКТ
Самара, Россия
e-mail: ShirinbegzodaSI@yandex.ru

1. Аннотация

Актуальность исследования определяется необходимостью внедрения цифровых технологий в инженерное образование. Современные образовательные платформы часто ограничены в функционале и не обеспечивают глубокого взаимодействия студентов с инженерными моделями.

Цель. Настоящая работа направлена на разработку виртуального кабинета конструкции самолётов, объединяющего инженерные модели и обеспечивающего интерактивное обучение. Создание подобного ресурса позволит интегрировать современные цифровые технологии в образовательный процесс, повысить уровень вовлечённости студентов и сформировать у них практические компетенции, необходимые для инженерной деятельности в авиационной отрасли.

Методы. Проанализирована существующая платформа cnit.ssau.ru, выполнен реверс-инжиниринг агрегатов с использованием 3D-сканера Scanform L5, произведено моделирование в Компас-3D, постобработка в Blender. Интерфейс спроектирован в Figma, веб-реализация — на Tilda Publishing, интеграция 3D-моделей — через Sketchfab.

Результаты. Разработана новая версия платформы с современным интерфейсом, интерактивными 3D-моделями агрегатов и встроенным расчётным модулем. Представлена сравнительная таблица характеристик старой и новой версии. Платформа опубликована в сети.

Заключение. Предложенное решение позволяет повысить качество подготовки инженеров-конструкторов, сняв ограничения традиционного формата обучения. Платформа может быть интегрирована в учебный процесс и расширена новыми агрегатами и модулями.

Ключевые слова: виртуальный кабинет; инженерные модели; 3D-моделирование; реверс-инжиниринг; авиационная техника; цифровое обучение.

Keywords: virtual cabinet; engineering models; 3D modeling; reverse engineering; aircraft technology; e-learning

2. Введение

Современное инженерное образование переживает этап цифровой трансформации, обусловленный развитием технологий и изменением требований к подготовке специалистов. В авиационной отрасли особенно важно формирование у студентов практических компетенций, связанных с пониманием конструкции летательных аппаратов. Однако традиционные образовательные ресурсы зачастую не удовлетворяют этим требованиям.

Одним из таких ресурсов является виртуальный кабинет cnit.ssau.ru, разработанный более десяти лет назад. Несмотря на ценность собранных материалов, платформа имеет ряд существенных недостатков: устаревший дизайн, ограниченный функционал, отсутствие интерактивности и адаптации к мобильным устройствам, наличие ошибок в текстах. Кроме того, ресурс привязан к конкретной лаборатории авиационных конструкций и не позволяет удалённого доступа.

В связи с этим возникает необходимость в разработке новой версии виртуального кабинета конструкции самолётов, обеспечивающей доступность и интерактивность учебных материалов, а также интеграцию инженерных моделей.

3. Анализ проблем инженерного образования и потенциал e-learning

Инженерное образование в классическом формате сталкивается с рядом ограничений:

- ограниченный доступ студентов к дорогостоящему оборудованию;
- устаревшие формы подачи материала;
- низкий уровень интерактивности;
- разрыв между теорией и практикой;
- слабая цифровая компетентность части преподавателей и студентов.

Эти проблемы также отмечаются в исследованиях по цифровизации образования [1]. С другой стороны, e-learning и смешанные форматы позволяют существенно повысить качество

подготовки специалистов за счёт доступности, интерактивности и адаптивности обучения [2].

Виртуальные лаборатории и цифровые двойники агрегатов дают студентам возможность:

- изучать конструкцию авиационной техники в интерактивном формате;
- выполнять расчёты в цифровой среде;
- формировать компетенции в области цифрового инжиниринга;
- осваивать навыки работы с CAD/CAE-системами.

Таким образом, разработка виртуального кабинета конструкции самолётов напрямую решает выявленные проблемы и соответствует задачам цифровой трансформации образования, закреплённым в государственной программе «Цифровая экономика Российской Федерации» [3].

4. Материалы и методы

Методологическая основа исследования включает несколько этапов.

1. Анализ платформы cnit.ssau.ru. Выявлены сильные стороны (набор информации, иллюстрации) и недостатки (устаревший интерфейс, ошибки, отсутствие 3D и адаптивности).
2. Реверс-инжиниринг агрегатов. Для получения цифровых двойников использовался 3D-сканер Scanform L5 [8]. Объекты сканирования находились в лаборатории авиационных конструкций.
3. CAD-моделирование. Обработка данных сканирования и создание параметрических моделей осуществлялись в Компас-3D [9].
4. Постобработка моделей. Для оптимизации геометрии и текстурирования применялся Blender [10].
5. Проектирование интерфейса. Макеты пользовательских экранов созданы в Figma.
6. Разработка веб-ресурса. Вёрстка и публикация выполнены с помощью конструктора Tilda Publishing.
7. Интеграция 3D-моделей. Для взаимодействия с моделями в браузере использовалась платформа Sketchfab [4].
8. Разработка расчётного модуля. Добавлен функционал для базовых инженерных вычислений по нормам прочности.

5. Разработка виртуального кабинета

Результатом работы стала обновлённая платформа виртуального кабинета, опубликованная в сети Интернет.

Ключевые особенности новой версии:

- современный интерфейс, соответствующий стандартам UX/UI;
- упрощённая навигация и адаптивность к мобильным устройствам;
- интеграция интерактивных 3D-моделей агрегатов самолётов;
- расчётный модуль для выполнения инженерных вычислений;
- расширенные описания агрегатов и справочные материалы.

Таблица 1 – Сравнение характеристик старой и новой версии виртуального кабинета

Параметр	**Старая версия**	**Новая версия**
Дизайн интерфейса	Устаревший, статический	Современный, адаптивный
Интерактивность	Отсутствует	Поддержка вращения и масштабирования 3D-моделей
Доступ к материалам	Только в лаборатории	Онлайн, без ограничений по месту
Ошибки в текстах	Многочисленные	Исправлены, добавлены новые материалы
Поддержка мобильных устройств	Отсутствует	Реализована
Учебный функционал	Ограниченный	Расчётный модуль, расширенные справочные разделы

Пойдем по порядку.

Анализ платформы cnit.ssau.ru показал, что сайт уже не развивается и остался с тем интерфейсом, с которым его создали. И поэтому я взял всё лучшее оттуда, а именно – информацию и некоторые картинки и схемы.

На картинках приведу примеры как выглядит сайт и его страницы.



Рис.1 . Главная страница старой версии платформы



Рис. 2. Страница «Общие сведения о самолётах»



Рис. 3. Страница «Фюзеляж»



Рис. 4 . Страница «Шасси»



Рис. 5. Страница «Крыло»



Рис. 6. Страница «Оперение»

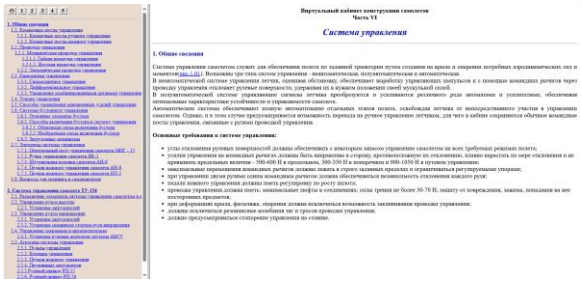


Рис. 7. Страница «Система управления»



Рис. 8. Страница «Оперение» - раздел «Оперение самолёта «Супер-Аэро»» - ошибки в тексте

Современные образовательные ресурсы должны обладать высоким уровнем интерактивности, поддержкой трёхмерных моделей и адаптивностью к мобильным устройствам. Отсутствие этих элементов делает традиционные сайты менее эффективными и не интересным. Поэтому этот проект можно отнести к студентам начальных курсов авиационных конструкций для первичного анализа объекта их исследования – авиационной техники.

Следующая задача — это реверс инжиниринг.

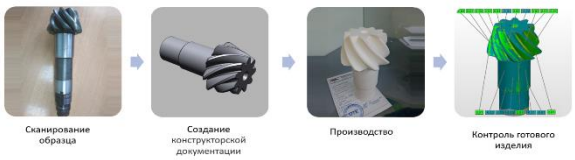


Рис. 9. Схема реверс-инжиниринга

Использование реверс-инжиниринга позволило воспроизвести геометрию сложных агрегатов с высокой точностью. Программные средства, такие как CAD/CAE системы, обеспечили последующую обработку полученных моделей.



Рис. 10. Процесс сканирования агрегата

6. Результаты

Результаты исследования включают:

- создание обновленной версии платформы виртуального кабинета;
- интеграцию современных элементов интерфейса и интерактивных моделей;
- публикацию ресурса в сети Интернет;
- разработку расчетного модуля на основе норм прочности.



Рис. 11. Предварительная главная страница в Figma



Рис. 12. Предварительная страница с одним из разделов

Новый интерфейс ориентирован на студентов, которые привыкли к современным веб-приложениям. Навигация упрощена, а визуальные элементы соответствуют современным стандартам UX/UI.

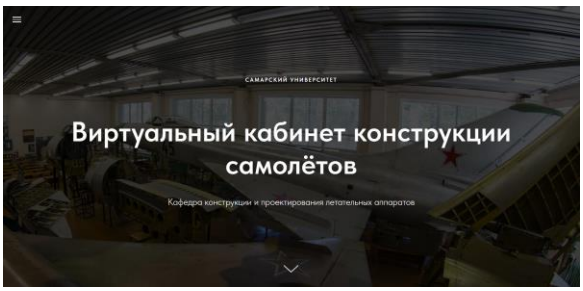


Рис. 13. Главная страница в Tilda

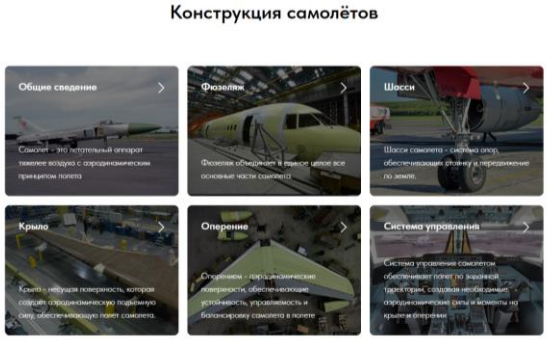


Рис. 14. Схема выбора раздела

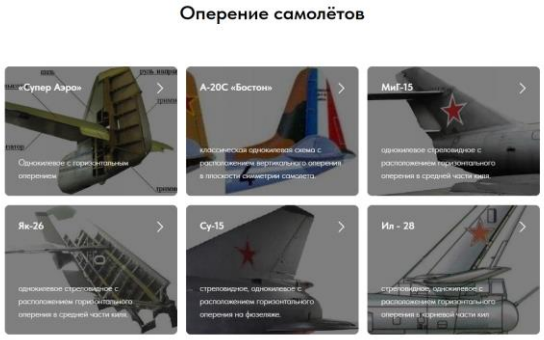


Рис. 15. Схема с примерами оперения



Рис. 16 . Схема с примерами крыла

Важной частью результатов является интеграция 3D-моделей агрегатов самолётов через Sketchfab. Эти модели могут вращаться, масштабироваться и изучаться пользователем в браузере без установки дополнительных программ.

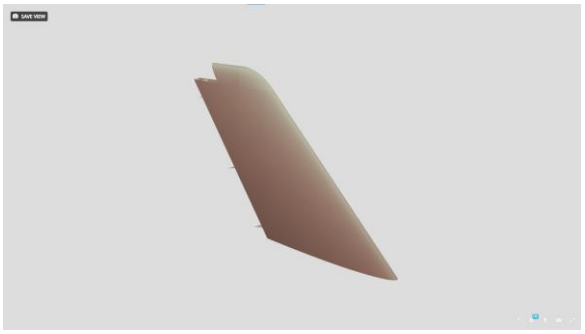


Рис. 17. Пример 3D-модели агрегата



Рис. 18. Окно с описанием агрегата на новой версии платформы



Рис. 19. Окно с описанием агрегата на предыдущей версии платформы

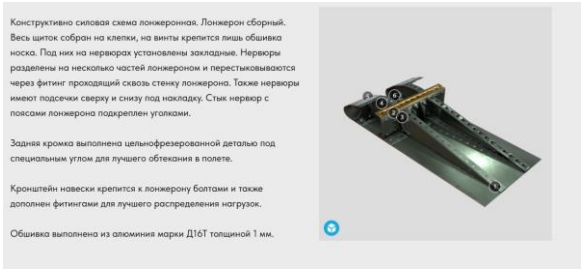


Рис. 2. Окно с описанием агрегата на новой версии платформы



Рис. 21. Окно просмотра 3D модели



Рис. 22. Окно просмотра 3D модели

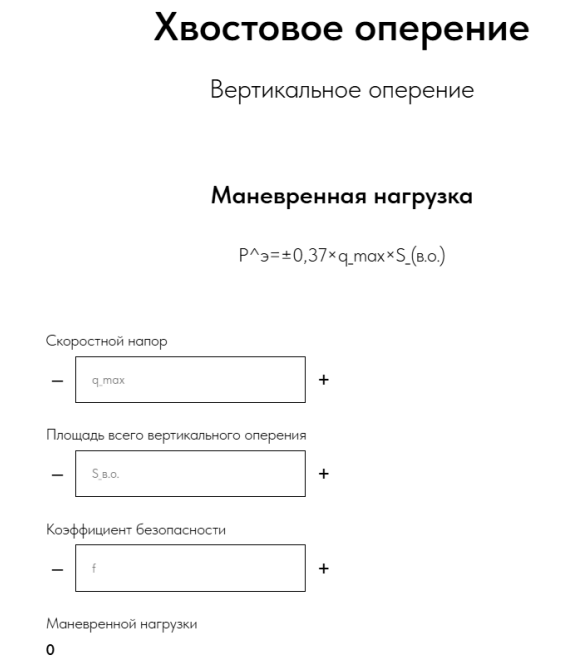


Рис. 23. Окно расчётного модуля

7. Заключение

- Основные выводы исследования:
- Разработана новая версия виртуального кабинета конструкции самолётов.
- Реализована интеграция интерактивных 3D-моделей агрегатов.
- Обеспечена поддержка учебного процесса средствами цифровой платформы.
- Платформа опубликована в сети Интернет и готова к использованию в образовательной практике.
- Результаты могут быть расширены за счёт интеграции новых агрегатов и расчётных модулей.

Список используемых источников

1. Boyko A., Savitskaya T. Improving the Quality of E-learning and Blended Learning in Engineering Education through the Psychometric Analysis. 2024.
2. Цифровая трансформация высшего образования в России. ResearchGate. 2024.
3. Программа «Цифровая экономика Российской Федерации». URL: https://static.government.ru/media/files/9gFM4FHj4PsB79I5v7yLVuPgu4bvR7M0.pdf (дата обращения: 01.05.2025).
4. Уголок неба – Большая авиационная энциклопедия [Электронный ресурс]. – URL: https://airwar.ru/ (дата обращения: 01.05.2025).
5. Виртуальный кабинет конструкции самолётов [Электронный ресурс]. – URL: https://cnit.ssau.ru/index.htm (дата обращения: 01.05.2025).
6. Sketchfab – платформа для публикации и просмотра 3D-моделей [Электронный ресурс]. – URL: https://sketchfab.com (дата обращения: 01.05.2025).
7. Реверсивный инжиниринг [Электронный ресурс]. – URL: https://3dcontrol.ru/services/revers-inzhiniring/reversivnyy-inzhiniring (дата обращения: 01.05.2025).
8. Scanform L5 – 3D-сканер [Электронный ресурс]. – URL: https://scanform.ru (дата обращения: 01.05.2025).
9. Компас-3D – CAD-система [Электронный ресурс]. – URL: https://kompas.ru (дата обращения: 01.05.2025).
10. Blender – графический редактор [Электронный ресурс]. – URL: https://www.blender.org (дата обращения: 01.05.2025).

Структура программы для автоматизации импорта модели

Ямали Д. Д.
Уфимский университет науки
и технологий, каф. ТК
Уфа, Россия
e-mail: den.yamali@yandex.ru

Родионова Л. Е.
Уфимский университет науки
и технологий, каф. АСУ
Уфа, Россия
e-mail: Rodionovakf@yandex.ru

Аннотация[1]

В данной статье представлен разработанный программный модуль для программы Autodesk Maya, автоматизирующий поиск и импорт объектов в сцену путём предоставления функционала поиска и быстрого выбора способа импорта непосредственно через интерфейс Maya. В статье рассматриваются обычный импорт и импорт референсов, а также в чём их отличия: обычный импорт создаёт независимую копию объекта, в то время как референс ссылается на файл объекта, что существенно уменьшает занимаемое место на диске. Рассмотрено понятие «namespace», для чего он нужен и в каких случаях используется: «именное пространство» нужно для устранения конфликтов одинаковых имён у объектов. Перечислены недостатки исходного функционала Maya и функционал предлагаемого решения, упрощающий работу. Показана структура предлагаемого решения.

Ключевые слова: программный модуль; компьютерная графика; референс; импорт; Autodesk Maya, скрипт.

1. Введение

Производство компьютерной графики – зачастую кропотливый и ресурсозатратный процесс. На производство полнометражного мультфильма продолжительностью полтора часа в среднем уходит более трёх лет – это обусловлено временем, затраченным на следующие этапы [1, 2]:

- Рисование концептов объектов и персонажей;
- Создание трёхмерной формы объектов;
- Рисование текстур;
- Настройка материалов для желаемого взаимодействия света с поверхностями;
- Сборка локаций: создание окружения путём расстановки объектов в сцене;
- Анимация, в том числе с технологией захвата движения (motion capture);
- Спецэффекты: симуляция волос, шерсти, воды, огня, взрывов и подобного;
- Расстановка и настройка освещения;
- Рендер: получение изображений в результате компьютерных вычислений;
- Композитинг: постобработка, корректировка после рендера;
- Монтаж: склейка отдельных шотов* в цельный видеоряд.

*Шот – это непрерывно снятая последовательность кадров.

Для проектирования окружения обычно используется множество отдельных объектов [3], однако создание этих объектов внутри самой сцены локации – неэффективно, поскольку это сильно увеличивает объём файла, а также делает проблематичным использование объектов повторно, например, в других сценах или проектах. Чтобы избежать лишней работы, каждый объект, будь то стул или дерево, создаётся в отдельном файле – такой подход позволяет добавлять одни и те же объекты в разные сцены, не тратя время на их повторное моделирование и текстурирование.

Детально проработанная сцена с большим массивом импортированных объектов может занимать от нескольких до сотен гигабайт на диске. С такими тяжёлыми сценами проблематично работать – они медленно открываются, с запозданием реагируют на действия художника, долго сохраняются, а также могут вовсе вызвать зависание программы, что часто означает впустую потраченные часы и повторение уже проделанной работы. Чтобы устранить лишнее дублирование информации, были придуманы «референсы» (reference) – ссылки на файлы с объектами [4].

Ключевые отличия импорта и референсов [5]:

[1] Труды XI Международной научной конференции "Информационные технологии интеллектуальной поддержки принятия решений", 11-13 ноября, Уфа-Чандигарх-Фандербейлпарк, 2025

Импорт создаёт отдельную независимую копию объекта, занимающую столько же места на диске, сколько и оригинал, что увеличивает занимаемое место кратно количеству объектов [6]. Преимущество такого подхода в том, что сцена не зависит от внешних файлов. Если открыть сцену на другом устройстве, она будет иметь тот же вид.

Референсы же ссылаются на файлы с объектами, поэтому они почти не занимают места [7, 8]. В сцене могут находиться тысячи референсов, но её файл при этом будет весить мало. Также преимуществом референсов является то, что при внесении изменений в исходный файл объекта, все сцены, которые на него ссылаются, автоматически обновят этот объект, и его не придётся заменять вручную.

В определённой ситуации подойдёт свой вариант: где-то более уместен обычный импорт, а где-то – референсы. В данной статье эти функции будут рассмотрены на примере программы «Autodesk Maya».

В Maya существует такое понятие как namespace – атрибут, осуществляющий уникальную идентификацию объектов [9]. Программа не позволяет иметь в сцене два объекта с одинаковым «именем», однако не редкость, когда в разных файлах имеются объекты, названные одинаково, например файлы с автомобилями могут иметь колеса с именем «wheel_fl» (от англ. front left – переднее левое колесо). Как раз в таких случаях namespace очень выручает: при его использовании, перед каждым объектом добавляется префикс с именем файла, откуда был загружен объект, что решает проблему уникальных имён. Если продолжать на примере колёс, то после импорта в одну сцену они могли бы иметь имена: «sport_car:wheel_fl» и «truck:wheel_fl», где «sport_car» и «truck» – имена файлов с машинами.

2. Недостатки исходного функционала

В течение всего времени производства нередко приходится обращаться к файлам, созданным несколько месяцев назад – поэтому важно соблюдать порядок в структуре проекта, такой подход позволит быстро ориентироваться в его ресурсах, даже если их сделал другой человек.

Однако встроенный функционал Maya имеет свои ограничения, такие как отсутствие поиска: приходится вручную указывать путь к нужному файлу и выставлять настройки, что делает импорт и добавление референсов времязатратным и изнурительным.

Также встроенный функционал Maya требует указывать в настройках импорта, создавать namespace или нет (рис. 1). Namespace нужен не всегда, поэтому для каждого отдельного случая нужно переходить в настройки и ставить соответствующую «галочку».

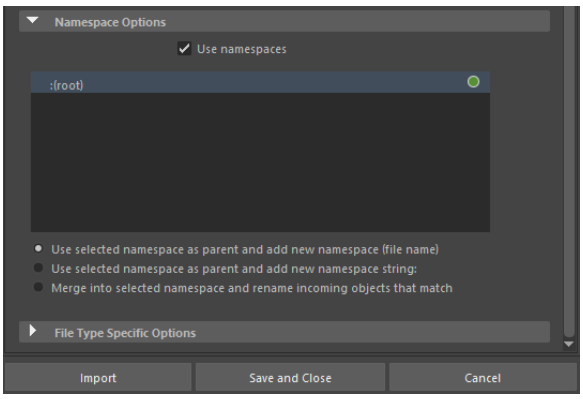


Рис. 1. Настройки встроенного импорта Autodesk Maya

3. Предлагаемое решение

В данной статье рассматривается программный модуль (скрипт) для Maya, который упрощает и ускоряет работу с файлами [10]. Он предоставляет следующий функционал:

- Поиск файлов по имени по всей файловой структуре проекта (рис. 2);
- Предпросмотр содержимого файлов, дополнительная информация (рис. 3);
- Импорт, добавление референса с namespace и без него (рис. 4);
- Возможность открыть директорию файла во встроенном проводнике (рис. 4);

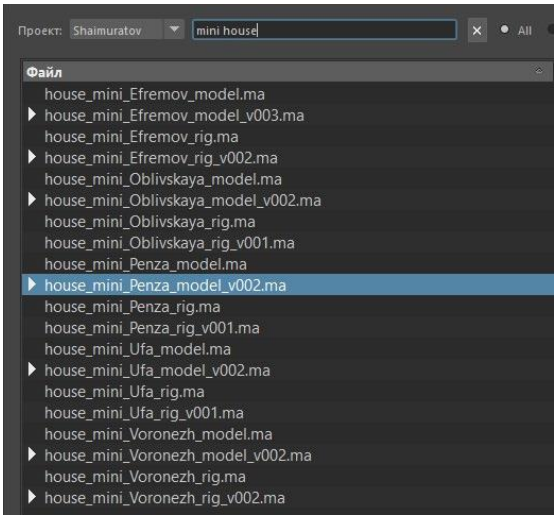


Рис. 2. Поиск по имени через программный модуль

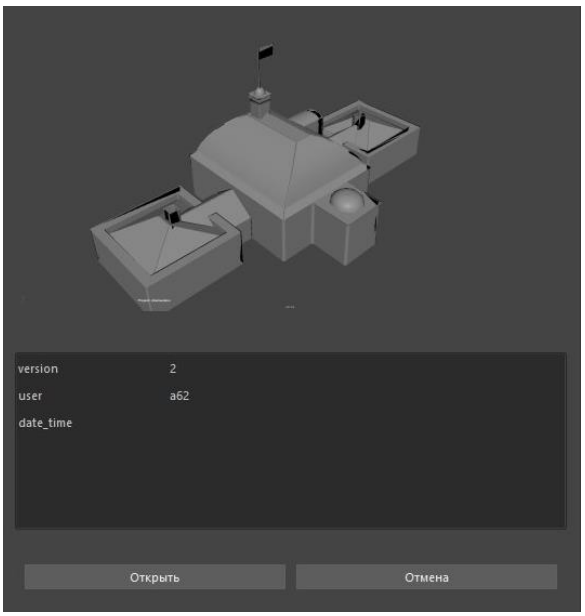


Рис. 3. Предпросмотр в программном модуле

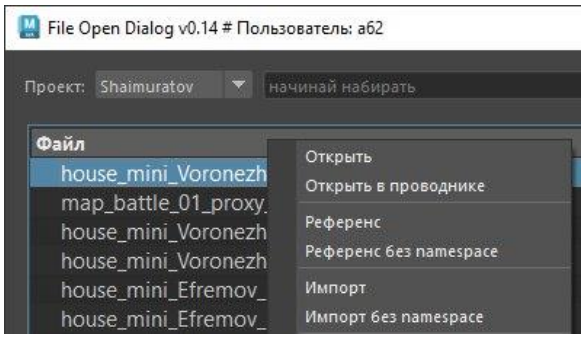


Рис. 4. Контекстное меню в программном модуле

4. Структура предлагаемого модуля

Ниже представлена структура разработанного программного модуля (рис. 5).

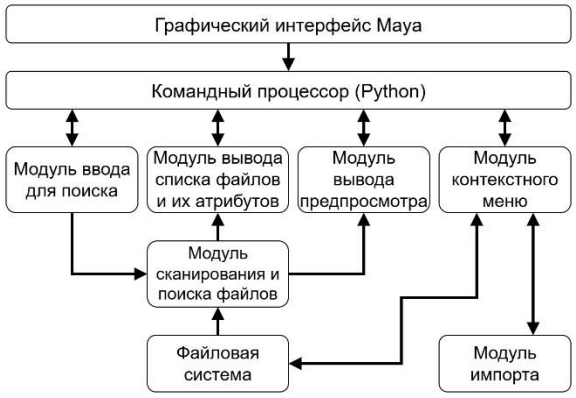


Рис. 5. Структура программы для автоматизации импорта

Данный модуль работает следующим образом:

1. После запуска скрипта открывается графический интерфейс (окно);
2. Считывание строки поиска;
3. Сканирование директории проекта;
4. Отбор и вывод файлов, удовлетворяющих условию поиска;
5. Считывание изображения из директории активного файла и вывод в окно предпросмотра;
6. Обработка контекстного меню;
7. Импорт из файла с соответствующими настройками;
8. Открытие директории в проводнике.

5. Заключение

Предлагаемый программный модуль значительно ускоряет и упрощает работу над проектами, исключая недостатки встроенного функционала Maya. Он позволяет находить файлы по названию, что даёт возможность поиска без знания точного имени и директории, а также упрощает добавление объектов в сцену – будь то импорт или добавление референсов с namespace или без.

Список используемых источников

1. Пайплайн 3D моделирования // skypro URL: https://sky.pro/wiki/digital-art/pajplajn-3d-modelirovaniya/ (дата обращения: 26.09.2025). (электронный ресурс).
2. Что такое CGI // Яндекс Практикум URL: https://practicum.yandex.ru/blog/chto-takoe-cgi-grafika/ (дата обращения: 26.09.2025). (электронный ресурс).
3. Разработка детализированных 3d сцен и ландшафтов // Apptask URL: https://apptask.ru/blog/razrabotka-detalizirovannyx-3d-scen-i-landsaftov (дата обращения: 26.09.2025). (электронный ресурс).
4. Руководство Maya // Autodesk URL: https://help.autodesk.com/view/MAYAUL (дата обращения: 26.09.2025). (электронный ресурс).
5. Maya Referencing // howto URL: https://howto.im/q/how-to-use-mayas-referencing-feature-effectively (дата обращения: 26.09.2025). (электронный ресурс).
6. Maya Import Options // help.autodesk URL: https://help.autodesk.com/view/MAYAUL/2023/ENU/?guid=GUID-37B31221-7BBC-40DA-AB33-47C333EF4E14 (дата обращения: 26.09.2025). (электронный ресурс).
7. Building Complex Maya Scenes // toxigon URL: https://toxigon.com/how-to-build-a-scene-in-maya-through-refereining (дата обращения: 26.09.2025). (электронный ресурс).
8. Getting the most out of File Referencing in Maya // 3dfiggins URL:

https://3dfiggins.com/writeups/mayaReference/ (дата обращения: 26.09.2025). (электронный ресурс).

9. Maya Namespace and Prefix // help.autodesk URL: https://help.autodesk.com/view/MAYAUL/2023/ENU/?guid=GUID-37B31221-7BBC-40DA-AB33-47C333EF4E14 (дата обращения: 26.09.2025). (электронный ресурс).
10. Creating Complex Qt UIs for Maya // dilenshah URL: https://www.dilenshah.com/post/creating-complex-qt-uis-for-maya-and-houdini (дата обращения: 26.09.2025). (электронный ресурс).

Алгоритм для определения маршрута движения наземного робота с учетом препятствий

Саллам М.
каф. «Системное программирование»,
Южно-Уральский государственный университет
Челябинск, Россия
e-mail: mohamedslam2000@yahoo.com

Макаровских Т. А.
Центр образовательных программ топ-уровня в сфере искусственного интеллекта «ВиртУм», Южно-Уральский государственный университет
Челябинск, Россия
e-mail: makarovskikh.t.a@susu.ru

Аннотация[1]

В статье рассматривается алгоритм построения маршрутов наземных роботов для проведения мониторинга в замкнутом пространстве (заводы, больницы, банки, школы, жилые дома и зоны экстренных ситуаций) при наличии неподвижных (стены и двери), полуподвижных (например, мебель и оборудование) и подвижных (живые объекты, люди и другие движущиеся транспортные средства) препятствий. Рассмотренный в статье алгоритм позволяет определить кратчайший маршрут робота, начинающийся в депо (парковке) и проходящий через максимально возможное число точек мониторинга. Подвижные и полуподвижные препятствия учитываются в режиме реального времени, и при их обнаружении на пути следования робота выполняется перерасчет оптимальной траектории и проверка доступности всех непройденных точек для мониторинга. Таким образом, представлен алгоритм для решения задачи прокладки маршрута транспортного средства (Vehicle Routing Problem) для данных ограничений. Разработана программа, с помощью которой создаются карты местоположений и маршрутов, определяются пути между несколькими точками с возвратом к месту парковки для подзарядки аккумуляторов. Программа сохраняет историю путей, успешно пройденных роботом, с учётом типов и проходимости различных препятствий, встреченных на маршруте. Хранение истории позволяет сократить время для поиска оптимальных путей при каждом последующем запуске робота.

Ключевые слова: маршрутизация, симуляция робота, комплекс программ, VRP, принятие решений, мониторинг.

Keywords: routing, routing, robot simulation, software, VRP, decision making, monitoring.

1. Введение

Проблема построения траекторий роботов в настоящее время является актуальной и находится в центре внимания многих научных институтов и компаний, стремящихся разработать оптимальные методы управления движением роботов и интегрировать их в повседневную жизнь вместо того, чтобы полагаться на человеческий контроль и выполнять опасные задачи, когда использование стационарных камер и датчиков может быть не применимо (разбор последствий чрезвычайных ситуаций, мониторинг промышленных садов). Кроме того, стационарные камеры позволяют осуществлять движение только вокруг фиксированных осей, что делает невозможным получение качественных снимков для проведения интеллектуального анализа объекта мониторинга. Даже дорогостоящие камеры зачастую имеют достаточно много слепых зон, что затрудняет получение полной информации об объекте наблюдения.

Известно немало разработок прототипов робота для проведения экспериментов и мониторинга объектов в реальных условиях с учетом препятствий, например, [1-3], а также различных алгоритмов распознавания препятствий [4,5].

Во всех вышеперечисленных случаях исследуемая территория содержит ряд препятствий, которые можно разделить на три типа: неподвижные (деревья, стены, столбы, заборы), полустационарные (мебель, столы, оборудование и

[1] Труды XI Международной научной конференции "Информационные технологии интеллектуальной поддержки принятия решений", 11-13 ноября, Уфа-Чандигарх-Фандербейлпарк, 2025

инструменты) и подвижные (люди, животные и другие транспортные средства, находящиеся в помещении). Устройство наблюдения (робот) должно учитывать эти препятствия при расчёте оптимального метода наблюдения и планировании своего маршрута. Известные методы [6–9], как правило, не учитывают препятствия. Классическая формулировка задачи учитывает количество объектов, которые необходимо посетить. Другой подход к решению той же задачи обсуждался в [10], но в данной статье препятствия не рассматривались.

Можно отметить, что все статьи можно подразделить на три категории. (1) Работы, в которых рассматривается математическая постановка задачи [6-9], приводится доказательная база и рассматриваются теоретические аспекты задачи. Тут подробно рассматриваются алгоритмы решения задачи коммивояжера и задачи прокладки маршрута транспортного средства с различными ограничениями. Однако в этих работах нет тестирования на реальных прототипах. (2) Работы, посвященные развитию методов распознавания изображений и построению архитектур нейронных сетей [2, 3]. (3) Работы, в которых обсуждается создание прототипа. Отдельно хотелось бы отметить работу [3], в которой приведен метод, позволяющий автоматически определить траектории движения робота и элементы управления карданным подвесом и адаптирует планы в режиме реального времени в зависимости от изменений окружающей среды. Наибольший интерес представляют исследования [11, 12], в которых авторы предлагают алгоритмы определения траекторий роботов, которые тестируются на реальных прототипах.

Задача, рассматриваемая в данной статье, аналогична так называемой задаче «дрон-грузовик» [10], в которой также рассматриваются устройство наблюдения (дрон) и устройство «дрон-сервер» (грузовик). Однако эти известные подходы также не учитывают, что робот должен уметь преодолевать препятствия, и что грузовик требуется не всегда. Существует также несколько исследовательских работ, посвященных траекториям с учетом препятствий, например, [13, 14]. Большинство из них посвящены физическим аспектам, таким как поиск точки поворота устройства наблюдения на основе свободных сегментов; однако цели остаются теми же — определение кратчайшего безопасного пути.

Эта тема хорошо изучена, и существует множество задач и эвристик [6, 9, 12–14] для решения задач с различными ограничениями. Например, в [2] представлены эвристики для эвристического поиска в реальном времени для планирования движения с динамическими препятствиями.

В данной работе обсуждается подготовка карты местности с указанием постоянных и временных препятствий для последующего построения кратчайшего пути для робота-наблюдателя и алгоритм SPPA, позволяющий построить кратчайший путь робота, проходящий через все точки мониторинга. Робот управляется программой, разработанной авторами [15], которая хранит базу данных успешно пройденных роботом путей для более быстрого определения его траектории с учетом встречающихся на пути препятствий. Несмотря на то, что существует большое количество алгоритмов маршрутизации с известной информацией об окружающей среде [11-13], в реальных приложениях информация об окружающей среде, получаемая роботом, часто неполна, поскольку могут происходить динамические изменения. Кроме того, рассматриваемые в данных статьях подходы чаще всего основаны на работе известного и хорошо изученного алгоритма A* [16, 17]. В данной работе мы покажем, что разработанный алгоритм SPPA позволяет превзойти A* не только по времени, но и по длине результирующего пути.

2. Алгоритм работы системы мониторинга объектов

Структурная схема системы моделирования траектории перемещения робота для проведения мониторинга приведена на рис.1.



Рис. 1. Технологический цикл работы системы мониторинга

Для задания исследуемой области и расположения на ней препятствий различных типов, а также отметки парковочного места робота и точек для проведения мониторинга, было создано приложение с простой графической средой. Приложение позволяет создать план помещения, в котором определяются границы исследуемого пространства, координаты точек мониторинга и координаты точек парковки, а также задаются препятствия, информация о которых имеется в базе данных.

Данные, получаемые от пользователя в этом приложении удобно для дальнейшей обработки представить в виде графовой модели [18], в которой всю область, используемую для перемещения робота, разделим на небольшие квадратные плитки Rj, j=1,…,M. Длина стороны каждой плитки равна расстоянию, проходимому роботом, между моментами передачи информации оператору *l*.

Робот может перемещаться между углами плитки по горизонтали либо вертикали (тогда длина пути между парой моментов обмена информацией равна l), либо по диагонали (в этом случае робот преодолеет расстояние $\sqrt{2}l$ до следующего момента передачи информации).

Представим исследуемую область как граф $G=(V,E)$, множество вершин V которого являются точками, в которых робот обменивается информацией с прикладной программой, а ребра E – связями между ближайшими вершинами (как отмечалось выше, переходы осуществляются по горизонтали, вертикали либо диагонали).

Подразделим все вершины графа $V = V_p \cup V_m \cup V_d \cup V_t$ на следующие подмножества:

- вершины «парковки» ($p_i \in V_p$) – вершины, с которых робот может подзарядить аккумулятор либо получить сервисное обслуживание, стартовые вершины;
- вершины мониторинга ($m_i \in V_m$) – вершины, в которых проводится мониторинг;
- вершины маршрута объезда ($d_i \in V_d$) вершины, принадлежащие маршруту объезда препятствия;
- транзитные вершины $v \in V_t$ – все остальные вершины графа, которые могут быть включены в маршрут.

Пользуясь информацией для этого графа можно построить производный взвешенный граф $G^* = (V_P \cup V_M, E^*)$, вес каждого ребра которого равен длине кратчайшего пути между соответствующими парами вершин в графе G. Рассмотрим вершины $v, w \in V_P \cup V_M$ в графе G. Если путь из вершины v в вершину w проходит через некоторую вершину $v_k \in V_P \cup V_M$, то вершины v и w не будут смежны в графе G^*. Если для всех пар вершин $v, w \in V_P \cup V_M$ пути, их соединяющие, не проходят ни через одну из вершин $v_k \in V_P \cup V_M$, то граф G^* является полным.

Для решения задачи построения маршрута движения робота – воспользуемся следующим алгоритмом.

Алгоритм SPPA

Входные данные: карта объекта F области с указанными границами, местоположением парковки ($P_0 \in V_p$) и точками мониторинга $m_i \in V_m$, отсортированными по возрастанию расстояния от P_0, последней точкой в этом списке будет $P_0 \in V_p$.

Выходные данные: кратчайший маршрут D, который начинается и заканчивается в точке парковки $P_0 \in V_p$ и проходит через каждую из точек мониторинга $m_i \in V_m$, избегая при этом обнаруженных препятствий.

Шаг 0. Загрузить карту области F и перечень точек для мониторинга V_m.

Шаг 1. Определить внешние рабочие границы F, постоянные известные препятствия s_i, назначить веса точкам, в которых находятся препятствия $\omega(s_i) \in [1;100]$.

Шаг 2. Инициализировать список достижимых вершин L_{open}. Для каждой $v \in L_{open}$ задать вес $w(v) = \omega(v_{ij})$, где v_{ij} - вершина графа, находящаяся в ячейке с координатами (i, j). Этот вес будет использоваться в функции выбора приоритета при определении кратчайшего маршрута.

Шаг 3. Инициализировать список недоступных вершин L_{closed}. Задать начальную точку $v_{curr} = P_0 \in L_{open}$ и результирующий путь $D = P_0$. Пусть текущая целевая точка находится в вершине $m_{curr} \in V_m$, где $m_{curr} \in L_{open}$. Определить границы области закрытого списка L_{closed} таким образом, что $\forall v \in L_{closed} : v \in F$.

Шаг 4. Пока список $L_{open} \neq \varnothing$, выполнить следующие действия.

Шаг 4.1. Найти $v_i \in L_{open}$, ближайшую к v_{curr}.

Шаг 4.2. Извлечь v_i из L_{open}: $L_{open} = L_{open}$, v_i.

Шаг 4.3. Сформировать список последователей $\mathrm{succ}(v_i)$, который включает вершины, посещенные после выхода из v_i.

Шаг 4.4. Для каждого $v \in \mathrm{succ}(v_i)$:

Если $v = P_0$, прекратить поиск, маршрут завершен.

Для каждой $v \in \mathrm{succ}(v_i)$ вычислить

$$h_1 = \mathrm{dist}(P_0, v_i) + \mathrm{dist}(v_i, v);\ h_2 = \mathrm{dist}(m_{curr}, v);$$
$$f(v) = h_1 + h_2.$$

Найти $w \in L_{open}$, что $w = v$. Если $f(w) < f(v)$, пропустить v.

В противном случае добавить v к результирующему маршруту: $D = D \bigcup v$, положить $v_{curr} = v$.

Шаг 4.5. Если достигнута целевая точка $v = m_{curr}$, то пометить эту точку как пройденную, добавив ее в закрытый список: $L_{closed} = L_{closed} \bigcup m_{curr}$. Выбрать следующую цель $m_{curr} \in V_m$ и повторить шаг 4. Если $L_{open} = \varnothing$, но $v_{cur} \neq P_0$, завершить работу алгоритма. Задача не имеет решения.

Шаг 5. D представляет собой кратчайший путь, проходящий через все точки $m_i \in V_m$.

Конец алгоритма SPPA.

Доказательство результативности алгоритма приведено в [19].

Алгоритм SPPA встроен в функционирование общего алгоритма проведения мониторинга [20].

Алгоритм МАРШРУТИЗАЦИЯ

Вход:

- база данных построенных ранее маршрутов,
- технические характеристики робота,
- вектор точек мониторинга M_i,
- координаты точки (точек) парковки P_i.

Выход: Протокол мониторинга объектов.

Шаг 1. Построить карту, определить точку парковки, точки для проведения мониторинга.

Шаг 2. Определить идеальный маршрут P_{ideal} между точкой P_0 (парковка) и всеми точками мониторинга M_i, $i = 1,\ldots,N$, с учетом расхода аккумуляторной батареи и учетом информации об известных препятствиях.

Шаг 3. Двигаться по маршруту P_{ideal} до тех пор, пока это возможно. Осуществлять мониторинг и передачу информации о достигнутых целях в обозначенных точках M_i, $i = 1,\ldots,N$.

Шаг 4. Если робот встретил препятствие, обратиться к базе данных и проверить наличие ранее построенного обходного маршрута *D*. Если маршрут *D* найден, загрузить его из базы данных и перейти на шаг 6. Иначе: решить задачу оптимальной маршрутизации с учетом доступности еще не посещенных точек мониторинга и известных роботу препятствий. Если удается построить маршрут *D*, продолжить движение по нему и перейти на шаг 6.

Шаг 5. Если маршрута *D* не существует, сообщить пользователю перечень объектов *E*, для которых невозможно провести мониторинг. Перейти на шаг 6.

Шаг 6. Если осмотрены ли все объекты, которые не помещены в множество *E*, перейти на шаг 7. Иначе продолжить движение и перейти на шаг 3.

Шаг 7. Конец алгоритма. Вывести протокол проведения мониторинга с указанием осмотренных и недоступных для осмотра точек.

Таким образом, представленный алгоритм управления роботом состоит из этапов, приведенных на рис.2. Алгоритм SPPA работает на втором этапе (Прокладка маршрута).

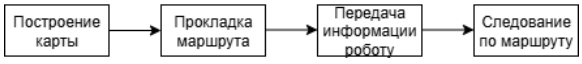


Рис. 2. Основные этапы работы системы мониторинга объектов роботами

Предложенный алгоритм предполагает, что оптимальный путь – это путь, не препятствующий движению робота. В полученном пути учитываются такие внешние воздействия и технические факторы, как возможность подзарядки аккумуляторов робота или замены карт памяти, используемых для хранения шагов пути, в случае недостаточности ресурсов для выполнения полного цикла мониторинга. Движение по заданному пути означает перемещение робота на одну позицию вдоль пути. Затем проверяется возможность продолжения движения (шаги 4 и 5), и мониторинг завершается (шаг 6). Пространство, исследуемое роботом, делится на элементарные сектора (плитки) одинаковой длины и ширины, которые будем называть ячейками. Это обеспечивает равноудалённость размеров во всех направлениях, что позволяет сравнивать количество ячеек независимо от направления движения (продольного, поперечного или диагонального) по разным путям, не учитывая математические относительные различия, которые могут привести даже к незначительной ошибке в расчётах движения. Непрерывное движение определяется как способность достичь ближайшей точки без препятствий или с минимальными препятствиями по местности с перепадами высот, которые затрудняют или даже замедляют движение от центра первого сегмента до следующей точки. Путь определяется на основе ранее приобретенного опыта, независимо от того, является ли новый желаемый путь частью существующего пути или более крупным, путем поиска в базе данных.

Таким образом, решена задача маршрутизации, позволяющая минимизировать число подзарядок и уменьшить общее время мониторинга всех выбранных точек.

Задача имеет существенные ограничения: когда некоторые данные неизвестны, становится невозможным выполнение всех ограничений для всех случайных переменных. Таким образом, может требоваться выполнение некоторых условий с заданной вероятностью, либо построение корректирующей модели, выполняющейся при нарушении каких-либо ограничений.

Например, можно рассмотреть следующие способы коррекции:

(1) вернуться на парковку, когда требуется подзарядка, а затем продолжить путь по определенному ранее маршруту;

(2) вернуться на парковку, когда потребуется подзарядка, и заново оптимизировать оставшуюся часть маршрута;

(3) запланировать досрочный возврат на парковку, даже если аккумулятор не полностью разряжен. В данном случае, решение может зависеть от

количества осмотренных точек, и расстояния до парковки. Робот может вернуться на парковку, если известно, что для осмотра следующей точки мониторинга и возврата на парковку не будет достаточно заряда аккумулятора.

Таким образом, на основе построенной алгоритмической схемы могут быть поставлены следующие задачи.

1. Построение маршрута мониторинга с учетом внешних воздействий.

2. Определение возможности продолжения движения с помощью алгоритмов компьютерного зрения.

3. Поиск обходного маршрута с учетом внешних воздействий и обнаруженных препятствий.

4. Обновление информации в базе данных о препятствиях и обходных путях.

Препятствия на пути робота могут отслеживаться несколькими способами (рис. 3):

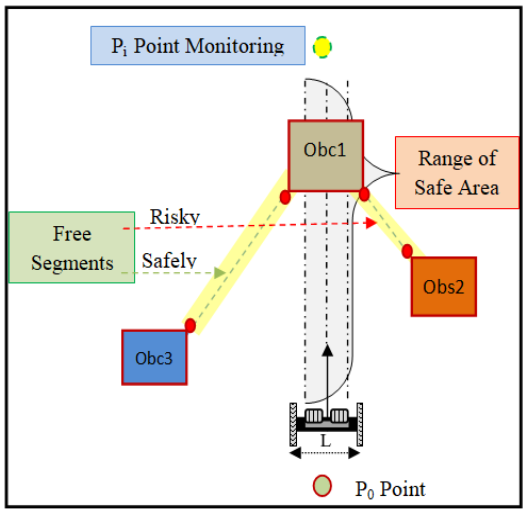


Рис. 3. Использование звуковых волн для обхода препятствий, оценки расстояния, выбора оптимального пути, соответствующего размеру робота, и прохождения пути

- Ультразвуковые волны, которые считаются эффективными и быстрыми, позволяют рассчитывать расстояние с точностью до трёх метров (это значение может быть увеличено в зависимости от типа установленного ультразвукового датчика), учитывая, что роботу требуется всего 5–10 метров для безопасной остановки благодаря инерции скорости робота во время движения. Также можно определить расстояние между препятствием и роботом, а также его размеры, которые определяются тремя датчиками спереди, двумя сзади и одним с каждой стороны. Это позволяет роботу изучать различные типы препятствий, проходя через патрубок трубы, небольшое отверстие или небольшой туннель с максимально возможной точностью.
- Интеллектуальные камеры ESP32, которые можно запрограммировать на отслеживание новых препятствий на пути робота при подключении к Wi-Fi и создании модели машинного обучения для обучения модели классификации объектов и, таким образом, точной классификации и даже распознавания препятствий.

3. Вычислительные эксперименты

Для тестирования алгоритма SPPA проведены эксперименты с разными известными метриками для вычисления $\mathrm{dist}(a,b)$: манхэттенская, евклидова, евклидова NoSQR и максимальное расстояние по Чебышеву.

В проведенном вычислительном эксперименте сравнивается трудоемкость алгоритма SPPA с трудоемкостью одной из современных модификаций алгоритма A* (A-star) [21] для решения одной и той же задачи. Алгоритм A* – это традиционный и хорошо известный подход к планированию маршрута [22]. На первый взгляд наиболее очевидным алгоритмом поиска ближайшего соседа представляется давно известный и хорошо изученный алгоритм Дейкстры. Но этот алгоритм позволяет найти расстояние между фиксированной парой вершин и фиксированными расстояниями. В рассматриваемой задаче расстояние между вершинами может меняться, кроме того, на каждом шаге мы получаем эвристическую информацию, которая может помочь ускорить поиск решения. В таком случае лучшим выбором являются алгоритмы из семейства A* и подобные подходы. В проведенном вычислительном эксперименте использована модифицированная версия алгоритма A* [22], которая допускает диагональные перемещения.

Может показаться целесообразным использование для сравнения волнового алгоритма (известного также как алгоритм Ли). Этот алгоритм привлекает своей простотой в реализации, поскольку использует для поиска решения хорошо известный BFS-поиск, однако он не учитывает веса ребер и неэффективен на больших графах. Кроме того, в статье [23] представлен результат эксперимента, где этот алгоритм превосходит A* на графе с 3600 вершинами. Однако в нашей статье рассматриваются взвешенные графы, с которыми алгоритм Ли не работает.

Также представлено сравнение с алгоритмом муравьиной колонии и KNN. Алгоритм муравьиной колонии имитирует поведение муравьёв, которые коллективно находят кратчайшие пути от муравейника к источнику пищи, он является одной из достаточно быстрых эвристик, позволяющих найти близкие к оптимальным решения. Алгоритм KNN позволяет решить не только задачу поиска оптимального пути, но и задачу кластеризации, когда требуется определить подмножества близко расположенных друг к другу объектов.

Эксперимент посвящен созданию путей между точками, P_0 и $m_i \in V_m$, без учета препятствий (см. таблицу 1). Этот эксперимент позволяет исследовать скорость вычислений и длину пути с использованием различных известных показателей.

Описание условий эксперимента

- Каждый эксперимент с одной и той же картой.
- В каждом эксперименте одни и те же объекты на каждой карте (количество, сортировка, тип и местоположение).
- Одинаковое количество точек для мониторинга на карте.
- Использование четырех наиболее распространенных метрики подсчета расстояния.
- Унифицированное максимальное количество точек поиска.
- Унифицированные размеры ячеек в каждом эксперименте.

Критерии оценки результатов эксперимента:

- Расчет времени в минутах, секундах и миллисекундах.
- Пространство, представляющее собой количество точек поиска оптимального маршрута.
- Отсортировать результаты времени для каждого алгоритма отдельно от наибольшего значения (1) до наименьшего (4) по времени, затраченному каждым алгоритмом на процесс поиска.
- Анализ проходимости препятствий для робота.

Было проведено достаточно много вычислительных экспериментов. Приведем результаты 16 из них, рассматривающих построение маршрута между парковкой и 15 точками для проведения мониторинга (таблица 1).

Таблица 1. Результаты вычислительных экспериментов

№ экс.	Алгоритм	Метрика	Время	Длина пути
1	SPPA	Манх.	1.54.20	180
2		Евкл.	3.05.03	180
3		Евкл. noSQR	3.07.55	180
4		Чебыш.	3.03.16	180
5	A*	Манх.	2.11.75	255
6		Евкл.	3.32.29	255
7		Евкл. noSQR	3.34.70	255
8		Чебыш.	3.02.60	255
9	Мурав. колония	Манх.	3:11.75	230
10		Евкл.	1:48.75	210
11		Евкл. noSQR	1:42.20	210
12		Чебыш.	2:04.75	230
13	KNN	Манх.	3:02.75	200
14		Евкл.	2:54.10	200
15		Евкл. noSQR	2:34.334	198
16		Чебыш.	3:05.02	230

Вычислительный эксперимент показал следующее.

1. Алгоритм KNN уступает остальным как по производительности, так и по длине найденного маршрута. Проложенный алгоритмом путь (рис. 4) следует прямолинейному пути, согласно его теории и методу поиска ближайшей клетки между текущей точкой и прямым направлением к целевой точке. Это исключает возможность диагональных переходов между клетками, что требует больше времени для поиска решения и большего количества точек для достижения оптимального пути к цели.

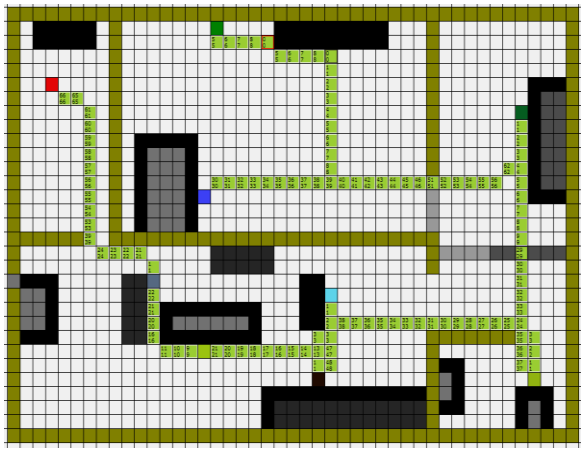


Рис. 4. Построение кратчайшего пути по метрике Euclidean no SQR для алгоритма KNN

2. Алгоритм муравьиной колонии проходит не по диагонали, а скорее напоминает движение по лестнице, чтобы добраться до следующей ближайшей точки, достигая таким образом ближайшей точки к цели (рис.5). Он также позволяет получить допустимое решение за приемлемое время (для ряда метрик даже превосходящее результаты других алгоритмов), но также уступает по длине найденного пути.

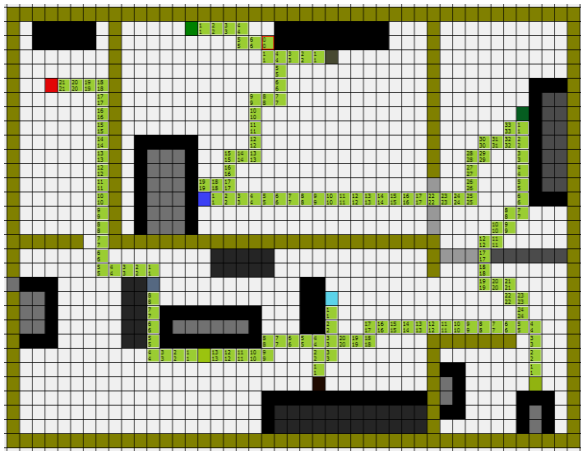


Рис. 5. Построение кратчайшего пути для мониторинга по метрике Euclidean no SQR для алгоритма Ant colony

3. Алгоритм A*, находящийся в приоритете у исследователей и разработчиков подходов определения пути с учетом препятствий, работает в 2 раза медленнее алгоритма муравьиной колонии, а его современные модификации, позволяющие учитывать движение по диагонали, не позволяют получить путь длины, сравнимой с результатами других алгоритмов (рис.6).

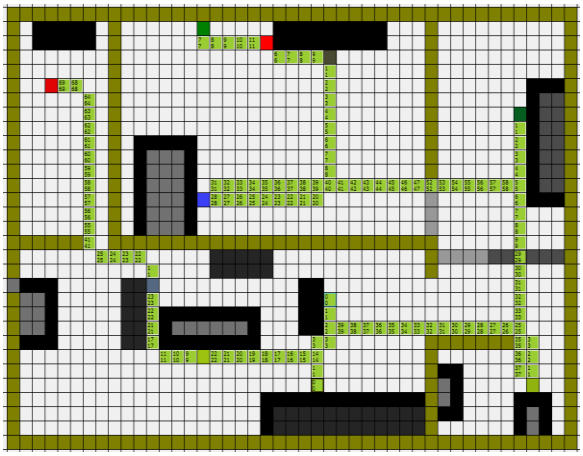


Рис. 6. Построение кратчайшего пути по манхэттенской метрике для алгоритма A*

4. Наконец, показано, что разработанный авторами статьи алгоритм SPPA, для проведенного эксперимента показал наилучшие результаты, к тому же данный метод позволил сократить длину пути за счет учета переходов между точками по диагонали (рис.7). Несмотря на то, что время работы алгоритма для всех метрик (кроме манхэттонской) превышает время работы ряда известных алгоритмов, он позволяет получить более короткий путь, проходящий через все точки интереса.

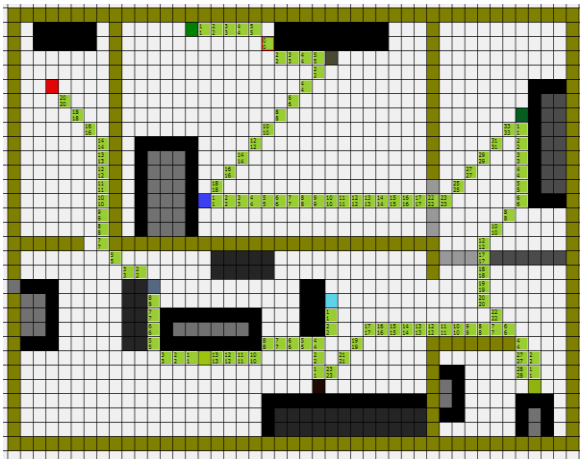


Рис. 7. Построение кратчайшего пути по манхэттенской метрике для алгоритма SPPA

4. Заключение

Таким образом, в статье рассмотрено решение проблемы планирования траектории мониторинга в области с препятствиями. Авторами разработан алгоритм SPPA, который улучшает заданную идеальную траекторию (решение классической задачи TSP), принимая во внимание как стационарные, так и подвижные препятствия. Информация о неподвижных препятствиях хранится в базе данных, и к ней можно мгновенно получить доступ, что позволяет сократить время на поиск наиболее подходящего решения. Для

разработанного алгоритма предполагается, что местоположение подвижных препятствий на карте известно. Однако в дальнейшем предполагается обнаруживать препятствия с помощью системы интеллектуального зрения, размещать их на карте мониторинга и пересчитывать найденный маршрут с учетом вновь поступившей информации. В этом контексте алгоритм должен использовать модуль искусственного интеллекта для распознавания препятствий, определения их контуров и включения их в карту.

Разработанный подход может быть использован для наземного мониторинга различных объектов.

Проведенные вычислительные эксперименты показывают, что алгоритм SPPA работает быстрее и позволяет найти более короткий путь по сравнению с алгоритмами A*, KNN, алгоритмом муравьиной колонии.

Список используемых источников

1. Sakahara, H., Masutani, Y., Miyazaki, F. (2007). Real Time Motion Planning in Dynamic Environment Containing Moving Obstacles Using Spatiotemporal RRT. Transactions of the Society of Instrument and Control Engineers. 43. 277-284. 10.9746/ve.sicetr1965.43.277.
2. CannonJ., RoseK., RumlW. Real-time motion planning with dynamic obstacles // Fifth Annual Symposium on Combinatorial Search, volume 3. 2012. In Vol. 3 No.1 (2012).
3. T. Nägeli, J. Alonso-Mora, A. Domahidi, D. Rus and O. Hilliges, "Real-Time Motion Planning for Aerial Videography With Dynamic Obstacle Avoidance and Viewpoint Optimization," in IEEE Robotics and Automation Letters, vol. 2, no. 3, pp. 1696-1703, July 2017, doi: 10.1109/LRA.2017.2665693
4. Kuric I, Bulej V, Saga M, Pokorny P. Development of simulation software for mobile robot path planning within multilayer map system based on metric and topological maps. International Journal of Advanced Robotic Systems. 2017;14(6).
5. Jichao Cui, Guanghua Nie, "Motion Route Planning and Obstacle Avoidance Method for Mobile Robot Based on Deep Learning", Journal of Electrical and Computer Engineering, vol. 2022, Article ID 5739765, 11 p., 2022.
6. Sundarrajan, M., Jothi, A., Prabakar, D., Kadry, S. (2023). The Smart Coverage Path Planner for Autonomous Drones Using TSP and Tree Selection. In: Kadry, S., Prasath, R. (eds) Mining Intelligence and Knowledge Exploration. MIKE 2023. Lecture Notes in Computer Science, vol 13924. Springer, Cham.
7. Chentsov, A.G. A bottleneck routing problem with a system of priority tasks. Izvestiya IMI UdGU, 61 (2023), pp. 156-186 DOI: 10.35634/2226-3594-2023-61-09.
8. Chentsov, A.G.; Chentsov, P.A. An extremal two-stage routing problem and procedures based on dynamic programming, Trudy InstitutaMatematikiiMekhanikiUrO RAN, 2022, vol. 28, no. 2, pp. 215–248.
9. Wu, J.; Li, M.; Gao, C.; Liu, Z.; Zhang, B.; Zhang, X. Research on Deployment Scheme and Routing Optimization Algorithm of Distribution Cable Condition Monitoring Devices. Energies 2023, 16, 6930.
10. Chung, S.H., Sah, B., Lee, J. Optimization for drone and drone-truck combined operations: A review of the state of the art and future directions. Computers&Operations Research, 2020, 123, ID 105004.
11. Wang, H., He, L., Zhang, S., Bai, R., & Wang, Y. (2025). Mobile Robot Path Planning Considering Obstacle Gap Features. Applied Sciences, 15(11), 5979. https://doi.org/10.3390/app15115979
12. Li, J. Optimization method of obstacle avoidance path for dual-arm cooperative robot based on machine vision. Discov Appl Sci 7, 559 (2025). https://doi.org/10.1007/s42452-025-07146-3
13. Yong P. D. et al. Path planning optimization for swine manure-cleaning robots through enhanced slime mold algorithm with cellular automata // Animal sciencejournal. 95(1), 2024. doi: 10.1111/asj.13992.
14. Hu, Y., Li, J., Jiang, L. “Local Path Planning Combined with the Motion State of Dynamic Obstacles”, Bio-Inspired Computing: Theories and Applications. BIC-TA 2022. Communications in Computer and Information Science, 2023, vol. 1801. Springer, Singapore. https://doi.org/10.1007/978-981-99-1549-1_46
15. Макаровских Т.А., Саллам М.Э.Х.М.М. Свидетельство Роспатента о государственной регистрации программы для ЭВМ “Программный модуль для управления системой контроля для устройств проведения мониторинга объектов” №2024616243 от 19.03.2024 , правообладатель: ФГАОУ ВО "ЮУрГУ (НИУ)
16. Dai Y, Lv W, Li S, Zong M. Improving the Lifelong Planning A-star algorithm to satisfy path planning for space truss cellular robots with dynamic obstacles. Robotica. Published online 2025:1-15. doi:10.1017/S0263574725000256
17. Chatzisavvas, A.; Dossis, M.; Dasygenis, M. Optimizing Mobile Robot Navigation Based on A-Star Algorithm for Obstacle Avoidance in Smart Agriculture. Electronics 2024, 13, 2057. doi: 10.3390/electronics13112057

18. Саллам М., Макаровских Т.А. Проектирование оптимальной траектории проведения мониторинга объектов // Вестник ЮУрГУ. Серия: Вычислительная математика и информатика. 2024. Т. 13, № 3. С. 32–46. DOI: 10.14529/cmse240302.

19. Makarovskikh, T., Sallam, M. (2025). Optimal Trajectory for Monitoring Objects with Obstacles. In: Kochetov, Y., Khachay, M., Eremeev, A., Pardalos, P. (eds) Mathematical Optimization Theory and Operations Research. MOTOR 2025. Lecture Notes in Computer Science, vol 15681. Springer, Cham. https://doi.org/10.1007/978-3-031-97077-1_18

20. Саллам М., Макаровских Т.А. Маршрутизация наземного робота, проводящего мониторинг в замкнутом пространстве // В сборнике: Информационные технологии интеллектуальной поддержки принятия решений (памяти проф. Н.И. Юсуповой) ITIDS'2024. Труды X Международной научной конференции. В 2-х томах. Уфа, 2024. С. 86-91.

21. Dai Y, Lv W, Li S, Zong M. Improving the Lifelong Planning A-star algorithm to satisfy path planning for space truss cellular robots with dynamic obstacles. Robotica. Published online 2025:1-15. doi:10.1017/S0263574725000256

22. Chatzisavvas, A.; Dossis, M.; Dasygenis, M. Optimizing Mobile Robot Navigation Based on A-Star Algorithm for Obstacle Avoidance in Smart Agriculture. Electronics 2024, 13, 2057. doi: 10.3390/electronics13112057

23. Михайлов Илья Евгеньевич Практическое сравнение алгоритма a* с алгоритмом волновой трассировки (алгоритмом ли) по быстродействию // Наука, техника и образование. 2016. №1 (19).

Разработка образовательной платформы для курса "Силовые установки" для студентов авиационных специальностей

Шишкин Д. А.
каф. «Конструкции и проектирования летательных аппаратов»
Институт авиационной и ракетно-космической техники
Самарский национальный исследовательский
университет имени академика С. П. Королева
Самара, Россия
e-mail: Cr1stra61@yandex.ru

Аннотация[1]

Рассматривается опыт разработки и внедрения образовательной платформы «Engine Club», предназначенной для сопровождения курса «Силовые установки» в подготовке студентов авиационных специальностей. Актуальность исследования обусловлена необходимостью интеграции электронного обучения (ЭО) или дистанционного обучения (ДО) и дистанционных образовательных технологий в современный образовательный процесс, что позволяет повысить гибкость и индивидуализацию обучения, обеспечить круглосуточный доступ к материалам, повысить качество усвоения теоретических и практических знаний, а также снизить нагрузку на преподавателей. Разработанная платформа включает модули учебных материалов, интерактивные тесты и справочные ресурсы, что упрощает поиск информации и способствует более комфортному и эффективному освоению сложных профессиональных тем. Проведённые опросы студентов подтвердили высокую востребованность данного формата обучения и его значимость для подготовки будущих инженеров. Результаты исследования свидетельствуют о целесообразности интеграции платформы в образовательный процесс и её потенциале для дальнейшего применения в инженерном образовании. Дополнительный опрос пользователей позволил собрать рекомендации по совершенствованию функционала платформы и определению направлений её дальнейшего развития.

Ключевые слова: электронное обучение, дистанционные технологии, платформа, авиастроение, силовые установки, e-learning, remote technologies, platform, aircraft manufacturing, power plants

1. Введение

Современное высшее образование находится в стадии активной цифровой трансформации, что обусловлено необходимостью повышения гибкости и качества учебного процесса. Традиционные формы преподавания, основанные на аудиторных занятиях, всё чаще дополняются электронными и дистанционными технологиями, которые позволяют студентам обучаться в индивидуальном темпе и получать доступ к материалам в любое время [3,4].

В инженерном образовании, где освоение сложных дисциплин требует повторного обращения к материалу и практической отработки, внедрение электронного обучения становится особенно актуальным. Оно способствует не только повышению эффективности усвоения знаний, но и снижению нагрузки на преподавателей за счёт автоматизации ряда процессов [5].

Электронное обучение меняет саму дидактическую структуру образовательного процесса, делая его более интерактивным и ориентированным на самостоятельную работу студентов [1,2]. Это создаёт предпосылки для формирования профессиональных компетенций в условиях цифровой образовательной среды.

Настоящая работа посвящена разработке и внедрению образовательной платформы «Engine Club», предназначенной для сопровождения курса «Силовые установки» для студентов авиационных специальностей. Цель исследования заключается в оценке потенциала платформы как инструмента повышения качества обучения и оптимизации учебного процесса.

[1] Труды XI Международной научной конференции "Информационные технологии интеллектуальной поддержки принятия решений", 11-13 ноября, Уфа-Чандигарх-Фандербейлпарк, 2025

2. Теоретические основы ЭО: методика и актуальность

Современная система образования находится на этапе глубокой трансформации, обусловленной активным внедрением цифровых технологий и медиасредств. Этот процесс отражает переход от традиционной лекционно-семинарской модели к электронным и дистанционным форматам, которые делают обучение более гибким, индивидуализированным и доступным для широкого круга обучающихся.

Дистанционное обучение представляет собой интеграцию существующих педагогических подходов в новую образовательную среду, основанную на цифровых ресурсах и современных средствах коммуникации. Оно объединяет электронные библиотеки, мультимедийные материалы, учебные пособия и обеспечивает взаимодействие преподавателей и студентов через телекоммуникационные технологии.

Особое внимание при проектировании дистанционных курсов уделяется развитию самостоятельной работы студентов, коллективной деятельности и выполнению исследовательских мини-проектов. Большое количество практических заданий сочетается с возможностью регулярных консультаций, что, как показывает международный опыт, способствует более частому и продуктивному взаимодействию с преподавателем по сравнению с традиционными формами обучения.

Актуальность дистанционного обучения сегодня очевидна во всём мире, поскольку оно обеспечивает доступ к образованию тем категориям обучающихся, для которых очная форма является затруднительной. Так, в США значительная часть студентов – это взрослые люди, совмещающие обучение с работой и семейными обязанностями, для которых гибкий формат является оптимальным. Использование дистанционных технологий позволяет не только организовать обучение в удобном режиме, но и значительно сократить временные и финансовые затраты, связанные с посещением очных занятий.

Отправной точкой развития открытого дистанционного образования можно считать начало 1970-х годов (рис. 1). В этот период во многих странах начали активно создаваться университеты нового типа. Однако в отдельных регионах, таких как Ближний Восток и Центральная Америка, развитие дистанционного образования по-прежнему отстаёт, несмотря на низкий уровень охвата населения образованием [1].

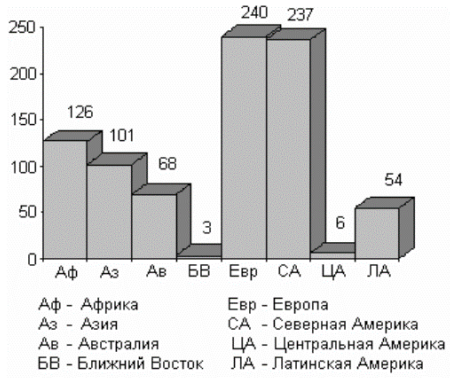


Рис. 1. Число учебных заведений, предлагавших в середине 90-х годов дистанционные образовательные услуги

Система дистанционного образования в России приобретает особую значимость в силу ряда социально-экономических и географических факторов. К числу ключевых причин относятся значительные размеры территории страны, концентрация ведущих научно-образовательных центров преимущественно в крупных городах, внутренняя миграция населения, а также появление новых образовательных запросов со стороны различных категорий граждан. Развитие дистанционного обучения в этих условиях не только расширяет доступ к качественному образованию, но и способствует укреплению позиции России на международном рынке образовательных услуг.

Потенциальная аудитория дистанционного образования в России оценивается экспертами в 1,5–3 млн человек. К ней относятся:

- учащиеся из сельских районов и малых городов;
- менеджеры и руководители предприятий;
- военнослужащие и лица, уволенные в запас;
- граждане, получающие второе высшее образование или проходящие профессиональную переподготовку;
- люди с ограниченными возможностями здоровья;
- русскоязычное население, проживающее за рубежом.

Важным этапом развития системы дистанционного образования стало поручение Правительства Российской Федерации, данное в 1996 году, о разработке концепции и программы его внедрения. Несмотря на то, что программа была реализована лишь частично, данная инициатива привлекла внимание научного и образовательного сообщества и стала импульсом для дальнейшего интенсивного развития дистанционного обучения в российских вузах [2].

2.1. Проблематика ЭО и дидактические подходы

В дидактическом аспекте при разработке и использовании информационно-коммуникационной среды электронного обучения выделяются три ключевые задачи:

1. организация самостоятельной познавательной деятельности обучающихся;
2. обеспечение индивидуальной поддержки учебного процесса со стороны преподавателя;
3. организация коллективной учебной деятельности, включая дискуссии и совместную работу над проектами [2].

В рамках электронного обучения формируются два основных подхода к организации образовательного процесса. Первый ориентирован на воспроизведение традиционной групповой работы в цифровой среде – это виртуальные дискуссии, онлайн-семинары и совместные проекты. В данной модели студенты самостоятельно ищут и анализируют учебную информацию, а преподаватель направляет их деятельность, организуя осмысление и закрепление знаний в формате коллективного взаимодействия.

Второй подход акцентирует внимание на самостоятельной работе студентов с заранее подготовленными электронными интерактивными материалами, которые могут использоваться как локально, так и в сетевой среде. Взаимодействие с преподавателем и другими обучающимися при этом остаётся значимым, но не является критически необходимым, так как ключевые этапы усвоения знаний – освоение, осмысление и закрепление – происходят индивидуально.

С организационной точки зрения первый подход требует более высокой квалификации преподавателей и является значительно более трудоёмким. Второй подход позволяет небольшой команде разработчиков создать учебные материалы, которые затем могут использоваться множеством преподавателей для обучения больших групп студентов [1].

Внедрение электронного обучения предполагает пересмотр дидактических принципов и структуры учебного процесса. Современная цифровая образовательная среда характеризуется следующими особенностями:

- интерактивность – активное участие обучающихся в процессе усвоения знаний и получение обратной связи в режиме реального времени;
- модульность – структурирование учебного материала на самостоятельные логические блоки, что обеспечивает гибкую адаптацию курса под уровень подготовки студента;
- доступность – возможность работы с материалами в любое время и из любого места, что повышает мотивацию и вовлечённость обучающихся;
- адаптивность – настройка сложности и последовательности подачи материала в зависимости от индивидуальных особенностей обучающегося [1,5].

Особую значимость электронное обучение приобретает в инженерном образовании, где подготовка специалистов предполагает не только усвоение теоретических знаний, но и развитие практических навыков. Использование цифровых технологий в этой области позволяет создавать виртуальные лаборатории, интеллектуальные обучающие системы, симуляторы и интерактивные тренажёры, тем самым максимально приближая учебный процесс к реальным условиям профессиональной деятельности [1].

Наиболее эффективным средством с дидактической точки зрения является использование учебных мультимедийных комплексов, которые поддерживают самостоятельную познавательную деятельность обучающихся на всех этапах обучения – от первоначального знакомства с материалом до решения сложных, профессионально-ориентированных задач [2].

3. Внедрение образовательной платформы в учебный процесс

С учётом изложенных выше дидактических и методических принципов была разработана образовательная платформа для дисциплины «Силовые установки», предназначенная для студентов авиационных специальностей Самарского университета. Платформа представляет собой структурированную цифровую среду, обеспечивающую системное изучение теоретических и практических аспектов курса. Её содержание организовано в виде пяти основных разделов: «Расположение двигателей», «Крепление двигателей», «Двигатели», «Тестирование» и «Новости» (рис. 2).



Рис. 2. Структура платформы

Особое место занимает раздел «Двигатели», который включает четыре основных типа силовых установок, каждая из которых представлена несколькими (не менее четырёх) примерами конкретных моделей с их характеристиками и особенностями. Раздел «Тестирование» содержит шесть уникальных проверочных модулей, соответствующих тематическим блокам платформы и предназначенных для самоконтроля и закрепления знаний. Такое структурное решение позволяет организовать обучение по принципу последовательного погружения в материал – от изучения базовых конструктивных решений до практической проверки полученных знаний.

3.1. Архитектура платформы

Информационное наполнение образовательной платформы формировалось на основе учебных пособий и методических материалов Самарского университета, используемых в учебном процессе по дисциплине «Силовые установки». Эти материалы послужили базой для структурирования и представления контента платформы, что позволило обеспечить его соответствие учебной программе и актуальным требованиям подготовки студентов авиационных специальностей [7,8].

При разработке визуализации образовательной платформы особое внимание уделялось выбору интерфейсного стиля, обеспечивающего комфортное восприятие информации пользователями. В результате был выбран минималистический стиль, позволяющий снизить визуальную нагрузку и упростить навигацию. Минимализм способствует концентрации на ключевой информации о двигателях, исключая отвлекающие элементы.

На главной странице платформы представлены основные направления и тематические блоки платформы (рис. 3). Начальный баннер демонстрирует назначение ресурса и включает навигационный элемент, направляющий пользователя к ключевым разделам (кнопка меню).



Рис. 3. Начальный экран платформы

Ниже располагается блок основных разделов, включающий «Двигатели», «Расположение двигателей», «Новости» и «Тестирование». Каждый раздел сопровождается графической иконкой и краткой аннотацией (рис. 4).

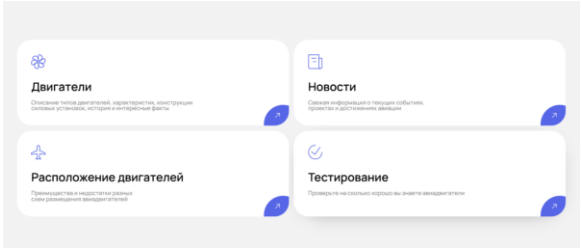


Рис. 4. Блок основных разделов платформы

Далее представлен графический блок «Двигатели» с четырьмя типами авиационных двигателей: поршневые (ПД), турбовинтовые (ТВД), турбовальные (ТВаД) и двухконтурные турбореактивные (ТРДД) или турбовентиляторные двигатели. Каждый тип сопровождается фирменной иконкой и кратким пояснением (рис. 5).

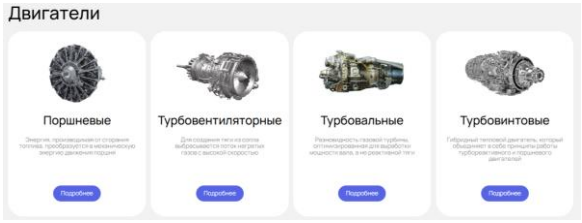


Рис. 5. Блок четырёх типов двигателей

Следующий блок содержит анонсы актуальных новостей, визуализируемых в виде карточек публикаций (рис. 6).

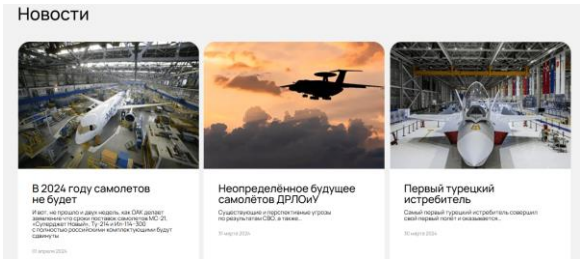


Рис. 6. Блок с новостями

Особое внимание уделено визуализации типов расположения двигателей. Один из четырёх вариантов расположений, которые можно пролистывать с помощью навигационных кнопок. Все типы размещений включают: на пилонах крыла, в корневой части крыла, внутри фюзеляжа и на хвостовой части фюзеляжа (рис. 7).

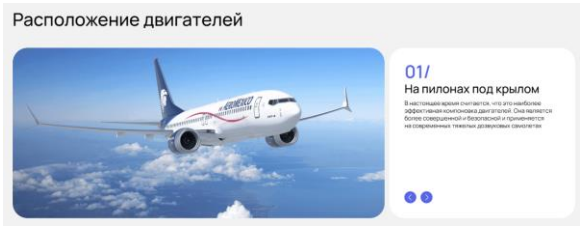


Рис. 7. Блок расположения двигателей

Далее представлен блок, отражающий типовые схемы крепления двигателей для каждого типа (ПД, ТВД, ТРДД), демонстрирующий основные

варианты конструктивной интеграции агрегатов (рис. 8).



Рис. 8. Блок крепления двигателей

Завершающий блок посвящён разделу тестирования. Здесь показана общая структура и цели обучающих тестов, предоставляемых платформой (рис. 9).

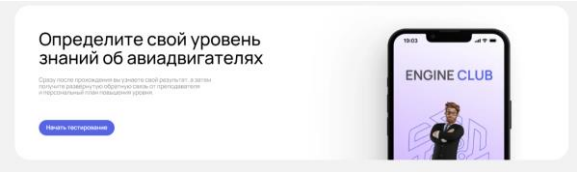


Рис. 9. Блок с тестированием

При переходе в раздел «Двигатели» пользователю открывается интерфейс с отображением четырёх двигателей, принадлежащих к выбранному типу (например, только ПД или только ТВД). Такой шаблон повторяется для всех четырёх категорий силовых установок, формируя наглядную базу агрегатов с ключевыми параметрами (рис. 10). В каждом случае указана модель и иллюстрация агрегата [9,10].

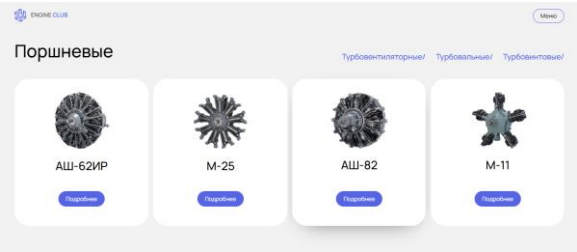


Рис. 10. Блок «Поршневые двигатели»

После выбора конкретного двигателя открывается страница детализированной информации, где в верхней части отображается его изображение и основная техническая справка (рис. 11).



Рис. 11. Страница по двигателю АШ–82

Ниже – интерактивное меню, содержащее семь тематических подразделов: история, модификации, характеристики, применение, конструкция, видеоматериалы и интересные факты (рис. 12,13,14).

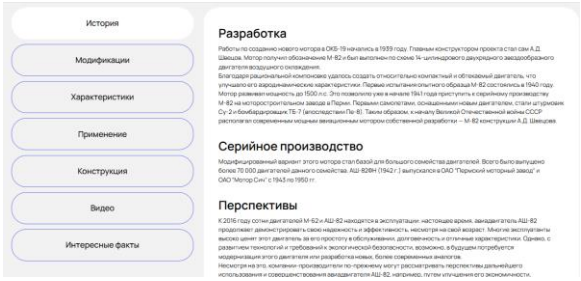


Рис. 12. Подраздел «История двигателя АШ–82»

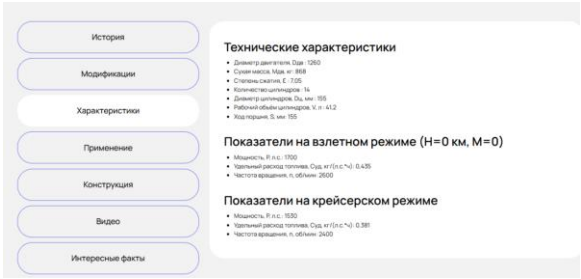


Рис. 13. Подраздел «Характеристики двигателя АШ–82»

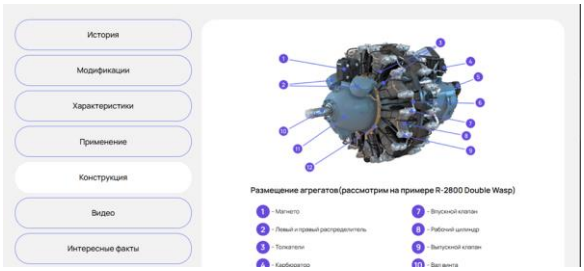


Рис. 14. Подраздел «Конструкция ПД»

При переходе в раздел крепления двигателей отображается информация по креплению: описание, конструктивные особенности, преимущества и недостатки (рис. 15)

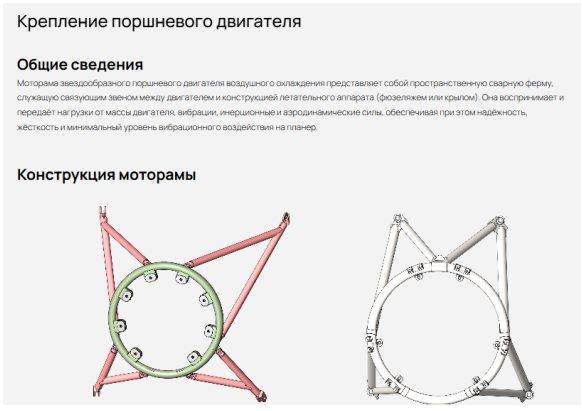


Рис. 15. Страница по креплению ПД

Раздел «Тестирование» реализован как самостоятельный образовательный модуль, доступный пользователям (рис. 16). Здесь показан стартовый экран с шестью доступными тестами: «Всё об авиадвигателях», «Расположение

двигателей», а также четыре теста по каждому из представленных типов двигателей (рис. 17).

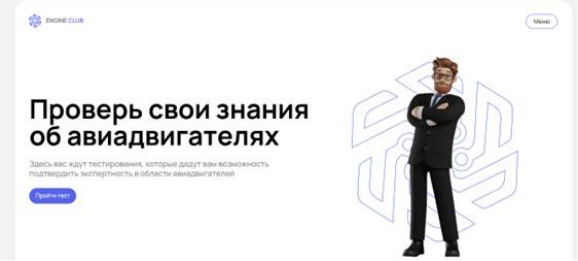


Рис. 16. Стартовый экран тестирования

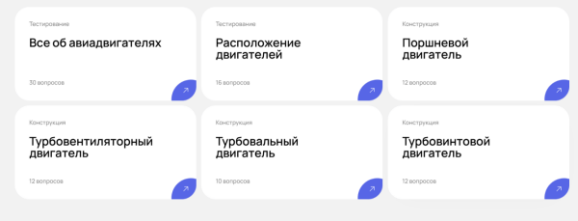


Рис. 17. Блок с 6 тестами

На следующих рисунках продемонстрирован пример работы тестирования. Здесь представлен пример тестового вопроса с одним правильным ответом. После выбора варианта ответа пользователь получает статистику голосования других участников (в процентном выражении), а также цветовую индикацию правильности ответа. Правильный вариант подсвечивается зелёным, а неправильный – красным (рис. 18)

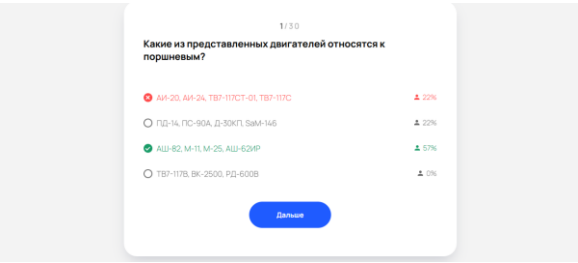


Рис. 2. Тестовый вопрос с результатом вопроса

Итоговый экран прохождения теста, где показаны общий балл, и короткая рекомендация (рис. 19).

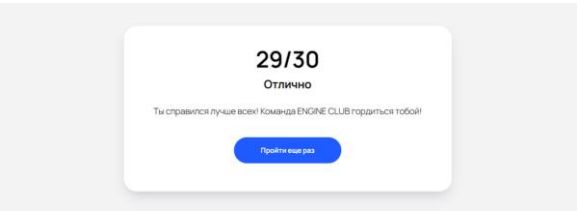


Рис. 19. Итоговый экран прохождения теста

Раздел «Расположение двигателей» представляет собой обобщённую классификацию компоновок силовых установок на самолётах.

При выборе конкретной схемы, например, размещения двигателя на пилоне под крылом, где представлено наглядное изображение данной конфигурации. Ниже следуют три логически выстроенных блока: общее описание, преимущества, а также недостатки выбранной схемы, что позволяет оценить её применимость в зависимости от задач проектирования (рис. 20).



Рис. 20. Пример оформления расположения двигателя

Раздел «Новости» представлен в виде динамической ленты актуальных публикаций и обновлений в сфере авиационных технологий. При нажатии на заголовок выбранной статьи пользователь попадает на страницу с полной версией материала, включающей текстовое описание, изображения и возможные ссылки на дополнительные источники (рис. 21).

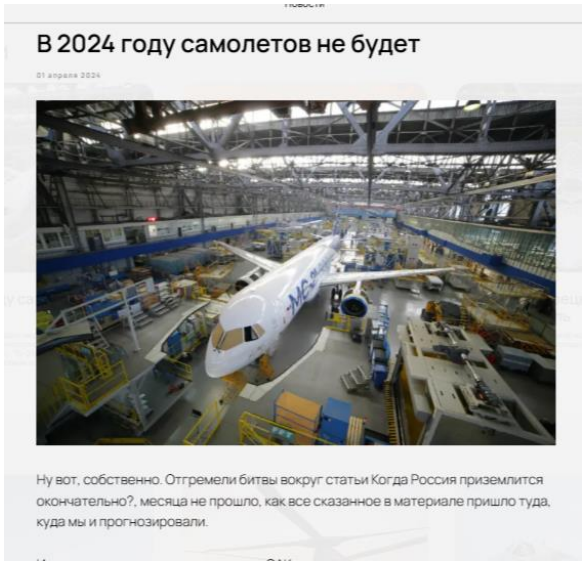


Рис. 21. Содержание статьи

4. Исследование эффективности использования платформы

Важной частью проведённой работы стал сбор и анализ эмпирических данных, направленных на оценку необходимости разработки и внедрения образовательной платформы «Engine Club», а также

выявление её сильных и слабых сторон после первичного внедрения. Исследование проводилось в форме онлайн-опроса с использованием сервиса Google Forms и было ориентировано на студентов, преподавателей и выпускников Самарского университета авиационного профиля.

Цель опроса заключалась в определении уровня заинтересованности целевой аудитории в создании специализированной платформы, оценки востребованности цифровых инструментов в учебном процессе и выявлении направлений для дальнейшего совершенствования разработанного решения. В исследовании приняли участие более 40 респондентов, что позволило получить репрезентативные данные о текущих потребностях и ожиданиях пользователей в области электронного обучения.

4.1. Анализ результатов опроса студентов

Результаты опроса показали, что значительная часть респондентов отмечает недостаток специализированных интернет-ресурсов, посвящённых тематике авиационных двигателей, и испытывает трудности при поиске структурированной и достоверной информации (рис. 22). Вместе с тем большинство участников указали, что при необходимости изучения материалов по силовым установкам они всё же обращаются к профильным сайтам, несмотря на их ограниченное содержание и фрагментарность представленных данных (рис. 23).

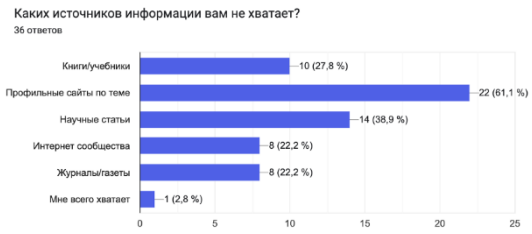


Рис. 22. Статистика про источники



Рис. 23. Статистика про виды источников

Кроме того, опрос выявил предпочтения обучающихся в выборе форм контроля знаний: подавляющее большинство респондентов отдают предпочтение онлайн-тестированию по сравнению с традиционными способами проверки знаний, считая его более удобным и доступным (рис. 24).

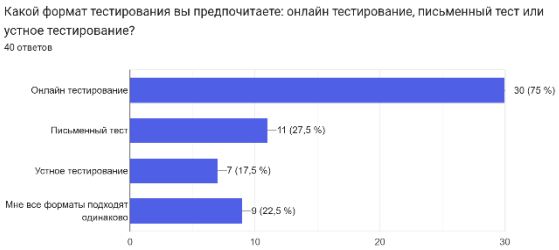


Рис. 24. Статистика про формат тестирования

Полученные результаты подтверждают актуальность разработки специализированной образовательной платформы, объединяющей структурированные материалы по силовым установкам и современные инструменты контроля знаний. Недостаток профильных источников и выраженное предпочтение студентов в пользу онлайн-тестирования указывают на необходимость создания единого цифрового ресурса, который будет одновременно выполнять функции учебного пособия и интерактивной среды обучения. Это позволит повысить качество усвоения материала, упростить доступ к необходимой информации и сделать образовательный процесс более гибким и соответствующим современным требованиям.

4.2. Выявленные рекомендации и направления развития

На этапе тестового использования платформы была собрана дополнительная обратная связь от студентов и преподавателей, направленная на оценку функциональных возможностей и удобства интерфейса. Анализ отзывов позволил выделить как сильные стороны платформы, так и области, требующие доработки (рис. 25, 26).

Среди основных преимуществ респонденты отметили:

- удобную навигацию и логичную структуру размещения материалов;
- возможность быстрого доступа к необходимой информации;
- наличие тестирования для проверки знаний;
- наличие интерактивных элементов, способствующих лучшему усвоению сложных тем;
- снижение необходимости поиска данных в сторонних источниках.

В то же время были обозначены и предложения по улучшению платформы, среди которых наиболее часто упоминались:

- расширение базы учебных материалов и включение дополнительных тематических разделов;
- внедрение адаптивной системы тестирования с возможностью подбора сложности;

- увеличение количества информации по различным составляющим двигателей.

Рябухин·Никита,·гр.1312¶

Недостатки:¶

1.В·разделе·расположение·двигателя·не·хватает·таких·расположений,·как:·над·крылом,·под·¶

фюзеляжем·и·комбинированная·силовая·установка·для·СВВП.¶

2.Неактивность·ленты·новостей.¶

3.На·данный·момент·малое·количество·двигателей.¶

4.В·классификации·отсутствуют·такие·типы·двигателей,·как:··ракетные·двигатели,·воздушно-¶

реактивные·двигатели.¶

6.Отсутствие·графиков·изменения·давления,·скорости·и·температуры·потока·воздуха·на·всех·¶

ступенях·двигателей.¶

Пожелания:¶

Желаю·данному·сайту·исправить·все·недочёты·и·выйти·в·массовое·пользование.·Данный·сайт·¶

может·помочь·многим·конструкторам·двигателей,·самолётов·и·вертолётов.¶

Рис. 25. Отзыв №1

Карабанов Руслан, гр. 1312-240304D

Плюсы:

1. Удобный интерфейс: веб-сайт имеет чистый и современный дизайн, позволяющий легко ориентироваться и находить нужную информацию.

2. Комплексные ресурсы: Engine Club предоставляет широкий спектр ресурсов, включая статьи, учебные пособия и курсы по различным темам

3. Поддержка сообщества: на веб-сайте есть форум, где пользователи могут задавать вопросы, делиться своим опытом и получать помощь от других членов сообщества.

4. Регулярные обновления: веб-сайт регулярно пополняется новым контентом, обеспечивая пользователям доступ к самой последней информации и тенденциям в авиационной промышленности.

Минусы:

1. Языковой барьер: Сайт создан на русском языке, что может стать барьером для пользователей, не владеющих этим языком.

2. Ограниченная поддержка английского языка: хотя на сайте есть некоторые переводы на английский язык, они не являются полными, а некоторые разделы могут быть трудными для понимания нерусскоязычным пользователям.

3. Отсутствие экспертной проверки: хотя на веб-сайте много полезной информации, неясно, проверяется ли ее содержимое экспертами в этой области.

4. Устаревший контент. Некоторый контент на веб-сайте может быть устаревшим, что может стать проблемой в отрасли, где технологии и лучшие практики постоянно развиваются.

Рис. 26. Отзыв №2

Выявленные рекомендации будут использованы в дальнейшей работе над платформой «Engine Club» и позволят повысить её функциональность, удобство и соответствие образовательным потребностям студентов.

5. Заключение

Проведённое исследование подтвердило, что внедрение электронных и дистанционных образовательных технологий является необходимым условием повышения качества подготовки студентов авиационных специальностей. Разработанная образовательная платформа «Engine Club», основанная на современных дидактических принципах, зарекомендовала себя как эффективный инструмент сопровождения курса «Силовые установки».

Платформа обеспечивает структурированный доступ к учебным материалам, интерактивные средства контроля знаний и справочные ресурсы, что способствует более глубокому усвоению материала и развитию практических навыков. Она также снижает нагрузку на преподавателей за счёт автоматизации рутинных процессов и обеспечивает гибкость организации обучения.

Результаты опросов подтвердили востребованность платформы и обозначили направления её дальнейшего совершенствования: расширение учебного контента, внедрение адаптивного тестирования и развитие инструментов обратной связи.

Таким образом, платформа «Engine Club» может рассматриваться как значимый элемент цифровой трансформации учебного процесса, способствующий повышению его эффективности и имеющий потенциал для применения в других инженерных дисциплинах.

Список используемых источников

1. Электронное обучение: проблематика, дидактика, технология / Соловов А. В. – Самара: СГАУ, 2015. – 148 с. (книга).
2. Соловов А. В. Введение в проблематику электронного дистанционного обучения. – Самара: Самар. гос. аэрокосм. ун-т, 2007. – 88 с. (учебное пособие).
3. Полат Е. С. Новые педагогические и информационные технологии в системе образования. – Москва: Академия, 2002. – 273 с. (книга).
4. Ахметшин Э. М. Цифровая трансформация высшего образования: состояние, проблемы и перспективы // Инновации. – 2020. – № 2 (256). – С. 83–87. (статья из журнала).
5. Лапчик М. П. Информационные технологии в образовании. – Москва: Академия, 2001. – 562 с. (учебное пособие).
6. Ваганова О. И., Ильина И. В. Электронные образовательные ресурсы как средство повышения качества образования // Балтийский гуманитарный журнал. – 2020. – Т. 9. – № 2 (31). – С. 203–207. (статья из журнала).
7. Боргест. Н.М. Авиационные силовые установки. – Самара: Издательство Самарского университета, 2020. – 128 с. (учебное пособие)
8. Созонов А.И. Авиационные двигатели. – Ульяновск: УВАУ ГА, 2008. – 72 с. (альбом схем)

9. Зрелов В.А. Отечественные ГТД. Основные параметры и конструктивные схемы (Часть 1): – Самар. гос. аэрокосм. ун-т. Самара, 2002. – 210 с. (учебное пособие)

10. Зрелов В.А. Отечественные ГТД. Основные параметры и конструктивные схемы (Часть 2): – Самар. гос. аэрокосм. ун-т. Самара, 2002. – 250 с. (учебное пособие)

Асимптотический анализ эффективности алгоритма соединения циклов для задачи коммивояжера

Леонова Ю. Ф.
Южно-Уральский государственный университет (НИУ)
кафедра системного программирования
Челябинск, Россия
e-mail: yuliya.igosheva@gmail.com

Панюков А. В.
Южно-Уральский государственный университет (НИУ)
кафедра системного программирования
Челябинск, Россия
e-mail: paniukovav@susu.ru

Аннотация[1]

Задача коммивояжера (ЗК) является одной из наиболее изучаемых NP-трудных задач, для которой активно разрабатываются эвристические методы. В данной работе рассматривается алгоритм соединения циклов (Cycle Merging Algorithm, CMA), основанный на объединении циклов в 2-факторе. Цель работы – провести сравнительный анализ теоретических оценок точности CMA и эмпирических результатов его применения к различным классам экземпляров ЗК. Для анализа использовались вычислительные эксперименты на симметричных, асимметричных и метрических экземплярах ЗК размерностью до 3000 вершин. Эмпирические ошибки аппроксимировались регрессионными моделями. Показано, что для асимметричной ЗК на минимум ошибка CMA убывает с ростом числа вершин, тогда как для симметричной ЗК на минимум наблюдается выраженный рост погрешности. Для метрической ЗК на максимум ошибка быстро убывает, приближаясь к степенной зависимости с показателем порядка –1, при этом полученные ошибки существенно меньше известных теоретических границ, что подтверждает практическую эффективность CMA.

Ключевые слова: задача коммивояжера; алгоритм соединения циклов; эвристика; регрессионный анализ; точность решений.

1. Введение

Задача коммивояжера (Traveling Salesman Problem, TSP) является одной из наиболее известных NP-трудных задач комбинаторной оптимизации, имеющей широкое применение в логистике, планировании маршрутов, организации вычислительных процессов и сжатии данных. Многообразие практических постановок, в которых необходимо находить эффективные маршруты или упорядоченные последовательности действий, привело к формированию целого направления исследований, посвященных разработке точных, приближенных и эвристических методов решения [1, 2, 6].

В последние десятилетия значительное внимание уделяется приближенным и эвристическим алгоритмам, позволяющим находить решения за полиномиальное время при приемлемом отклонении от оптимума. Классическими результатами в данной области являются алгоритм Кристофидеса для метрического варианта MIN-TSP, обеспечивающий гарантию аппроксимации $3/2$ [1], а также семейство локальных эвристик типа Lin–Kernighan, которые показывают высокое качество решений на практике [2]. Существенный вклад внесли и стохастические подходы, в частности анализ поведения TSP на случайных графах, выполненный А. Фризом [3]. Работы О. Свенссона и соавторов [4] уточнили границы точности для вариаций алгоритма Кристофидеса и связали их с оценками интегральности LP-релаксаций.

Несмотря на накопленные результаты, исследование алгоритмов с улучшенными гарантиями качества остается актуальной задачей, особенно для варианта MAX-TSP, где оптимизируется максимальная длина цикла. В этом направлении интерес представляет алгоритм соединения циклов (Cycle Merging Algorithm, CMA), обладающий рядом преимуществ: простота реализации, возможность эффективной параллельной обработки и наличие строгих оценок качества решения. Дальнейшее исследование CMA позволяет не только уточнить его теоретические свойства, но и сопоставить гарантированные оценки

[1] Труды XI Международной научной конференции "Информационные технологии интеллектуальной поддержки принятия решений", 11-13 ноября, Уфа-Чандигарх-Фандербейлпарк, 2025

с эмпирическими результатами, получаемыми на практических наборах данных.

2. Цель

Целью настоящего исследования является анализ и обобщение свойств алгоритма соединения циклов (Cycle Merging Algorithm, CMA) для метрической задачи коммивояжера на максимум. В работе акцент сделан на сравнении теоретических оценок точности алгоритма с эмпирическими результатами его работы.

3. Методы

В исследовании применялись как аналитические, так и экспериментальные методы.

3.1. Теоретический анализ

Были рассмотрены известные оценки точности алгоритма соединения циклов (Cycle Merging Algorithm, CMA) для метрической задачи коммивояжера на максимум. В основу анализа положено доказательство гарантированной нижней границы относительной точности $5/6$, установленной для худшего случая.

3.2. Экспериментальные исследования

Для сопоставления с теоретическими оценками проведён статистический анализ качества решений, получаемых CMA на случайных экземплярах задачи различной размерности. В качестве входных данных использовались полные графы с метрическими весами рёбер, сгенерированные по равномерному распределению.

3.3. Методы оценки качества

Для каждого экземпляра задачи вычислялось отношение полученного веса гамильтонова цикла к решению задачи о назначении. Далее результаты усреднялись по сериям запусков, что позволило построить регрессионные зависимости и выявить характер изменения относительной ошибки с ростом числа вершин.

4. Результаты

В настоящем разделе представлены результаты анализа качества алгоритма соединения циклов (CMA) применительно к различным постановкам задачи коммивояжера. Рассмотрены как теоретические оценки точности, так и эмпирические результаты, полученные в ходе вычислительных экспериментов. Основное внимание уделено сопоставлению доказанных оценок для худшего случая с фактическими характеристиками алгоритма на реальных данных, что позволяет выявить пределы применимости теоретических границ и обосновать практическую эффективность CMA.

4.1. Теоретические результаты

В теоретической части исследования была доказана следующая оценка точности алгоритма CMA:

$$\frac{W_{alg}}{W_{opt}} \geq \frac{5}{6},$$

где W_{alg} – вес решения, построенного CMA, а

W_{opt} – оптимальное значение задачи.

Эта граница является лучшей из известных для CMA при числе вершин до примерно 500 000 и дополняет результаты В. В. Шенмайера, который ранее получил асимптотическую оценку, стремящуюся к нулю при росте размерности. В отличие от него, полученная нами граница не является асимптотической и дает более сильную характеристику поведения алгоритма в области практически значимых размеров задачи.

Важно подчеркнуть, что и данная оценка, и результаты В.В. Шенмайера относятся к худшему случаю, тогда как на практике фактическая ошибка алгоритма обычно оказывается существенно меньше. Поэтому особый интерес представляет сопоставление теоретической нижней границы с эмпирическими зависимостями, которые будут рассмотрены далее.

4.2. Эмпирический анализ

Для оценки практической эффективности алгоритма соединения циклов (CMA) были проведены вычислительные эксперименты на различных классах задач коммивояжера: симметричных и асимметричных задачах на минимум, а также метрической задаче на максимум.

Методика экспериментов

Тестирование проводилось на выборках, включающих задачи размерностью от 100 до 3000 вершин. Для каждой размерности выполнялась серия из 100 запусков, что обеспечивало статистическую устойчивость полученных оценок. В качестве показателей качества рассматривалась относительная ошибка алгоритма:

$$\varepsilon = \frac{W_{alg} - W_{opt}}{W_{opt}} \cdot 100\%.$$

Симметричная задача на минимум

Для симметричной задачи коммивояжера выбор адекватной формы регрессионной зависимости оказался нетривиальным. Были протестированы степенная, полиномиальная и логарифмическая модели, каждая из которых по-разному описывает характер изменения относительной погрешности при росте размерности задачи.

Построенная *степенная модель* имеет вид

$$y \approx 0{,}105 \cdot n^{0{,}491},$$

где y – относительная погрешность, n – число вершин. Оба коэффициента модели статистически значимы ($p \leq 2 \cdot 10^{-16}$). Коэффициент детерминации $R^2 = 0{,}9997$ свидетельствует о практически полном соответствии модели данным.

Для параметров модели получены следующие 99%-доверительные интервалы:

для свободного члена: $\beta_0 = -2{,}2555, [-2{,}2896, -2{,}2214]$;

для коэффициента при $\log n$: $\beta_1 = 0{,}4911, [0{,}4863, 0{,}4959]$.

Таким образом, показатель степени b надежно оценивается в окрестности $0{,}49$, что указывает на близкую к корневой зависимость относительной погрешности от числа вершин. Однако оценка регрессии выявила гетероскедастичность и несоответствие ряду предпосылок линейной регрессии. Это делает её менее надёжной при экстраполяции.

Полиномиальная модель второго порядка

Построенная полиномиальная регрессионная модель второго порядка имеет вид:

$$\log y = \beta_0 + \beta_1 \cdot \log n + \beta_2 \cdot (\log n)^2.$$

С использованием уровня значимости 99% были получены доверительные интервалы коэффициентов:

$$\beta_0 \in [-2{,}221;\ -1{,}993],$$

$$\beta_1 \in [0{,}411;\ 0{,}482],$$

$$\beta_2 \in [0{,}00084;\ 0{,}00616].$$

Отрицательное значение β_0 отражает смещение базового уровня зависимости при малых значениях n. Коэффициент $\beta_1 > 0$ статистически значим и характеризует основную тенденцию роста относительной погрешности при увеличении числа вершин. Положительное значение β_2 показывает небольшое ускорение роста при малых n, которое постепенно сглаживается для больших размерностей. Таким образом, модель учитывает как линейный, так и слабый квадратичный компонент в логарифмической шкале, что обеспечивает более точное описание эмпирических данных по сравнению с простой степенной регрессией.

Логарифмическая модель имеет вид

$$y = a + b \cdot \log n\ ,$$

где y – относительная погрешность, а n – число вершин. Подставив оценки коэффициентов, получаем уравнение:

$$y = -6{,}291 + 1{,}404 \cdot \log n.$$

Статистический анализ показал, что оба коэффициента модели являются значимыми ($p < 2 \cdot 10^{-16}$). Коэффициент детерминации $R^2 = 0{,}949$ (скорректированный $R^2 = 0{,}947$) указывает на достаточно хорошее соответствие модели данным, хотя точность аппроксимации оказалась ниже, чем у степенной и полиномиальной моделей. Стандартная ошибка остатков составляет 0,282, что существенно выше по сравнению с альтернативными моделями.

Для коэффициентов построены доверительные интервалы с уровнем значимости 0,99:

свободный член принимает значения в интервале $[-7{,}51; -5{,}08]$, коэффициент при $\log n$ лежит в диапазоне $[1{,}23; 1{,}57]$.

Интерпретация полученных коэффициентов заключается в следующем. Отрицательный свободный член отражает низкие значения ошибки при малых n. Положительный коэффициент при $\log n$ указывает на линейный рост погрешности с логарифмом размерности задачи.

Проверка предпосылок регрессии выявила ряд нарушений. Тест Дарбина–Уотсона показал сильную положительную автокорреляцию остатков ($DW = 0{,}219, p < 10^{-12}$), тест Бреуша–Пагана подтвердил наличие гетероскедастичности ($p = 8{,}3 \cdot 10^{-4}$), а глобальный тест (GVLMA) указал на несоответствие предпосылкам линейности (в частности, нарушение связи переменных и асимметрии распределения остатков).

Таким образом, логарифмическая модель удовлетворительно описывает общий замедляющийся рост относительной погрешности с увеличением числа вершин, однако по качеству аппроксимации и статистическим характеристикам уступает полиномиальной модели второго порядка. Это ограничивает её применение для прогнозирования, хотя в рамках исследованных диапазонов n она может использоваться для анализа тенденций.

Сравнительный анализ трех регрессионных моделей показал следующее (рис 1). Степенная модель обеспечивает высокую точность аппроксимации ($R^2 = 0{,}9997$), однако нарушает предпосылки регрессии (гетероскедастичность, отклонения в глобальном тесте), что снижает её надежность. Логарифмическая модель имеет более простую форму и наглядно отражает замедляющийся рост погрешности, однако демонстрирует заметно худшие статистические показатели ($R^2 = 0{,}949$) и сильные нарушения предпосылок. Наиболее сбалансированным вариантом оказалась полиномиальная модель второго порядка: она сочетает высочайшее качество аппроксимации ($R^2 = 0{,}9999$) с удовлетворительным выполнением ключевых статистических требований и предоставляет более гибкое описание зависимости. Поэтому для дальнейшего анализа и прогнозирования выбрана именно полиномиальная модель.

Прогноз для задач до 10000 вершин (рис 2) подтверждает тенденцию замедления роста погрешности. Таким образом, поведение ошибки CMA в симметричном случае характеризуется устойчивостью: хотя абсолютная величина ошибки возрастает, её прирост становится всё менее значительным. Это согласуется с предположением о практической применимости алгоритма к задачам большой размерности.

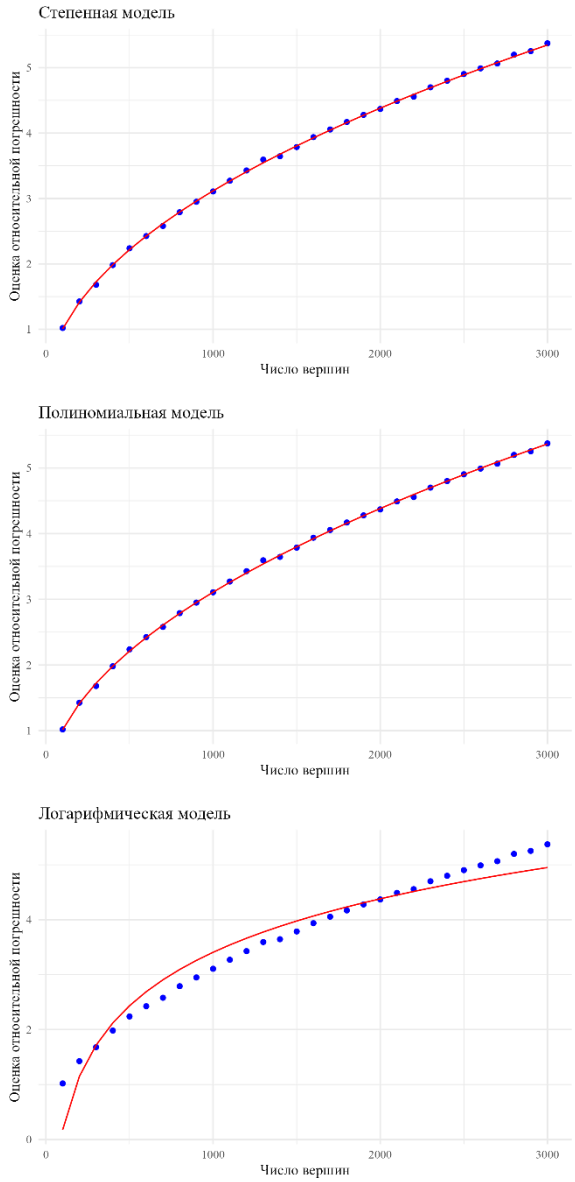


Рис 1. Сравнение моделей регрессии оценки относительной погрешности CMA в симметричной задаче на минимум

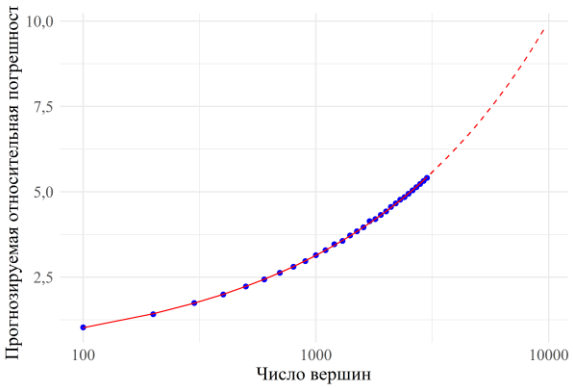


Рис 2. Прогноз оценки относительной погрешности CMA для симметричной задачи

Асимметричная задача на минимум

Для асимметричной задачи коммивояжера была построена степенная модель в логарифмическом виде:

$$\log(error) = 4{,}374 - 0{,}463 \cdot \log n.$$

Коэффициенты модели статистически значимы ($p < 2 \cdot 10^{-16}$), коэффициент детерминации $R^2 = 0{,}964$ свидетельствует о высоком качестве аппроксимации. Доверительные интервалы при уровне значимости 0,99 составили: для свободного члена [4,040; 4,707], для коэффициента при $\log n$ [–0,510; –0,416].

Проверка предпосылок линейной регрессии показала хорошее качество модели:

- тест Дарбина–Уотсона ($DW = 2{,}92, p = 0{,}994$) подтвердил отсутствие автокорреляции остатков;
- тест Бреуша–Пагана ($p = 0{,}094$) не выявил гетероскедастичности;
- глобальный тест (gvlma) показал, что все предпосылки модели выполняются ($p = 0{,}562$).

Таким образом, модель можно считать устойчивой и статистически корректной. Она отражает замедляющееся уменьшение относительной погрешности при росте размерности задачи (рис 3), что соответствует интуитивным ожиданиям: с увеличением числа вершин влияние локальных отклонений снижается, и алгоритм демонстрирует более стабильное поведение.

Сопоставление результатов для симметричной и асимметричной задач показывает принципиально различное поведение относительной погрешности решений, полученных алгоритмом соединения циклов.

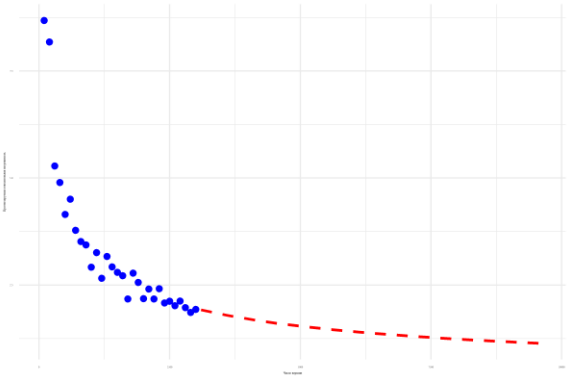


Рис. 3. Прогноз оценки относительной погрешности CMA для асимметричной задачи

Сложность симметричной неевклидовой задачи коммивояжера для алгоритмов построения тура, начинающихся с решения задачи о назначении, обусловлена низкой точностью нижней границы, полученной на основе этого метода. В частности, для подавляющего большинства симметричных экземпляров задача о назначении приводит к цикловому покрытию, содержащему большое количество коротких циклов (табл. 1). Это значительно затрудняет процесс их объединения в единый маршрут, увеличивая как вычислительные затраты, так и итоговую относительную погрешность.

Таблица 1. – Отношение числа начальных циклов к числу вершин задачи

Число вершин	Асимметричные экземпляры	Симметричные экземпляры
100	0,0382	0,4920
200	0,0225	0,4768
300	0,0162	0,4771
400	0,0139	0,4892
500	0,0118	0,4848
600	0,0101	0,4867
700	0,0086	0,4896
800	0,0076	0,4902
900	0,0071	0,4890
1000	0,0070	0,4878

Как видно из таблицы, для асимметричных экземпляров отношение числа начальных циклов к числу вершин существенно уменьшается с увеличением размерности задачи, снижаясь с 0,0382 (100 вершин) до 0,0070 (1000 вершин). Это указывает на тенденцию к формированию более длинных циклов, что способствует более эффективному соединению маршрутов на последующих этапах работы алгоритма.

В отличие от этого, для симметричных экземпляров данное отношение остается практически неизменным, колеблясь в диапазоне 0,4768–0,4920. Это свидетельствует о том, что при формировании 2-фактора для данного семейства задач образуется множество коротких подциклов, что затрудняет их объединение в итоговый маршрут.

Метрическая задача на максимум

Особое внимание уделялось метрической задаче на максимум, так как именно для неё доказана теоретическая граница точности 5/6. Эксперименты показали, что фактическая ошибка CMA значительно ниже этой оценки. Более того, полученные данные описываются убывающей степенной моделью, что подтверждает тенденцию устойчивого уменьшения ошибки при росте числа вершин.

После логарифмирования данных получено уравнение линейной регрессии:

$$\log(error) = -1{,}078 - 0{,}999 \cdot \log(n),$$

где n – число вершин.

В исходной шкале это соответствует степенной зависимости: $error(n) \approx 0{,}341 \cdot n - 0{,}999$.

Таким образом, относительная погрешность убывает обратно пропорционально числу вершин.

При этом оба коэффициента статистически значимы ($p < 2 \cdot 10^{-16}$), коэффициент детерминации $R^2 = 0{,}973$, что указывает на отличное соответствие модели данным. Остатки распределены симметрично и не демонстрируют выраженных систематических отклонений. Доверительные интервалы gри уровне значимости 0,01 (доверительность 99%):

свободный член: $[-1{,}699; -0{,}457]$,

коэффициент при $\log(n)$: $[-1{,}086; -0{,}913]$.

Это подтверждает устойчивость модели и надежность знака коэффициентов.

Тесты Дарбина–Уотсона ($DW = 2{,}00, p = 0{,}422$) и Бреуша–Пагана ($p = 0{,}787$) подтверждают выполнение предпосылок линейной регрессии. Глобальный тест предпосылок ($p = 0{,}669$), все частные тесты также не отвергают гипотезу о корректности модели.

Зависимость ошибки от числа вершин носит обратно пропорциональный характер: при увеличении размерности задачи относительная погрешность снижается примерно как $1/n$. Это согласуется с теоретическим нижним пределом точности алгоритма CMA для метрической задачи на максимум и подтверждает устойчивое поведение алгоритма: чем больше размерность, тем ближе решение к оптимальному (рис. 4).

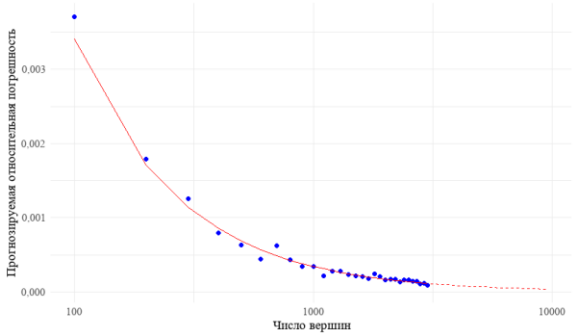


Рис. 4. Оценка относительной погрешности CMA для метрической задачи на максимум.

4.3. Сравнение с теоретическими оценками

Сравнение эмпирических результатов с теоретической границей точности, полученной В.В. Шенмайером, показывает существенный разрыв между прогнозируемыми и фактическими значениями ошибки. В работе Шенмайера нижняя граница точности CMA в метрической задаче на максимум задаётся в терминах худшего случая и асимптотически стремится к нулю при $n \to \infty$. Такая оценка корректна с точки зрения теории, но она чрезмерно консервативна: на практике наблюдаемые значения относительной ошибки значительно меньше и не демонстрируют тенденции к достижению предсказанного уровня даже при задачах размерности до нескольких тысяч вершин (рис. 5).

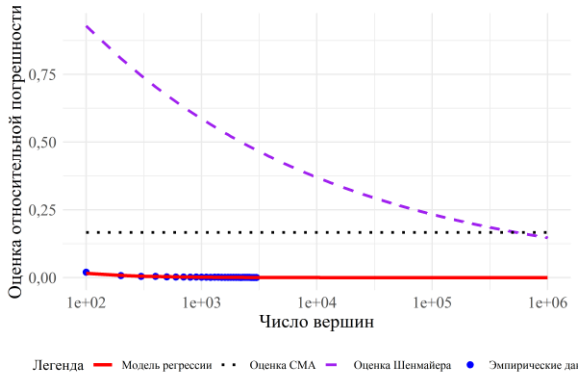


Рис. 5. Сопоставление теоретических границ и эмпирических ошибок CMA в метрической задаче на максимум

Причина этого расхождения заключается в различии между теоретическими худшими случаями и типичными экземплярами. Теоретическая граница учитывает возможность крайне неблагоприятных конфигураций графа, тогда как в реальных вычислительных экспериментах (в том числе на случайно распределённых и практико-ориентированных данных) такие конфигурации встречаются редко. В результате фактическая погрешность CMA оказывается существенно ниже, а снижение ошибки при росте числа вершин носит устойчивый и плавный характер.

Подобное расхождение между теоретическими гарантиями и эмпирическим поведением фиксируется и в других исследованиях. В частности, Фриз в ряде работ по вероятностному анализу задачи коммивояжера [3, 9] показал, что для случайных графов оптимальные и приближённые решения имеют существенно лучшие характеристики по сравнению с худшими случаями, а относительная ошибка снижается с ростом размерности задачи. Гловер, Гутин и Зверович в обзоре по эвристикам построения тура [7] отмечают, что методы объединения циклов, несмотря на относительно слабые строгие оценки, на практике демонстрируют гораздо более высокое качество решений для симметричных и асимметричных TSP. Аналогичные результаты были получены в работах О. Свенссона и соавторов [10], которые анализировали поведение алгоритмов на LP-релаксациях: фактические ошибки алгоритмов оказывались существенно ниже, чем следовало бы из общих теоретических границ.

Таким образом, эмпирический анализ подтверждает, что CMA:

- стабильно обеспечивает меньшую погрешность, чем теоретически гарантированная оценка в худшем случае;
- демонстрирует тенденцию к быстрому снижению ошибки с ростом числа вершин;
- соответствует наблюдаемым закономерностям для других классов алгоритмов, где практическое качество решений значительно выше теоретически гарантированного уровня.

Это подтверждает практическую ценность CMA: в реальных условиях он не только укладывается в строгие теоретические гарантии, но и значительно превосходит их, что делает данный алгоритм конкурентоспособным по совокупности точности и трудоемкости.

Таким образом, эмпирический анализ подтверждает, что CMA демонстрирует устойчивое и предсказуемое поведение, обеспечивая высокое качество решений на задачах большой размерности и значительно превосходя ожидания, основанные на теоретических худших оценках.

5. Заключение

Основные результаты работы:

1. Проведен анализ точности алгоритма соединения циклов (Cycle Merging Algorithm, CMA) для задачи коммивояжера на классах симметричных, асимметричных и метрических экземпляров.
2. Построены регрессионные модели, описывающие зависимость относительной погрешности CMA от числа вершин.

3. Установлено, что для асимметричных задач на минимум ошибка CMA убывает с ростом размерности, что подтверждает устойчивость алгоритма в таких случаях. Для симметричных задач на минимум погрешность растет.
4. Для метрической задачи на максимум выявлено быстрое убывание ошибки, описываемое степенной зависимостью с показателем близким к –1.
5. Показано, что эмпирические ошибки CMA существенно меньше, чем известные теоретические границы, что подтверждает практическую эффективность алгоритма на задачах большой размерности.

Список используемых источников

1. Christofides N. Worst-case analysis of a new heuristic for the travelling salesman problem // Carnegie-Mellon University, 1976. 20 p.
2. Lin S., Kernighan B.W. An effective heuristic algorithm for the traveling-salesman problem // Operations Research. 1973. Vol. 21. № 2. P. 498–516.
3. Frieze A.M. On the value of a random minimum spanning tree problem // Discrete Applied Mathematics. 1984. Vol. 10. № 1. P. 47–56.
4. Svensson O., Tarnawski J., Végh L.A. A new bound for the integrality gap of the subtour elimination LP for the TSP // Mathematics of Operations Research. 2018. Vol. 43. № 4. P. 1482–1500.
5. Shenmaier V. Asymptotic analysis of a cycle-patching heuristic for the Euclidean traveling salesman problem // Journal of Applied and Industrial Mathematics. 2022. Vol. 16. № 4. P. 636–647.
6. Applegate D.L., Bixby R.E., Chvátal V., Cook W.J. The Traveling Salesman Problem: A Computational Study. — Princeton: Princeton University Press, 2006. — 593 p.
7. Glover F., Gutin G., Zverovich A. Construction heuristics for the traveling salesman problem: an empirical evaluation // European Journal of Operational Research. – 2001. – Vol. 129, № 3. – P. 552–559. DOI: 10.1016/S0377-2217(99)00464-5.
8. Johnson D. S., McGeoch L. A. Experimental analysis of heuristics for the STSP //The traveling salesman problem and its variations. – Boston, MA : Springer US, 2002. – P. 369-443.
9. Frieze A.M. An analysis of a simple heuristic for the traveling salesman problem // Random Structures & Algorithms. – 1984. – Vol. 5, № 3. – P. 297–318. DOI: 10.1002/rsa.3240050302.
10. Svensson O., Tarnawski J., Végh L.A. A new approximation for the asymmetric traveling salesman problem // Proceedings of the 50th Annual ACM SIGACT Symposium on Theory of Computing (STOC'2018). – New York: ACM, 2018. – P. 853–860. DOI: 10.1145/3188745.3188750.

Диверсификация поиска в метаэвристических алгоритмах одномерной упаковки

Усманова А. Р.
Институт информатики, математики и робототехники
Уфимский университет науки и технологий
Уфа, Россия
e-mail: kfmn2004@mail.ru

Валиахметова Ю. И.
Институт информатики, математики и робототехники
Уфимский университет науки и технологий
Уфа, Россия
e-mail: julikas@inbox.ru

Аннотация[1]

Сложность решения многих оптимизационных задач связана, во-первых, с их NP-полнотой и невозможностью применения методов перебора, а, во-вторых, с дискретностью данных задач и невозможностью использования классических методов математического анализа. Большинство такого рода задач решается с помощью метаэвристических алгоритмов. К ним относятся как алгоритмы, имитирующие поведение биологических объектов – эволюционные и генетические, алгоритмы муравьиной и пчелиной колонии, так и имитирующие физические процессы – алгоритм моделирования отжига. В большинстве подобных алгоритмов при поиске в пространстве решений существует проблема выхода из локального оптимума – невозможности улучшить текущее решение, когда оно является локально оптимальным. В данной работе рассматривается задача одномерной упаковки BPP (Bin Packing Problem) и описывается процедура диверсификации, позволяющая существенно преобразовать текущее решение с целью выхода из окрестности решений с локальным оптимумом. Кроме того, авторы предлагают рассматривать нестандартную постановку задачи – в виде задачи расписания параллельных процессоров.

Решения, подлежащие обработке, оказываются неприемлемыми применительно к изначальным условиям задачи упаковки. Нахождение приемлемого решения в рамках задачи составления расписания чаще всего подразумевает достижение оптимальных результатов в задаче упаковки. Данная работа описывает методику локального поиска среди близких вариантов полиномиальной размерности.

Среди задач одномерной упаковки особую сложность представляют собой, так называемые триплеты – задачи, в которых веса всех предметов близки к трети объема контейнера. Предлагаемая диверсификация разработана в первую очередь именно для таких задач. Вычислительный эксперимент проводился с целью определения параметров предлагаемой процедуры диверсификации. В заключении сформулированы полученные результаты и выводы.

Ключевые слова: Bin Packing Problem; метаэвристики; упаковка; локальный поиск; диверсификация; триплеты.

1. Введение

Задачу одномерной упаковки контейнеров - Bin Packing Problem (BPP) можно сформулировать следующим образом.

Задача 1. Имеется множество $L = \{w_1, w_2, ..., w_n\}$ весов объектов, а так же задано значение емкости контейнера, обозначаемое символом C. Требуется разместить все объекты в наименьшем возможном количестве контейнеров так, чтобы суммарный вес каждого контейнера оставался в допустимых пределах и не превышал его емкости. Доказано, что BPP является NP-полной задачей [1,12].

Далее мы рассмотрим альтернативную проблему – распределение заданий по параллельному ряду вычислительных устройств.

Задача 2. Предположим, дано m центральных процессоров и n операций с соответствующими

[1] Труды XI Международной научной конференции "Информационные технологии интеллектуальной поддержки принятия решений", 11-13 ноября, Уфа-Чандигарх-Фандербейлпарк, 2025

длительностями исполнения $L = \{w_1, w_2, ..., w_n\}$. Необходимо распределить каждую операцию на какой-то процессор таким образом, чтобы общее максимальное время обработки T на любом из процессоров было минимальным. Ясно, что найденное оптимальное расписание, при котором $T \le C$, автоматически решает первоначальную задачу BPP первого типа с параметром целевого функционала равным количеству используемых процессоров m.

Большинство метаэвристических алгоритмов для решения проблемы размещения реализуются посредством пошагового алгоритма. Исходное предварительное размещение подвергается серии последовательных модификаций согласно определенной схеме действий. Рассмотрим обобщенную структуру алгоритма.

Алгоритм «Упаковка»

1. Присвоить переменной i значение 0, соответствующее номеру текущего шага. Используя метод RPP, сформировать начальное приближенное решение x_0, при котором значение целевой функции составит Nrpp. Зафиксировать текущее количество контейнеров N_i= Nrpp и установить наилучший достигнутый показатель целевой функции N*= N_i.
2. Запустить процедуру удаления одного контейнера, $N_i=N_i-1$.
3. Выполнить запуск алгоритма табу-поиска TS. Обнулить счетчик итераций улучшения решения. Для текущего решения сформировать две окрестности решений. Выбрать наилучшее решение из этих окрестностей, обновить список запретов. Если новое решение имеет лучшее значение TO, запомнить новое решение как наилучшее из достигнутых и обнулить счетчик итераций улучшения решения, иначе – увеличить счетчик итераций улучшения решений.
4. Перейти к шагу 5, если выполнено одно из условий:

   А) превышено общее допустимое число итераций;

   Б) **в течение заданного числа не происходило смены наилучшего решения** – это и будет критерием вызова диверсификации;

   В) получено допустимое решение без превышения емкости контейнеров.

   Если оба условия не соблюдаются, вернуться к выполнению шага 3.
5. Присвоить i новое значение: i=i+1. Если найдено допустимое размещение, обновить лучший результат N*= N_i, и перейти к шагу 5. В противном случае – перейти к шагу 6.
6. Если достигнуто минимально возможное количество контейнеров (т.е. $N_i=N_0$), значит, найдено оптимальное решение, завершить выполнение алгоритма. Иначе продолжить с шага 2.
7. Завершить работу алгоритма с выводом итогового значения N*.

Поясним следующие моменты:

1. Метод RPP является простой вероятностной эвристикой на основе алгоритма FFD (первый подходящий с упорядочиванием), описан в [3,4].
2. Алгоритм «уничтожения» контейнера и структура данных представления плана упаковки описана в [5,6].
3. После уничтожения контейнера решение становится недопустимым в Задаче 1, и поиск с запретами осуществляется уже для Задачи 2.
4. Нижняя граница вычисляется как $N_0 = \left\lceil \sum_{i=1}^{n} w_i / C \right\rceil$, её достижение признак заведомой оптимальности решения в **Задаче 1**.
5. Допустимость полученного решения, указанного в пункте 4, определяется путем контроля нулевого показателя общей степени перегрузки всех контейнеров TO (Total Overfilling). Пусть S_j представляет массу груза внутри j-го контейнера, где j=1,…,m, где m – число используемых контейнеров. Для оценки объема превышения грузоподъемности введем дополнительные коэффициенты d_j, j=1,…,m, отражающие численно степень заполнения каждого отдельного контейнера сверх нормы следующим образом:

$$d_j = \begin{cases} S_j - C, & \text{если } C < S_j \\ 0, & \text{иначе} \end{cases}$$

   Тогда суммарная переполненность контейнеров

$$TO = \sum_{j=1}^{m} d_j$$

6. Метод поиска с запретами подробно описан в работе [6,14], а также в работе [7] описаны две окрестности полиномиальной мощности, в которых ведется поиск. Поскольку вид данных окрестностей повлиял на разработку предлагаемой процедуры диверсификации, приведём их краткое описание. Первая окрестность включает варианты размещений, формируемые путем перемещения единственного элемента из одной емкости в

другую. Вторая окрестность представлена решениями, получаемыми вследствие обмена двух элементов, находящихся изначально в разных контейнерах.

7. При выборе последующего варианта применяется критерий оценки, выраженный как разница масс двух контейнеров, из которых перемещаются предметы. Таким образом, использование указанной функции направлено на снижение дисбаланса нагрузок между контейнерами.

Многим итерационным методам [13] присуща проблема остановки в локальном оптимуме, когда целевая функция не претерпевает улучшения в течение некоторого числа итераций Решением этой проблемы и является процедура диверсификации. В некоторых эволюционных и генетических алгоритмах роль диверсификации играет процедура мутации. Качественная и эффективная процедура диверсификации должна на взгляд авторов удовлетворять следующим требованиям.

- Критерии и условия применения диверсификации должны быть обоснованы, нарушение этого требования будет приводить к тому, что алгоритм не будет успевать «спускаться» в локальные минимумы.

- Сама процедура диверсификации не должна обладать высокой вычислительной сложностью.

- Степень изменения плана упаковки с одной стороны должна быть значительной, чтобы «уводить» процесс из локального оптимума. С другой стороны, чрезмерное или просто случайное изменение плана может привести к недопустимо большим потерям времени при возврате к относительно «хорошим» решениям.

Далее изложена предлагаемая авторами процедура диверсификации.

2.Процедура диверсификации

Метод диверсификации заключается в модификации существующего распределения путем замены пары элементов из одного контейнера одним элементом другого контейнера. Выбор пар контейнеров для обмена неслучаен: два объекта берутся из контейнера с наибольшей степенью перегрузки, тогда как третий элемент – наиболее тяжелый объект – извлекается из резервуара с минимальной суммой масс сравниваемых объектов.

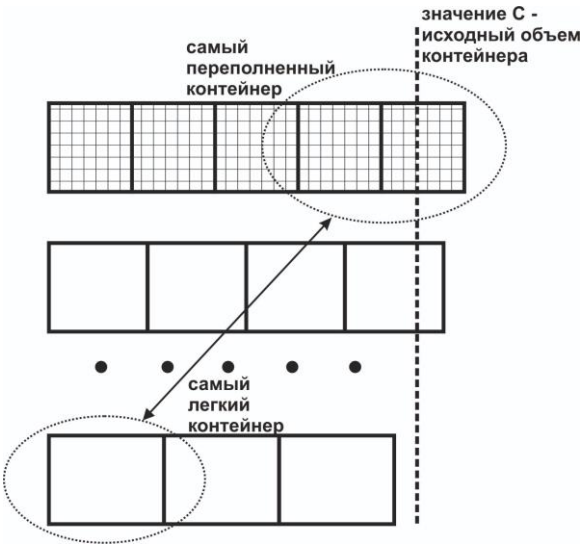


Рис.1 Обмен 2 предметов на 1.

Такая процедура диверсификации была разработана после тщательного анализа процесса поиска решения некоторых задач. После длительного поиска без улучшения целевой функции, то есть выполнения критерия применимости диверсификации, возникала следующая ситуация: в одном из контейнеров скапливалось 5 самых легких предметов, при этом он обычно становился самым переполненным. В дальнейшем поиске эти 5 предметов не подвергались обмену, потому что обмен по первой окрестности решений, когда какой-нибудь легкий предмет в этом контейнере перемещался бы в другой контейнер, приводил к слишком плохому значению оценочной функции – другой контейнер становился значительно более переполненным, чем до обмена. А обмен предметов по второй окрестности решений не применялся к этому контейнеру потому, что в нем скапливались самые лёгкие предметы, и замена предмета на более тяжелый также была невыгодна с точки зрения улучшения оценочной функции. Следует также отметить, что самый легкий контейнер обычно содержал два самых тяжелых предмета и так же большое число итераций не выбирался в качестве контейнера с предметами, подлежащими обмену.

В некоторых случаях описанный алгоритм упаковки всё-таки иногда производил обмены предметов из самого переполненного контейнера самостоятельно, благодаря использованию списка запретов, но в большинстве случаев такого не происходило, и именно эта ситуация означала наступление состояния локального оптимума и необходимость применения процедуры диверсификации.

Аналогичный механизм диверсификации описан в исследовании [9], где именовался как «существенная диверсификация» (influential diversification) и предназначался для оптимизации расписания процессов. Его суть заключалась в перераспределении задач между устройствами,

загруженными чрезмерно малым либо чрезмерно большим количеством трудоемких операций. Сложность реализации подобного подхода возрастала пропорционально квадрату количества объектов.

Предлагаемый авторами новый способ основан на представлении контейнерно-центрированного решения [5]: он сводится лишь к нескольким операциям присвоения, и его вычислительная сложность остается неизменной величиной, не зависящей ни от общего числа элементов, ни от количества контейнеров.

Строго говоря, алгоритм линейный относительно числа предметов в контейнере. Но, поскольку речь идет всегда о точно 4 предметах (мы из контейнера с пятью предметами убрали два и добавляем один новый) и нахождение места правильной вставки нового предмета требует не более трёх операций сравнения, то такую сложность вполне допустимо считать константной. Предлагаемая авторами процедура диверсификации демонстрирует высокую продуктивность в случаях, касающихся задач-триплетов [10], поскольку именно при их решении регулярно возникают ситуации, подобные ранее описанной. Что касается иных типов задач, известных как однородные, здесь данная методика служит эффективным инструментом для кардинального изменения текущего размещения объектов в решении.

Описанная ситуация с обменом двух предметов из пяти на один из двух просто служила обоснованием способа диверсификации. Авторы полагают, что анализ решения в случае локального оптимума и действия, направленного именно на преодоление подобной ситуации полезнее, чем выбор в качестве диверсификации просто какого-то случайного существенного изменения решения.

Также возможен вариант усовершенствованного механизма диверсификации, при котором операция «замены двух объектов на один» выполняется не исключительно для самого перегруженного контейнера, а последовательно охватывает группу наиболее наполненных контейнеров. Перемещение объектов производится для ограниченного набора из b_num контейнеров, отсортированных по убыванию показателей загрузки. Значение b_num рекомендуется выбирать небольшим, не превышающим число переполненных контейнеров в текущем варианте решения. Экспериментальные исследования выбора оптимальной величины этого параметра описаны в пункте 2.3 настоящей статьи.

Запуск процедуры диверсификации осуществляется при выполнении классического критерия – достижении установленного порога div_it последовательных итераций, в ходе которых отсутствовали обновления лучшего найденного решения. Важно отметить, что этот порог должен адаптироваться к масштабам решаемой задачи. Излишнее уменьшение этого параметра приведет к преждевременному выходу из области локального оптимума еще до достижения эффективных промежуточных решений.

С другой стороны, слишком большое значение будет негативно влиять на эффективность алгоритма, увеличивать время его работы без существенного прогресса в улучшении решения. В разделе 2.2 приведены результаты эксперимента по исследованию данного параметра процедуры диверсификации.

2.1 Тестовые задачи

Для эксперимента использовались 8 тестовых наборов задач из электронной OR-библиотеки, предложенные Фолкнером [8], в каждом наборе имеется 20 случайным образом сгенерированных задач, объединенных общими характеристиками исходных данных. В таблице 1 приведены эти характеристики для каждого набора.

Табл. 1 Характеристики тестовых задач

№ тестового набора	**n – число предметов**	**C – ёмкость контейнера**	**Интервал, содержащий веса предметов**
1	120	150	[20,100]
2	250	150	[20,100]
3	500	150	[20,100]
4	1000	150	[20,100]
5	60	100	[20,50]
6	120	100	[20,50]
7	249	100	[20,50]
8	501	100	[20,50]

Данные тестовые задачи являются достаточно сложными и решения, получаемые простыми эвристическими алгоритмами в большинстве случаев далеки от оптимального. Для всех задач известны решения, полученные Фолкнером с помощью генетического алгоритма. Почти во всех случаях эти решения являются заведомо оптимальными, так как совпадают с нижней границей. В таблице 2 приведены средние значения целевой функции - число контейнеров. Задачи наборов 5-8 представляют собой так называемые «триплеты» - в оптимальном решении обычно содержится по три предмета в каждом контейнере. Для триплетов из одного набора число контейнеров в оптимальном решении точно известно и одинаково.

Табл.2. Средние значения наилучшего известного числа контейнеров в тестовых задачах

№ тестового набора	Среднее число контейнеров
1	49.15
2	101.70
3	201.20
4	400.55
5	20.00
6	40.00
7	83.00
8	167.00

2.2 Исследование частоты запуска диверсификации

В данном разделе приведены результаты вычислительного эксперимента по определению параметра процедуры диверсификации – числа итераций, в течение которого не происходило улучшения целевой функции и происходил запуск процедуры диверсификации. В таблице приведены только данные для тестовых наборов 1,4,5 и 8. Значения числа итераций принимались равными 100,300,500 и 1000. В таблице фиксировалось среднее (для 20 задач каждого из наборов) отклонение от известного наилучшего значения целевой функции. Как и ожидалось, значение 100 является недостаточным – алгоритм упаковки не успевает достичь локально оптимальных областей. Интересным результатом является тот факт, что наихудшие отклонения получены для задачи наименьшей размерности – тестового набора 5 с n=60. В дальнейшем предлагается этот параметр диверсификации принимать равным 500 при n, не превышающем 500 и 1000 иначе. Для триплетов же предлагается принимать это значение равным 1000 вне зависимости от размерности задачи – числа контейнеров.

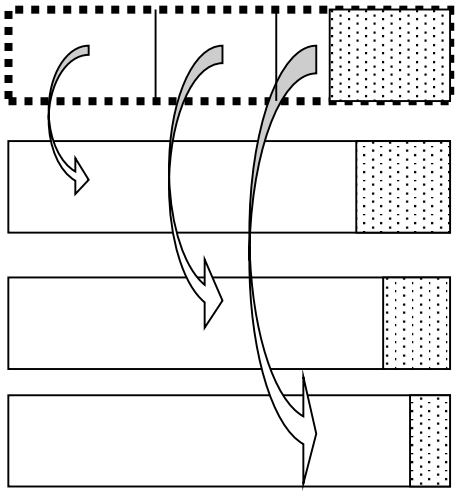
Табл. 3 Отклонение от оптимума в зависимости от частоты запуска диверсификации

№ тестового набора	Число итераций, после которых запускается диверсификация			
	100	300	500	1000
1	0,84	0,03	0	0
4	1,12	0,89	0,14	0
5	0,69	0,18	0,09	0,02
8	1,16	0,74	0	0

2.3 Исследование параметра диверсификации - числа контейнеров, подвергаемых диверсификации

В данном разделе приведём результаты вычислительного эксперимента по определению параметра b_num – числу переполненных контейнеров, к которым применялась процедура диверсификации. Опишем кратко возникновение переполненных контейнеров. Они возникали на шаге 2 алгоритма упаковки, описанного в разделе 1. Предметы из самого легкого контейнера перемещались в другие контейнеры, а сам контейнер считался «уничтоженным», тем самым целевая функция улучшалась на единицу. Алгоритм перемещения довольно простой и состоит в следующем: самый тяжелый предмет из «уничтожаемого» контейнера перемещается во второй из наименее заполненных контейнеров, второй по возрастанию веса предмет перемещается в третий из наименее заполненных контейнеров и так далее, до тех пор, пока все предметы из данного контейнера не будут перемещены.

Рис. 1 «Уничтожение» контейнера



Работа алгоритма демонстрируется на рис.1, где контейнеры расположены по возрастанию заполненности (суммарного веса сопоставленных им предметов) снизу вверх, а предметы в

«уничтожаемом» контейнере расположены по возрастанию веса слева направо. В рассматриваемых тестовых примерах даже сразу после процедуры уничтожения контейнера число переполненных контейнеров не превышало числа предметов в самом легком контейнере – обычно это число составляло 3-4 предмета. В результате работы алгоритма описанного в разделе 1, это число, как правило, сокращалось. Поэтому авторы решили рассматривать только два значения числа контейнеров, подвергаемых диверсификации – 1 или 2. В таблице 4 приведены результаты вычислительного эксперимента, демонстрирующие достижимость наилучшего известного решения при использовании различных значений числа контейнеров, подвергающихся переупаковке. Так как общий алгоритм решения задачи является недетерминированным, то каждая индивидуальная задача решалась 10 раз. Для каждой из 20 задач тестового набора было подсчитано среднее число достижения наилучшего решения при 10 запусках. Именно это число и приведено в таблице 4.

Таблица 4. Исследование параметра b_num

№ тестового набора	B_num - число контейнеров, подвергающихся диверсификации	
	1	2
1	9,73	10,0
2	8,90	9,13
3	6,09	8,51
4	3,78	6,44
5	9,64	9,87
6	8,03	9,91
7	6,49	9,73
8	7,12	9,69

Как видно из приведенных результатов во всех случаях для каждого тестового примера достигалось наилучшее известное решение. Более того, для 2 задач тестового набора 5 и 1 задачи тестового набора 7 было найдено лучшее решение, чем у автора, предложившего данные тесты [8,11]. Что касается параметра b_num, авторы предлагают в дальнейшем использовать значение 1 для тестовых наборов 1,2,5,6 и значение 2 для остальных. Если сформулировать правило безотносительно к рассмотренным в вычислительном эксперименте задачам, то предлагается подвергать диверсификации 1 контейнер при размерности не более 500 для равномерных задач и не более 250 предметов для триплетов, и использовать 2 контейнера в иных случаях.

3 Заключение

- Авторы рассмотрели задачу BPP, привели описание алгоритма локального спуска, использующего две полиномиальные окрестности соседних решений
- Было приведено описание плана упаковки в случае достижения локального оптимума, чаще всего встречающееся при решении так называемых задач-триплетов
- Авторами предложена процедура диверсификации – существенного изменения плана упаковки с целью выхода из локального оптимума
- Приведены результаты вычислительного эксперимента по определению двух параметров предлагаемой диверсификации – критерия запуска и числа преобразуемых контейнеров, определены правила по выбору указанных параметров в зависимости от характеристик конкретной задачи

Список используемых источников

1. Garey M.R. and Johnson D.S. Computers and Intractability: A Guide to the Theory of NP-Completeness. W.H. Frecman, San Francisco. 1979 –324 p.
2. Scholl A., Klein R., Jurgens C. Bison: A fast hybrid procedure for exactly solving the one-dimensional bin-packing problem. Computers Ops.Res. Vol. 24, No 7, 1997.-pp. 627-645.
3. Усманова А.Р. Вероятностные жадные эвристики для задачи упаковки в контейнеры.- Сборник докладов Оптим-2001. Первая всероссийская научно-практическая конференция по вопросам решения оптимизационных задач в промышленности.- Спб,2001.-С.141-145.
4. Кочетов Ю.А., Усманова А. Р. Вероятностный поиск с запретами для задачи упаковки в контейнеры // Иркутск: XII Байкальская международная конференция. 2001. С. 22-27.
5. Тарасова Т.Д., Усманова А.Р. Контейнерно-ориентированное представление плана упаковки в задаче BPP. - Информационные технологии интеллектуальной поддержки принятия решений (ITIDS'2019). Труды VII Всероссийской научной конференции (с приглашением зарубежных ученых): в 3 томах. 2019. С. 62-67.
6. Усманова А.Р., Валиахметова Ю.И. Особенности метода поиска с запретами для задачи упаковки. - Системная инженерия и информационные технологии. 2022. Т. 4. № 2 (9). С. 37-42.

7. Usmanova A.R., Zemlyanov A.P. The local search algorithm in polynomial neighborhoods for the linear packing problem. - CSIT'2016. Proceedings of the 18th International Workshop on Computer Science and Information Technologies. 2016. C. 138-143.

8. Falkenauer E. A hybrid Grouping Genetic Algorithm for Bin Packing. Journal of Heuristics, Vol.2, No 1,1996.- pp.5-30.

9. Gutin G. Exponential neighborhood local search for the traveling salesman problem. Comput. Oper. Res. Vol.26, 1999.- pp. 313-320.

10. Усманова А.Р., Валиахметова Ю.И. Особенности метода поиска с запретами для задачи упаковки. Системная инженерия и информационные технологии. 2022. Т. 4. № 2 (9). С. 37-42.

11. Usmanova A.R., Valiakhmetova Ju. Two approaches to the nondeterminism of simple heuristics for the BPP. Information Technologies for Intelligent Decision Making Support (in memory of Prof. N.I. Yusupova) ITIDS'2024. Proceedings of the 10th International Scientific Conference. In 2 volumes. 2024. C. 146-153.

12. Shperling S. Bin-by-bin strategy for the two-dimensional irregular bin packing problem. Nonlinear analysis and extremal problems (NLA-2024). Proceedings of the 8th International School-Seminar. Irkutsk, 2024. C.252-253.

13. Li Y., Chen M., Huo J. 2022. A hybrid adaptive large neighborhood search algorithm for the large-scale heterogeneous container loading problem. Expert Systems with Applications 189: 115909.

14. Meliani Y., Hani Ya., Lissane Elhaq S., El Mhamedi. A Tabu search based approach for the heterogeneous fleet vehicle routing problem with three-dimensional loading constraints. Applied Soft Computing. 2022. T. 126. C. 109239.

Программное обеспечение для автоматического обнаружения пожаров на изображениях лесных территорий с использованием методов машинного обучения

Батыршин А. Д.
Уфимский университет науки и технологий
Институт информатики, математики и кибернетики
Уфа, Россия
kamelotbatyr@gmail.com

Сазонова Е. Ю.
Уфимский университет науки и технологий
Институт информатики, математики и кибернетики
Уфа, Россия
ekaterina_rassadnikova@mail.ru

Сметанина О. Н.
Уфимский университет науки и технологий
Институт информатики, математики и кибернетики
Уфа, Россия
smoljushka@mail.ru

Аннотация[1]

Статья посвящена разработке программного решения для распознавания пожаров на изображения лесных территорий. Обоснована актуальность решаемой задачи, приведена постановка задачи и структура решения задачи, предложена свёрточная нейронная сети для задачи распознавания пожаров, приведены особенности сбора набора данных и обучения модели, показаны результаты применения обученный модели, представлен аспекты разработки и интерфейс разработанного программного решения.

Ключевые слова: компьютерное зрение, нейронные сети, сверточные нейронные сети, обнаружение лесных пожаров.

1. Введение

Пожары несут за собой большие последствия, например, согласно статистике [1] за 2023 год в пожарах, сгорело около 12 миллионов гектаров леса, что занимает около 40 процентов от всех потерь леса за 2023 год. По статистке из того же источника каждый год Россия теряет около 2,5 миллионов гектаров леса в пожарах, что ставит её на первое место среди всех стран мира по объёму сгоревшего леса. По данным на 12 мая 2025 года [2], на территории Российской Федерации действовало 95 лесных пожаров. Площадь, пройденная огнём, составила 521 172 гектара, что говорит о том, что в этом году ситуация не улучшится, так как пик лесных возгораний ещё впереди. Рассмотрим подробнее какие меры предпринимаются для обнаружения пожаров

На сегодняшний день существует много методов обнаружения лесных пожаров, таких как наблюдение со смотровых вышек, слежка со спутников и камер, но все из них обладают недостатками. Вышки дорого поставить в труднодоступных местах, а территория покрытия вышек не такая большая относительно других методов, спутники, напротив, могут покрывать огромные расстояния, но замечают пожар уже на более поздних стадиях. Слежка с камер схожа с наблюдением с вышек, их область мала, а устанавливать их в крупных лесах невыгодно.

Создав систему автоматического распознавания пожаров, можно обнаружить пожар камерой с дрона. В зависимости от дрона расстояние, на которое он способен улететь может варьироваться, но его можно отправить в разные места, например при помощи других методов можно определить пожароопасные области и именно их патрулировать при помощи дрона. Такую систему можно использовать при видеонаблюдении с камер, установленных на вышках, которые будут оповещать наблюдающего при обнаружении огня, чтобы тот уже лично достоверность оповещения и предпринял дальнейшие действия.

Таким образом, можно сделать вывод о том, что разработка программного решения для распознавания пожаров на изображения лесных территорий – это актуальная задача, при решении которой стоит обратить внимание на два признака:

[1] Труды XI Международной научной конференции "Информационные технологии интеллектуальной поддержки принятия решений", 11-13 ноября, Уфа-Чандигарх-Фандербейлпарк, 2025

огонь и дым, из которых дым является приоритетным признаком из-за особенностей, подразумевающихся при использовании программы.

2. Постановка задачи распознавания лесных пожаров на изображениях

Рассмотрим задачу распознавания лесных пожаров. Задача заключается в обработке изображения таким образом, чтобы установить наличие на нем дыма или огня, чтобы в заключении сделать вывод о том, присутствует ли на изображении пожар. Входными данными является файл с изображением в расширении JPG, JPEG или PNG. Предполагается, что на изображении есть вид лесных территорий. Выходными данными является присвоенный изображению класс и прочие данные, полученные в ходе обработки.

Приложение должно выполнять следующие функции: функция выбора собственного изображения; возможность запуска распознавания по выбранному изображению; вывод полученных результатов.

Формальная постановка задачи представлена на Рисунке 1.



Рис. 1. Формальная постановка задачи

Структура решения задачи автоматического обнаружения пожаров на изображениях лесных территорий состоит из следующих компонентов: предварительная обработка изображения; распознавание пожара по изображению; обработка результатов и его вывод (Рис.2).

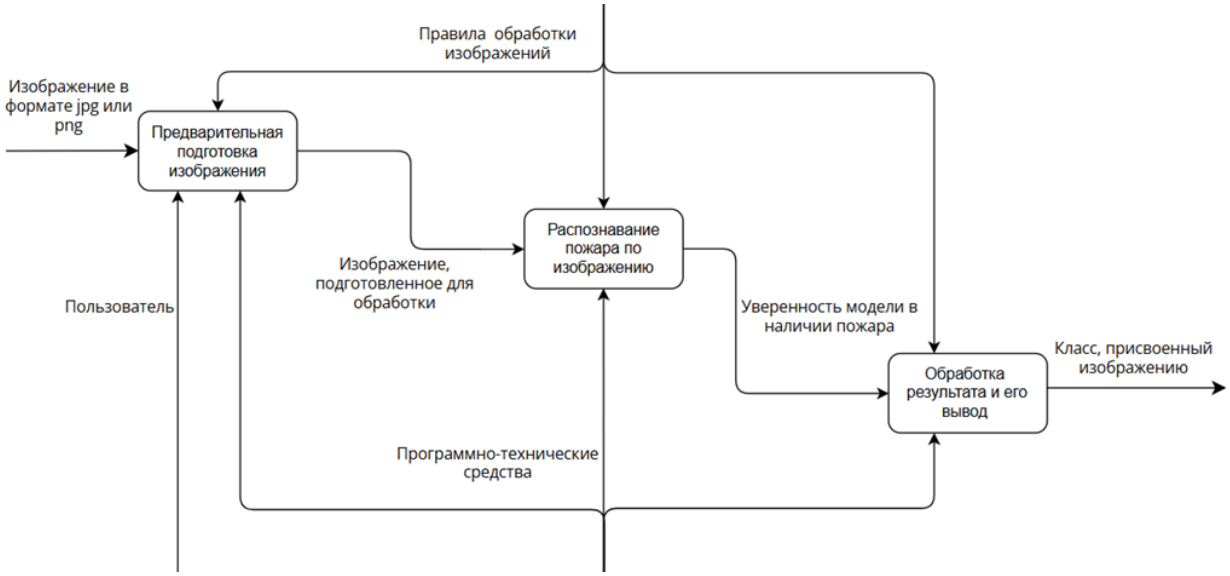


Рис. 2. Структура решения задачи

Входными данными является изображение лесных территорий в формате jpg. Задача заключается в обработке изображения таким образом, чтобы установить наличие на нем дыма или огня, чтобы в заключении сделать вывод о том, присутствует ли на изображении пожар. Выходными данными является вывод о том присутствует ли на изображении огонь и(или) дым, свидетельствующие о пожаре на изображении, а также уверенность модели в ответе.

Рассмотрим подробнее структуру решения задачи. Сначала пользователь загружает в приложении входные данные, jpg - изображение. Затем это изображении проходит предобработку, где оно изменяется в размере до размера 250 на 250 пикселей, а также повышается его контрастность для лучшего выделения признаков. После этого, в соответствии с полученными в обучении модели весами, модель производит оценку признаков, также выявленных ею при обучении, на обработанном изображении. По итогам оценки изображению присваивается класс, который выводится на экран для пользователя, вместе с уверенностью модели в её ответе.

Особенности распознавания пожаров и получения класса изображения будут рассмотрены ниже, так как это напрямую зависит от выбранного метода и реализации модели для распознавания.

3. Сравнительный анализ сервисов для классификации пожаров

В открытом доступе нет существующих программных решений специализирующихся на распознавании пожаров в качестве конкурентов разрабатываемому программному решению будут рассмотрены сервисы позволяющие обучить модель, не создавая её вручную.

Для анализа были выбраны следующие популярные платформы: Google Vertex AI, Teachable Machine, Roboflow (YOLOv8), Microsoft Azure. Рассмотрим каждую платформу подробнее:

- Google Vertex AI (AutoML Vision). Преимуществами являются интеграция с Google Cloud, что позволяет комфортно работать с облачным хранилищем; работа на удаленном сервере, что уменьшает требования к производительности ПК пользователя; поддержка решения для различных классов задач: классификация, детекция, OCR. К недостаткам можно отнести следующее: для постоянного использования требуется аккаунт Google Cloud с платной подпиской; ограничен производительностью удаленных серверов, на что никак не сможет повлиять пользователь.

- Teachable Machine (Google). Платформа полностью бесплатна, нет никаких ограничений в использовании; обучение происходит в браузере, что позволяет пользоваться Teachable Machine даже на самых слабых устройствах; модель сохраняется в веб-формате TensorFlow.js, что позволяет пользоваться ей в сайтах. Однако, функционал программного обеспечения крайне ограничен, решается лишь задача классификации (нет детекции/сегментации); точность модели низкая в сравнении с другими вариантами, что делает это решение подходящим лишь для простых задач.

- Roboflow. Данная платформа поддерживает различные архитектуры YOLOv8, Faster R-CNN; возможно изменять множество параметров, что позволяет добиться высокой точности при правильном обращении с моделью; у платформы есть бесплатный тариф (до 10000 изображений и 1000 предсказаний/месяц). Однако, большое количество настроек может быть препятствием для неподготовленного пользователя; бесплатный тариф не позволяет справляться с более продвинутыми задачами, а для разблокировки больших возможностей нужен Pro-аккаунт. Модель развертывается полностью на устройстве пользователя, что требует производительности устройства для комфортной работы.

- Microsoft Azure Custom Vision. Платформа обладает простым интерфейсом, позволяющим легко разметить объекты на изображениях для задачи детекции. Модель можно экспортировать в разных форматах (TensorFlow, Docker-контейнер и др.). Наличие поддержки мобильных устройств (iOS/Android через Core ML и TensorFlow Lite), бесплатного тарифа (до 2 проектов и 10000 предсказаний/месяц бесплатно). Однако, она хуже аналогов справляется со сложными задачами, связанными с, например, небольшими объектами, или там, где объект виден неполностью; также нет поддержки сегментации (только классификация и детекция).

Рассмотрим показатели этих моделей на двух датасетах MNIST и CIFAR-10. MNIST – это датасет (60 000 тренировочных и 10 000 тестовых изображений размером 28×28 пикселей) для простейшей задачи решения распознавания рукописных цифр. На нем можно оценить справляется ли модель с базовой задачи классификации. CIFAR-10 – это датасет изображений объектов десяти классов, на котором можно оценить способность модели справляться с уже более комплексной задачей классификации, что говорит о её способности выделять более сложные признаки, что станет ключевой особенностью ПО, необходимой при распознавании лесных пожаров. Метрики моделей на классических датасетах для оценки представлены в Таблице 1.

Таблица 1. Точности моделей на датасетах MNIST и CIFAR-10

Платформа	Accuracy на MNIST	Accuracy на CIFAR-10
Google Vertex AI	99.3%	93.1%
Teachable Machine	96.8%	78.5%
Roboflow (YOLOv8)	98.9%	89.7%
Microsoft Azure	97.5%	86.3%

По результатам тестирования на MNIST все варианты показали хороший результат, что говорит о том, что каждая из альтернатив способна решить простейшую задачу классификации. На CIFAR-10 показатели моделей заметно отличаются, что говорит о том, что Teachable Machine уже не так хорошо справляется с более сложными изображениями. Google Vertex AI и Roboflow показали лучшие результаты из приведенных моделей.

Рассмотрев возможные альтернативы для того, чтобы создать модель распознавания лесных пожаров, можно сделать вывод о том, что, несмотря на перечисленные преимущества, не все альтернативы способны в полной мере справиться с задачей в полной мере. Teachable machine хорошо подходит для простых задач, но для задачи распознавания пожаров, с учетом встречаемых проблем, это программное решение не подойдет. Другие представленных выше альтернатив имеют возможности для решения задачи, но у всех них есть свои слабые стороны, такие как недостаточно обширный инструментарий для редактирования

модели или необходимость платной подписки для получения доступа к полному списку возможностей.

Следовательно, разработка приложения для распознавания лесных пожаров является актуальной задачей.

На основе проведенного анализа подходов при распознавании пожаров на изображениях (классические методы компьютерного зрения; классические методы машинного обучения; глубокое обучение) было решено использовать сверточную нейронную сеть (CNN).

4. Архитектура CNN для решения задачи обнаружения лесных пожаров

Задача выделения областей возгорания решается с использованием свёрточной нейронной сети (CNN), разработанной для детектирования пожаров на изображениях. На вход модели подаётся изображение в формате RGB, которое затем масштабируется под размер 250×250 пикселей.

Модель состоит из нескольких блоков, каждый из которых в свою очередь состоит из нескольких слоев, а именно Conv2d, BatchNorm2d, ReLU, MaxPool2d, Dropout, Flatten, Linear. Рассмотрим подробнее каждый слой:

- Conv2d – сверточный слои, которые проходятся по изображению квадратом заданного размера, находя в нем признаки.
- BatchNorm – нормализует выводы сверточных слоев, что делает обучение модели более стабильным.
- ReLU – слои функции активации, выборочно пропуская входные данные
- MaxPool2d – выбирает самые важные, максимальные признаки, выделенные на предыдущих слоях.
- Dropout- обеспечивает защиту от переобучения тем, что иногда обнуляет какое количество признаков.
- Финальный слой, состоящий из Flatten, Linear и ReLU, где новые слои выполняют следующий задачи: Flatten превращает входные данные в вектор, Liniear помогает сначала выделить основные признаки, а затем выдает результат классификации, ReLU отвечает за нелинейность в модели, чего она добивается при помощи того, что не пропускает значения признаков меньше нуля, что делает обучение модели нелинейным, предоставляя возможность обучить её сложным признакам (Рис. 3).

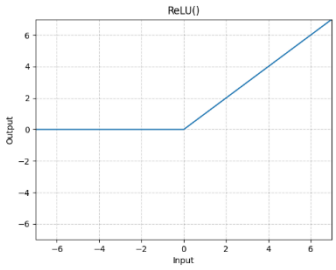


Рис. 3. График отношения входных и выходных данных функции ReLU

5. Анализ работы с CNN при решении задачи обнаружения лесных пожаров

При создании CNN модели следует учесть несколько особенностей, а именно то, на каких изображениях будет происходить обучение модели, а также то, что успешность финальной модели может сильно зависеть от параметров и количества эпох обучения [6], что значит, что для получения оптимального результата потребуется обучить несколько вариантов модели и постепенно выбрать оптимальные параметры модели. Рассмотрим особенности создания набора данных и обучения модели.

5.1. Особенности создания набора данных

При обучении модели требуется обратить особенное внимание на датасет, так как именно от его качества будет зависеть качество самой модели. Так как модель при реальном использовании будет работать с изображениями с дронов или камер на высоте, то это обязательно отразить в датасете, чтобы она была научена работать с подобными изображениями (Рис. 4).



Рис. 4. Пример оптимального изображения для датасета

Определять наличие пожара можно по двум признакам: огонь и дым, а также что у каждый из этих признаков требует внимания и специально подобранных изображений в датасете, которые помогут модели научиться отличать этот признак от изображений с похожими явлениями, чтобы это не привело к ложным ответам. Соответственно, датасет

должен включать в себя изображения с огнем и с дымом, для определения главных признаков наличия пожара, а также изображения леса, на которых не будет пламени или дыма. В группе изображений без признаков пожара должны быть снимки с проблематичными признаками, такими как закат или туман, чтобы модель на них с большей точностью научилась определять отсутствие пожаров в таких случаях (Рис. 5).



Рис. 5. Пример изображения с туманом, похожим на дым от пожара

В группе изображений с пожаром должны быть изображения, которые тоже будут отражать проблемные случаи в распознавании пожаров, такие как полное отсутствие огня на изображении, такие снимки, где пожар находится далеко, а дым занимает лишь небольшую часть изображения (Рис. 6).



Рис. 6. Пример изображения с дымом, занимающим небольшую часть изображения

В процессе разработки было рассмотрено более пятнадцати тысяч изображений из разных датасетов, найденных в Интернете, но далеко не все из них попали в финальный датасет. Финальный датасет имеет следующую структуру: Train выборка – 3500 изображений (2000 – с пожаров, 1500 – без); Validation выборка – 500 изображений (250 на каждый класс изображений). Итак, создав датасет, изображения следует предобработать для того, чтобы модели было проще их обработать. Размер каждого изображения был изменен до 250:250, а также было применено увеличение контраста, так как дым и особенно огонь отличается по свету.

5.2. Особенности обучения модели

Во время разработки модели были обучены разные варианты модели, в которых были использованы разные параметры и наборы изображений для датасета. Рассмотрим ключевые моменты, которые были выяснены во время разработки, а также решения, к которым они привели.

В начале разработки была рассмотрена идея разделить модель на две, где одна распознавала бы огонь, а другая дым, как два главных признака пожара. Изображения по этой задумке бы проходили сначала оценку на наличия огня, а затем на наличие дыма, а потом выдавалась бы общая вероятность пожара. От этой идеи пришлось отказаться, так ещё на ранних этапах тестирования стало ясно, что если рассчитывать финальную оценку, как отношение оценок от двух моделей, то при работе на изображениях с высоты огонь видно далеко не всегда, а главным признаком лесного пожара чаще всего является именно дым, то среднюю оценку результат по огню часто тянет вниз, если на изображении пожар, где не видно огня или его мало, а то это большинство изображений, а если пытаться рассматривать эти оценки отдельно, то получается, что модель для огня в довольно большой части случаев бесполезна, а модель реагирующая только на дым будет менее эффективна, в сравнении с моделью обученной на изображениях, содержащих все признаки пожаров. Так как CNN способна на выделение сложных признаков было принято решение не разделять решение на две модели, а обучить одну более сложную модель на датасете, содержащим все типы изображений: с пожарами и без, так как это приведет к оптимальному решению задачи.

Первые итерации модели были обучены на максимально большом датасете с разнообразием изображений из разных датасетов, найденных в интернете, но результаты такой модели не были высокими (не больше 0,75 на валидационной выборке), что привело к пересмотру выборки изображений. Выяснилось, что один из выбранных датасетов содержит большое количество повторяющихся изображений, а также изображений, где на изображение леса исскуственно добавлен дым, который выглядит неестественно и не отражает реальные изображения, с которыми столкнется модель, что портит качество модели, снижая её точность.

После удаления неудачных примеров результаты нового варианта модели, без изменения других параметров, повысились на 10% процентов на валидационной выборке.

В дальнейших итерациях модели датасет был расширен изображении в связи с тем, что проблемные случаи, описанными выше, занимали примерно пятьдесят датасета для обучения и больше

семидесяти процентов валидационной выборки, что приводило к тому, что полученная модель отлично справлялась с определением класса изображения на таких изображениях, но могла ошибаться на снимках леса, на которых нет тумана, закатов и других факторов, которые должны усложнять распознавание. Так как подразумевается, что обученная модель будет использоваться для патрулирования леса, она, с существенно большей вероятностью будет обрабатывать именно изображения, не содержащие осложняющих факторов, поэтому было принято решение расширить датасет с целью увеличить количество изображений без проблемных случаев для того, чтобы повысить её точность на таких изображениях. Финальный датасет содержит примерно семьдесят пять процентов обычных изображений и двадцать пять процентов проблемных изображений и в тренировочной и валидационной выборках.

6. Программное обеспечение задачи обнаружения лесных пожаров

Рассмотрим в пошаговую последовательность работы приложения с интерфейсом для демонстрации работы готовой модели:

- Загрузка и проверка входного изображения. На входе дан файл изображения, выбранный пользователем. Необходимо проверить корректность формата и подготовить изображение к обработке, а именно должно быть изображение в формате PNG/JPG, пригодное для дальнейшего анализа.

- Предварительная обработка изображения. На входе дано изображение в формате PNG/JPG. Необходимо преобразовать изображение в тензор для подачи в нейросеть. В результате предобработки получаем нормализованный тензор размера [1, 3, H, W] (например, 250×250).

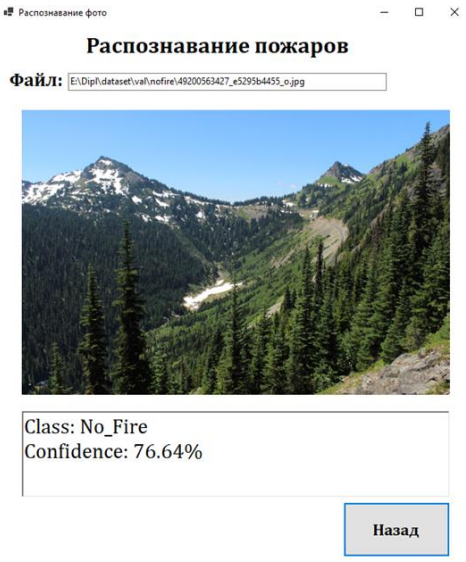
- Классификация изображения с помощью модели PyTorch. Необхоходимо загрузить преобученную модель FireDetector (файл .pth), далее инициализировать модель и загрузить веса. В результате получаем готовую к работе модель в режиме eval(). Далее осуществляется прямой проход (forward pass). Используем предобработанный тензор изображения для получения предсказания модели. В результате получаем логиты (сырые выходы нейросети).

- Постобработка и вывод результатов. На входе имеется предсказанный класс и уверенность модели. Необходимо преобразовать выход модели в текстовый формат. В результате пользователь получает строковые метки "fire"/"nofire" и значение уверенности в процентах.

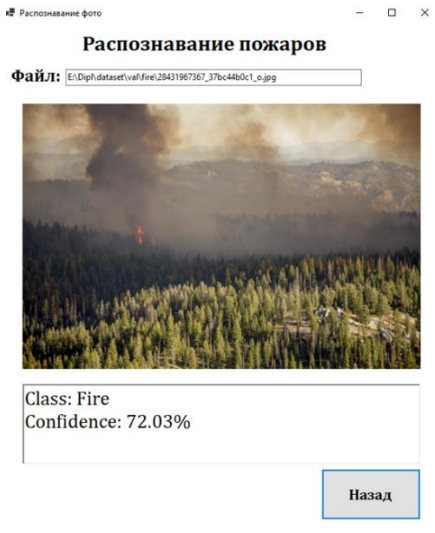
Визуализация результатов в интерфейсе. На входе дано исходное изображение и результаты классификации. Целью данного этапа отображение результата пользователю. В результате пользователь видит графическое окно с изображением и текстовой аннотацией.

При разработке программного обеспечения для обучения модели использовался язык программирования Python версии 3.11, так как на этом языке реализован фреймворк pytorch для удобной работы с библиотекой torch, предоставляющей возможности для создания сверточных нейронных сетей. Для создания интерфейса использовался Windows Forms, GUI библиотека. Код интерфейса реализован на языке программирования C#. Код распознавания написан на языке программирования Python.

Интерфейс приложения показан на рисунке 7.



а) Окно с результатом, вывод ответа «no_fire»



б) Окно с результатом, вывод ответа «fire»

Рис. 7. Интерфейс пользователя

На Рисунке 8 представлена диаграмма классов разработанного программного обеспечения, на которой показаны классы, используемые для реализации интерфейса: MainForm и ResultsForm, FireDetectorTrainer и FireDetector, отвечающие за обучение модели и за взаимодействие с уже обученной моделью соответсвенно.

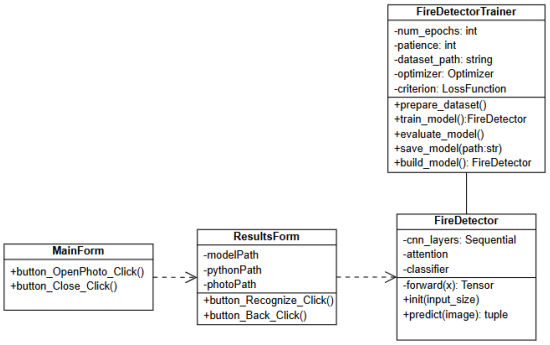


Рис. 8. Диаграмма классов

Класс FireDetector олицетворяет структуру модели, со всеми её слоями и параметрами для каждого из них, а также имеющий метод forward, реализующий механизм внимания для отслеживания важных признаков. Реализована функция prepare_dataset и train. prepare_dataset отвечает за предобработку датасета, необходимую для обучения, train же обучает модель заданное количество эпох, при этом отслеживая улучшение метрик.

Класс MainForm отвечает за работу основного окна, которое открывается при запуске приложения. Разработанные методы: button_OpenPhoto_Click, вызываемый при нажатии кнопки «Выбрать фото», который открывает проводник, позволяя выбрать фото для распознавания; button_Close_Click, закрывающий форму, при нажатии на соответствующую кнопку

Класс ResultsForm отвечает за работу окна с выбранной пользователем фотографией, которое открывается при выборе фото в прошлой форме. Разработанные методы: button_Recognize_Click, вызываемый при нажатии кнопки «Выбрать фото», который открывает проводник, позволяя выбрать фото для распознавания; button_Close_Click, закрывающий форму, при нажатии на соответствующую кнопку

Отдельно реализована функция use_model на языке python, которая позволяет вызвать распознавание изображения моделью, эта функция вызывается через отдельны python скрипт, вызываемый кодом при нажатии на кнопку «распознать» в приложении

7. Анализ эффективности предложенного решения

Эксперимент проводился на ЭВМ со следующими характеристиками: операционная система Windows 10 x64; интерпретатор Python версии 3.11; процессор Intel Core i5-12400f 2.5 GHz; количество ядер процессора: 6; ОЗУ объёмом 32 Гб.

Эксперимент проводился для трех разных выборок изображений, каждая из которых создана показать результаты модели в уникальных условиях:

- Изображения в идеальных условиях (изображения, подобные датасету для обучения, без сложных для распознавания ситуаций) – 1000 (пожар – 700, не пожар – 300);
- Проблемные изображения (выборка состоящая исключительно из изображений с проблемными условиями для распознавания, такими как туман, осенний лес и другое) – 100;
- Изображения, сделанные с земли (группа изображений, сделанных с земли, созданная проверить использование созданной программы в нестандартных для неё условиях) – 50;

Рассмотрим результаты модели на каждой из экспериментов:

Таблица 2. Результаты эксперимента

Вид эксперимента	Коли-чество изображе-ний	Коли-чество неверных ответов	Точность
Изображения в идеальных условиях	1000	74	92.6%
Проблемные изображения	100	28	72%
Изображения, сделанные с земли	50	16	68%

Проведенное тестирования разработанной модели показало, что точность: 92,6% - для изображений, схожих с теми, на которых обучалась модель, 72% - для выборки с исключительно проблемными случаями, 68% - для изображений, сделанных на земле.

Заключение

В работе представлена и экспериментально обосновано программное обеспечение для автоматической классификации лесных пожаров с использованием сверточной нейронной сети.

Основные результаты работы:

- Проведён анализ исследований в области автоматизированного распознавания лесных пожаров и обзор программных решений для создания моделей автоматизированного распознавания лесных пожаров, который показал, что в этой области отсутствует программные рещения для автоматического распознавания пожаров, обладающее высокой точностью и не требующее ежемесячной оплаты.

- Проведён обзор существующих методов решения задачи автоматизированного распознавания лесных пожаров, по итогам которого для решения задачи был выбран метод создания сверточной нейронной сети.

- Разработано математическое обеспечение для автоматизированного распознавания лесных пожаров, в основу которого была положена модель CNN, а также были определены функция ошибок, алгоритм обучения модели и параметры её слоев.

- Разработано программное обеспечение для автоматизированного распознавания лесных пожаров. Для реализации обучения и взаимодействия с моделью в качестве языка программирования был выбран язык программирования – Python; для разработки приложения использовалась среда разработки Visual Studio, также были использованы следующие библиотеки: Numpy, Matplotlib, Torch, Torchvision. Для создания интерфейса был использован язык программирования – C#, а также библиотека – Windows Forms.

- Проведено тестирование программного обеспечения и оценка его эффективности: точность: 92,6% - для изображений, схожих с теми, на которых обучалась модель, 72% - для выборки с исключительно проблемными случаями, 68% - для изображений, сделанных на земле.

Список используемых источников

1. МЧС России. Ежегодный отчет о пожарах и их последствиях на территории Российской Федерации за 2023 год [Электронный ресурс]. – М.: МЧС России, 2024. – Режим доступа: https://www.mchs.gov.ru/dokumenty/otchety (дата обращения: 25.04.2025).

2. Росстат. Пожары и противопожарная защита в России: Статистический сборник. – М.: Росстат, 2023. – 156 с. – ISBN 978-5-89476-XXX-X.

3. Гудфеллоу Я., Бенджио Ю., Курвилль А. Глубокое обучение – М.: ДМК Пресс, 2017. – 652 с. ISBN 978-5-97060-554-4.

4. Дуда Р., Харт П., Сторк Д. Распознавание образов. – М.: Вильямс, 2021. – 752 с. – ISBN 978-5-8459-2097-1.

5. Стивенс Э. PyTorch: глубокое обучение на Python. – СПб.: Питер, 2021. – 320 с. – ISBN 978-5-4461-1589-8.

6. Шолле Ф. Глубокое обучение на Python – СПб.: Питер, 2022. – 400 с. ISBN 978-5-4461-1668-8.

7. Muhammad K., Ahmad J., Lv Z. Efficient Deep CNN-Based Fire Detection System // IEEE Access. – 2019. – Vol. 7. DOI: 10.1109/ACCESS.2019.2895760.

8. Frizzi S., Kaabi R., Bouchouicha M. Convolutional Neural Network for Fire Detection – 2018. – Vol. 65, no. 1 DOI: 10.1109/TIE.2017.2725581.

9. Krizhevsky A., Sutskever I., Hinton G. ImageNet Classification with Deep Convolutional Neural Networks // Advances in Neural Information Processing Systems. – 2012. – Vol. 25. – P. 1097–1105.

10. Shorten C., Khoshgoftaar T.M. A Survey on Image Data Augmentation for Deep Learning // Journal of Big Data. – 2019. – Vol. 6, № 1. – P. 1–48. – DOI: 10.1186/s40537-019-0197-0.

11. Simonyan K., Zisserman A. Very Deep Convolutional Networks for Large-Scale Image Recognition // arXiv preprint. – 2014. – arXiv:1409.1556.

Архитектура серверной части медицинского информационного сервиса для взаимодействия врача и пациента

Янчиев Д. В.
Уфимский университет науки и технологий
Институт информатики, математики и робототехники
Уфа, Россия
e-mail: danilayanchiev@yandex.ru

Шакирзянов А. А.
Уфимский университет науки и технологий
Институт информатики, математики и робототехники
Уфа, Россия
e-mail: aydar102@gmail.com

Ермолов И. О.
Уфимский университет науки и технологий
Институт информатики, математики и робототехники
Уфа, Россия
e-mail: ermolovivan2018@gmail.com

Макаркин А. А.
Уфимский университет науки и технологий
Институт информатики, математики и робототехники
Уфа, Россия
e-mail: arseniy-makarkin1@mail.ru

Аннотация[1]

В статье рассматривается опыт разработки серверной части медицинского информационного сервиса, предназначенного для поддержки взаимодействия между врачами и пациентами. Система ориентирована на автоматизацию процессов формирования медицинских опросов, их назначения, планирования и последующего прохождения пациентами. Основное внимание сосредоточено на архитектуре серверного приложения, используемом технологическом стеке и реализации REST API. Также кратко рассматриваются подходы к тестированию и обеспечению корректности работы серверной логики. Материал носит обзорный характер и отражает практический опыт разработки распределённого программного решения, ориентированного на одновременную работу нескольких клиентских приложений.

Ключевые слова: медицинские информационные системы, серверная архитектура, backend-приложения, FastAPI, REST API, PostgreSQL, SQLAlchemy, аутентификация и авторизация, клиент-серверное взаимодействие, тестирование программного обеспечения.

1. Введение

Развитие цифровых медицинских сервисов в последние годы привело к активному внедрению информационных систем, ориентированных на удалённое взаимодействие между врачами и пациентами. Такие системы используются не только в рамках телемедицины, но и для регулярного анкетирования, мониторинга состояния пациентов, контроля соблюдения рекомендаций и накопления структурированных медицинских данных. Практика показывает, что востребованность подобных решений возрастает по мере увеличения числа пользователей мобильных и веб-приложений в сфере здравоохранения.

В подобных системах ключевую роль играет серверная часть, поскольку именно она отвечает за обработку пользовательских сценариев, управление данными, контроль доступа и интеграцию различных клиентских приложений. Ошибки или архитектурные просчёты на уровне бэкенда напрямую влияют на устойчивость всей системы и возможность её дальнейшего развития.

В рамках рассматриваемого проекта разрабатывался медицинский информационный сервис, предназначенный для практического использования врачами и пациентами. Целью проекта являлось создание единой серверной платформы, обеспечивающей централизованную реализацию бизнес-логики и доступ к данным для нескольких клиентских приложений. В статье представлен обзор архитектуры серверной части, используемых технологий и инженерных решений, принятых в ходе разработки.

2. Назначение и функциональные возможности системы

Разрабатываемая система предназначена для организации взаимодействия между врачами и

[1] Труды XI Международной научной конференции "Информационные технологии интеллектуальной поддержки принятия решений", 11-13 ноября, Уфа-Чандигарх-Фандербейлпарк, 2025

пациентами в рамках выполнения медицинских опросов и анкет. Основная функциональная нагрузка системы связана с автоматизацией процессов, которые в традиционной практике часто выполняются вручную и требуют значительных временных затрат со стороны медицинского персонала.

Со стороны врача система предоставляет инструменты для создания медицинских вопросов и формирования опросов различной структуры. Реализована возможность повторного использования ранее созданных вопросов, что позволяет формировать опросы на основе уже накопленного набора элементов. Дополнительно врач может задавать параметры прохождения опроса, включая временные интервалы, порядок выполнения и привязку к конкретным пациентам. Функциональность пациента ориентирована на получение назначенных опросов и их последовательное прохождение. Пациент получает уведомления о необходимости выполнения заданий, после чего может выбрать соответствующий опрос и ввести ответы через клиентское приложение. Все действия пациента фиксируются серверной частью системы, что обеспечивает целостность данных и позволяет врачу получать актуальную информацию о ходе выполнения назначенных опросов. Таким образом, система реализует полный цикл работы с медицинскими опросами — от их создания и планирования до получения и хранения результатов.

3. Общая архитектура серверного приложения

Серверная часть системы реализована в виде REST API и выступает центральным компонентом всей архитектуры приложения. Все клиентские приложения, независимо от типа и платформы, взаимодействуют с системой исключительно через сервер, что позволяет централизовать обработку данных, контроль доступа и реализацию бизнес-логики. Такой подход является типовым для современных распределённых систем и хорошо масштабируется при росте числа пользователей и функциональных возможностей.

Архитектура серверного приложения спроектирована с учётом принципа разделения ответственности между основными логическими слоями. На уровне API осуществляется маршрутизация HTTP-запросов, первичная валидация входных данных и формирование ответов клиентам. Для описания структуры запросов и ответов используются схемы данных, что обеспечивает строгую типизацию и единообразие интерфейсов взаимодействия.

Важной особенностью архитектуры является логическое разделение API по ролям пользователей. В проекте реализованы отдельные группы эндпоинтов для публичных операций, пациентов, врачей и администраторов. Такое разделение позволяет чётко определить границы доступа и упростить сопровождение системы, так как каждая группа эндпоинтов отвечает за строго определённый набор сценариев.

Уровень бизнес-логики реализует основные сценарии работы системы. К ним относятся управление пользователями, обработка медицинских опросов, формирование связей между врачами и пациентами, планирование прохождения опросов, а также обработка результатов. Выделение бизнес-логики в отдельный слой позволило избежать дублирования кода и повысить читаемость серверного приложения.

Уровень доступа к данным инкапсулирует взаимодействие с базой данных и отвечает за выполнение операций чтения и записи. Такой подход упрощает модификацию структуры хранения данных и снижает риск возникновения ошибок при изменении логики работы системы. Использование единого слоя доступа к данным также облегчает реализацию транзакционной логики и обработку ошибок.

Архитектурная схема серверного приложения и взаимодействие его основных компонентов представлена на рисунке 1.

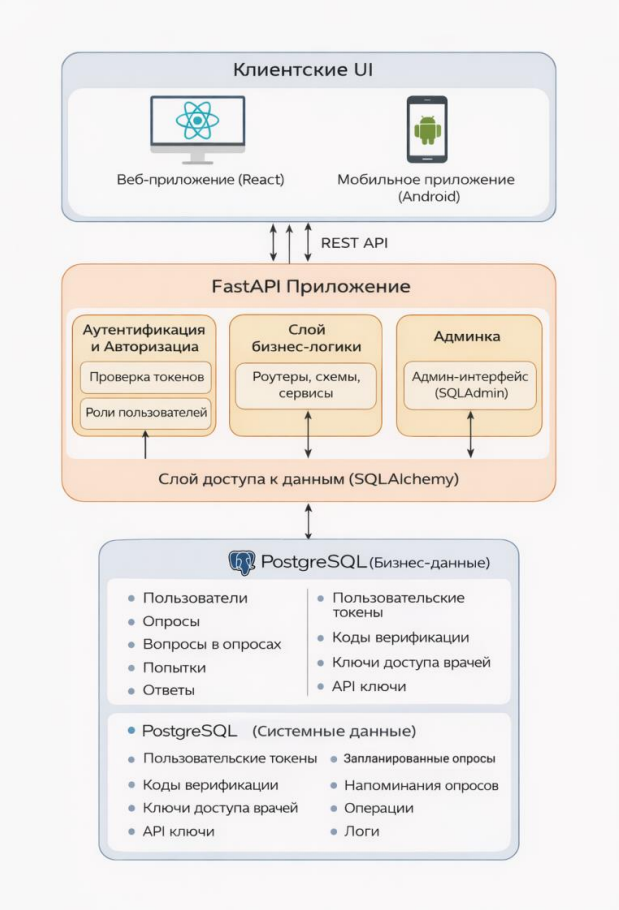


Рис. 1. Структурная схема взаимодействия компонентов серверного приложения

На схеме отражены клиентские приложения, единый REST API, основные модули FastAPI-приложения, а также слой доступа к данным и системы хранения информации.

4. Выбор технологий и реализация серверной части

Серверное приложение реализовано на языке Python с использованием фреймворка FastAPI. Выбор данного фреймворка был обусловлен рядом практических факторов, включая высокую производительность, поддержку асинхронной обработки запросов и строгую типизацию данных. FastAPI ориентирован на построение REST API с чётко определёнными контрактами взаимодействия, что является важным требованием при разработке серверной части, обслуживающей несколько клиентских приложений [2,10].

Архитектура backend-приложения организована по модульному принципу с разделением кода на логические компоненты. В проекте выделяются модули маршрутизации (routers), схемы данных (schemas), модели базы данных (models) и сервисный слой (services). Такое разделение позволяет снизить связанность компонентов, упростить сопровождение кода и обеспечить возможность его поэтапного расширения. Роутеры отвечают за описание API-эндпоинтов и приём HTTP-запросов, схемы данных используются для валидации и сериализации информации, модели описывают структуру данных в базе, а сервисный слой инкапсулирует бизнес-логику.

Важную роль в реализации серверного приложения играет механизм dependency injection, предоставляемый FastAPI. Он используется для централизованного подключения к базе данных, обработки аутентификации и авторизации, а также передачи общих зависимостей в обработчики запросов. Такой подход позволяет минимизировать дублирование кода и повысить прозрачность архитектуры приложения [6].

Аутентификация пользователей реализована на основе пользовательских токенов и API-ключей. Дополнительно применяется проверка ролей пользователей, что обеспечивает разграничение доступа к функциональности системы и предотвращает выполнение несанкционированных операций. Подобная модель доступа широко используется в распределённых системах и соответствует рекомендациям по обеспечению безопасности REST API [3,8].

Для хранения данных используется система управления базами данных PostgreSQL. Выбор реляционной СУБД обусловлен необходимостью работы со связанными сущностями и требованиями к целостности данных. В рамках проекта логически выделяются бизнес-данные и системные данные, что позволяет рассматривать возможность их разделения на уровне хранения. Подобный подход упрощает масштабирование системы, резервное копирование и администрирование данных [1,4,5].

Взаимодействие с базой данных реализовано с использованием ORM-библиотеки SQLAlchemy. Применение ORM позволило описывать основные сущности системы в виде объектных моделей и абстрагироваться от конкретных SQL-запросов. Это повышает читаемость кода и снижает вероятность ошибок при работе с данными. Кроме того, использование ORM облегчает сопровождение проекта и его развитие при изменении структуры базы данных.

Для административного управления данными используется инструмент SQLAdmin, предоставляющий веб-интерфейс для работы с сущностями системы. Административная панель позволяет выполнять операции управления пользователями, опросами и вспомогательными данными без прямого доступа к базе данных, что повышает безопасность и удобство эксплуатации системы.

5. Взаимодействие серверного приложения с клиентскими системами

Серверная часть системы используется несколькими клиентскими приложениями, включая веб-приложение и мобильное приложение для платформы Android. Все клиентские решения взаимодействуют с сервером через единый REST API, что соответствует распространённым архитектурным подходам к построению распределённых информационных систем [9,10]. Использование единого интерфейса взаимодействия позволяет централизовать контроль доступа, обработку данных и реализацию бизнес-логики. Ключевой особенностью реализованной архитектуры является разделение доступа не по типу клиентского приложения, а по ролям пользователей. В системе выделяются роли пациента, врача и администратора, каждая из которых имеет доступ к строго определённому набору эндпоинтов. Такой подход позволяет использовать одни и те же клиентские приложения для различных сценариев работы и соответствует современным принципам проектирования REST API [2].

На уровне API серверное приложение логически разделено на несколько групп эндпоинтов, включая публичные эндпоинты общего назначения, а также эндпоинты для пациентов, врачей и администраторов. Данное разделение упрощает сопровождение системы и повышает читаемость кода, поскольку каждая группа эндпоинтов отвечает за строго определённую область функциональности [9].

Каждый запрос от клиентского приложения проходит через слой middleware FastAPI, где выполняется обработка CORS-политик, проверка пользовательских токенов, API-ключей и ролей доступа. Такой подход обеспечивает единообразную обработку запросов и повышает уровень безопасности системы, что особенно важно для медицинских информационных сервисов [3,6,8].

После прохождения слоя аутентификации и авторизации запрос передаётся в соответствующий модуль бизнес-логики. Здесь осуществляется

обработка пользовательского сценария, включая проверку допустимости операции, работу с доменными сущностями и формирование ответа клиенту. Централизация бизнес-логики на сервере позволяет снизить связанность клиентских приложений и упростить их развитие [7]. Использование строгих контрактов взаимодействия и схем данных позволяет обеспечить согласованность обмена информацией между клиентами и сервером. Это упрощает тестирование клиент-серверного взаимодействия и снижает вероятность ошибок при изменении интерфейсов API [10].

С точки зрения масштабирования выбранная модель взаимодействия позволяет без существенных изменений серверной части подключать новые клиентские приложения или внешние сервисы. Такой подход соответствует современным требованиям к построению отказоустойчивых и расширяемых распределённых систем [1,4,5]. Практика эксплуатации показала, что использование единого REST API и разделение доступа по ролям пользователей создаёт устойчивую архитектурную основу для дальнейшего развития системы и внедрения новых функциональных возможностей.

6. Подходы к тестированию серверного API

Тестирование серверной части являлось важным этапом разработки и сопровождения проекта. Основное внимание уделялось проверке корректности работы API-эндпоинтов, устойчивости бизнес-логики и корректности обработки пользовательских сценариев при различных условиях эксплуатации [9,10].

В процессе разработки применялся комплексный подход к тестированию, включающий автоматизированные и ручные методы. Автоматизированное тестирование использовалось для проверки отдельных компонентов серверного приложения и основных сценариев взаимодействия клиентов с API. Такой подход позволял выявлять ошибки на ранних этапах разработки и снижать риск возникновения регрессий при внесении изменений в код [7].

Особое внимание уделялось корректности работы механизмов аутентификации и авторизации, а также проверке разграничения доступа по ролям пользователей. Это является критически важным аспектом для медицинских информационных систем, работающих с персональными данными [3,8].

Для проверки взаимодействия серверного приложения с базой данных использовались тесты, ориентированные на контроль целостности данных и корректности выполнения основных операций. Такой подход позволяет убедиться в правильности реализации бизнес-правил и снижает вероятность ошибок при эксплуатации системы. Дополнительно проводилось ручное тестирование взаимодействия клиентских приложений (веб-приложения и мобильного приложения) с сервером. Ручная проверка позволила оценить корректность интеграции, соответствие реализованных сценариев ожиданиям пользователей и выявить потенциальные проблемы, не покрываемые автоматизированными тестами.

В целом применяемые подходы к тестированию обеспечивают достаточный уровень надёжности серверного API и создают основу для дальнейшего развития и рефакторинга системы.

7. Практическая значимость и перспективы развития

Разработанный серверный сервис представляет практическую ценность как основа для создания и внедрения медицинских информационных систем, ориентированных на удалённое взаимодействие между врачами и пациентами. Архитектурные и технологические решения, реализованные в проекте, позволяют использовать систему в условиях реальной эксплуатации с одновременной работой различных категорий пользователей и клиентских приложений.

Практическая значимость проекта заключается в возможности его применения для автоматизации процессов анкетирования, мониторинга состояния пациентов и сбора структурированных медицинских данных. Использование единого серверного API обеспечивает централизованное управление бизнес-логикой и упрощает сопровождение системы при её развитии и масштабировании. Подобный подход может быть востребован как в частных медицинских учреждениях, так и в рамках корпоративных или ведомственных информационных систем.

Применённая модульная архитектура backend-приложения создаёт условия для поэтапного расширения функциональности без необходимости существенной переработки существующего кода. Это позволяет адаптировать систему под новые сценарии использования, включая добавление новых типов опросов, интеграцию дополнительных сервисов уведомлений и подключение внешних аналитических модулей.

С точки зрения перспектив развития, дальнейшая эволюция проекта может быть связана с внедрением механизмов интеллектуальной обработки данных, накопленных в ходе эксплуатации системы. Это включает использование методов статистического анализа и машинного обучения для выявления закономерностей в ответах пациентов и поддержки принятия решений медицинскими специалистами.

Отдельным направлением развития является повышение уровня информационной безопасности и отказоустойчивости серверного приложения. В частности, возможна реализация более строгих политик управления доступом, расширение механизмов аудита и логирования, а также внедрение средств резервирования и репликации данных. Такие меры позволят повысить надёжность

системы и обеспечить её устойчивую работу в условиях роста нагрузки и увеличения объёма обрабатываемых данных.

Заключение

В рамках данной работы был рассмотрен практический опыт проектирования и реализации серверной части медицинского информационного сервиса, ориентированного на взаимодействие между врачами и пациентами. В ходе разработки особое внимание уделялось архитектурным решениям, обеспечивающим масштабируемость, расширяемость и устойчивость серверного приложения при одновременной работе нескольких клиентских систем.
В результате выполненной работы были получены следующие основные результаты:

1. Спроектирована серверная архитектура на основе REST API с чётким разделением ответственности между уровнями маршрутизации, бизнес-логики и доступа к данным;

2. Реализовано логическое разделение API по ролям пользователей (пациент, врач, администратор), что позволило упростить контроль доступа и повысить безопасность системы;

3. Выбран и обоснован технологический стек серверной части, включающий FastAPI, PostgreSQL и SQLAlchemy, показавший свою эффективность при разработке распределённого приложения;

4. Реализована модульная структура backend-приложения с разделением на роутеры, схемы данных, модели и сервисный слой, что упростило сопровождение и развитие проекта;

5. Разработан комплексный подход к тестированию серверного API, обеспечивающий проверку корректности бизнес-логики, механизмов аутентификации и клиент-серверного взаимодействия.

Полученные результаты подтверждают целесообразность выбранных архитектурных и технологических решений и могут быть использованы при проектировании и реализации аналогичных медицинских информационных систем, а также других распределённых сервисов, ориентированных на работу с чувствительными данными и различными ролями пользователей.

Список используемых источников

1. Портнов К. В., Баранов м. П., мартынов д. Ю. Развитие науки и практики в глобально меняющемся мире в условиях рисков. – ооо" издательство алеф" конференция: развитие науки и практики в глобально меняющемся мире в условиях рисков москва, 27 декабря 2023 года организаторы: образовательный холдинг" институт развития образования и консалтинга", научно-исследовательский финансовый институт минфина россии, сумгаитский государственный университет азербайджанской республики.

2. Сермягин К. А., Пылькин а. Н. Построение эффективной архитектуры для высоконагруженных и масштабируемых сервисов на fastapi, принципы и лучшие практики. – рязанский государственный радиотехнический университет им. ВФ Уткина Конференция: новые информационные технологии в научных исследованиях Рязань, 27–29 ноября 2024 года Организаторы: Министерство науки и высшего образования РФ Рязанский государственный радиотехнический университет им. ВФ Уткина Библиометрические показатели: Входит в Ринц: да Цитирований в РИНЦ: 0 Входит в ядро Ринц: нет Цитирований из ядра РИНЦ: 0 Рецензии: нет данных Тематические Направления:.

3. Дрянкова Д. А., Халевин Т. А., Козлитин Р. А. Система аутентификации в fastapi: jwt токены и oauth2 //Вестник науки. – 2025. – Т. 1. – №. 12 (93). – С. 1264-1270.

4. Гвоздев в. Е. и др. Влияние дестабилизирующих воздействий на коммуникационные свойства холонических систем с динамическими структурами. – уфимский университет науки и технологий, уфа конференция: информационные технологии интеллектуальной поддержки принятия решений (памяти проф. Ни юсуповой) itids'2024 уфа, 12–14 ноября 2024 года организаторы: азербайджанский государственный университет нефти и промышленности, баку чандигархский университет, мохали уфимский федеральный исследовательский центр ран, институт биохимии и генетики, уфа российская ассоциация искусственного интеллекта, москва научный совет ран по методологии искусственного интеллекта и когнитивным исследованиям, башкирское отделение, уфа.

5. Гвоздев В.Е., Гузаиров М.Б., Ракипова А.С., Галимов Р.Р., Янчиев Д.В. Статистическое оценивание показателей структурной надежности мобильных систем передачи данных. Моделирование, оптимизация и информационные технологии. 2025;13(3). URL: https://moitvivt.ru/ru/journal/pdf?id=1989 DOI: 10.26102/2310-6018/2025.50.3.039

6. Халевин Т. А. Начало работы с fastapi. Простейшая система авторизации и аутентификации пользователя.

7. Макаренко Е. А., Басаргин А. А. Разработка автоматизированной системы управления заказами и запасами для интернет-магазина малого бизнеса с интеграцией сторонних API.

8. ЖУКОВ Н. П. и др. Информация и безопасность //информация и безопасность Учредители: Воронежский государственный технический университет. – 2025. – Т. 28. – №. 1. – С. 111-118.

9. Салтанов Д. С., Арсентьева Н. В. Сравнительный анализ архитектурных подходов к разработке rest api для высоконагруженных систем //Форум молодых ученых. – 2025. – №. 6 (106). – С. 329-340.

10. Худницкий А. В., Шауро Е. А. Сравнительный анализ Fastapi, Django и Flask: производительность, удобство разработки и выбор оптимального фреймворка для современных веб-приложений. – 2025.

Разработка информационно-аналитической системы распознавания представляющих опасность для человека бездомных животных в городских условиях

Сайфуллин М. В.
АО «Башкиравтодор»
Министерство транспорта и дорожного хозяйства Республики Башкортостан
Уфа, Россия
e-mail: maks.sajjfullin@gmail.com

Прокудина Е. И.
Институт информатики, математики и робототехники
Уфимский университет науки и технологий
Уфа, Россия
e-mail: preliv@gmail.com

Аннотация[1]

Разработана информационно-аналитическая система для автоматического распознавания бездомных животных, представляющих потенциальную опасность для человека в городской среде. В качестве основы для детекции объектов в реальном времени использована модификация алгоритма YOLO (You Only Look Once), адаптированная для задач распознавания образов бездомных собак и способная в режиме реального времени интегрировать данные с уличных камер видеонаблюдения, чтобы затем идентифицировать соответствующие кадры. Разработана архитектура системы, обучена модель YOLOv11m, проведена оценка качества по основным метрикам (94,5% Precision, 80% Recall).

Ключевые слова: искусственный интеллект, распознавание образов, компьютерное зрение, YOLO, бездомные животные, стаи собак, безопасность городской среды.

1. Введение

Актуальность проблемы взаимодействия бездомных животных и человека в городской среде обусловлена ростом числа упоминаний в средствах массовой информации случаев нападений на горожан стай бездомных собак, зачастую приводящих к травмам, а в некоторых случаях к летальному исходу. Современные технологии компьютерного зрения и машинного обучения позволяют автоматизировать процесс обнаружения угроз и не допустить человеческих жертв.

На сегодняшний день на территории Республики Башкортостан отсутствует концепция системы мониторинга нападений агрессивных бездомных животных. Распространение получил способ ОСВВ, известный как «отлов-стерилизация-вакцинация-возврат», однако в ряде районов была отмечена неэффективность проведения данной меры ввиду отсутствия контроля за подрядными организациями, осуществлявших программу.

Таким образом, разработка информационно-аналитической системы, которая бы на основе видеоданных в режиме реального времени позволяла распознавать бездомных животных, представляющих опасность для человека, является крайне актуальной.

2. Анализ существующих решений

Существующие решения в области видеоаналитики для городской безопасности включают методы на основе Haar-каскадов [2], SSD [3] и Faster R-CNN [4]. Однако их применение ограничено сравнительно низкой скоростью обработки и высокой долей ложных срабатываний. В результате исследования было выявлено, что алгоритм Ultralytics YOLO [5] демонстрирует лучший баланс между точностью и производительностью. Это подтверждается исследованиями в области мониторинга городской инфраструктуры [6]. Анализ существующих решений в области компьютерного зрения представлен в таблице 1.

Для проектируемой системы выбрана модель YOLO как оптимальная по соотноЫЫшению полученных метрик и скорости обработки данных.

[1] Труды XI Международной научной конференции "Информационные технологии интеллектуальной поддержки принятия решений", 11-13 ноября, Уфа-Чандигарх-Фандербейлпарк, 2025

3. Формализованное описание алгоритма работы системы

YOLO является мощным инструментом для обнаружения объектов в реальном времени и широко используется в различных областях, таких как компьютерное зрение, автоматическое управление и Таблица 1. – Сравнительный анализ различных моделей и метрик

Метрика / Модель	Precision	Recall	F1-score	IOU	Относительная скорость обработки кадров
Haar Cascades	0,70	0,65	0,67	0,60	Низкая
HOG+ SVM	0,75	0,70	0,72	0,65	
YOLO	**0,78**	**0,82**	**0,80**	**0,65**	**Высокая**
Faster R-CNN	0,85	0,75	0,80	0,70	Средняя
SSD	0,82	0,79	0,80	0,68	

безопасность. Модель использует сложную комбинацию математических операций и алгоритмов для распознавания образов, основные компоненты математического аппарата перечислены ниже:

1. Сверточная нейронная сеть (рисунок 1):

$$(f * g)(x, y) = \sum_m \sum_n f(m, n) \cdot g(x - m, y - n), \quad (1)$$

где f – входное изображение,

g – фильтр свертки,

(x, y) – координаты выходного изображения,

m и n – индексы суммирования, пробегающие по размеру фильтра.

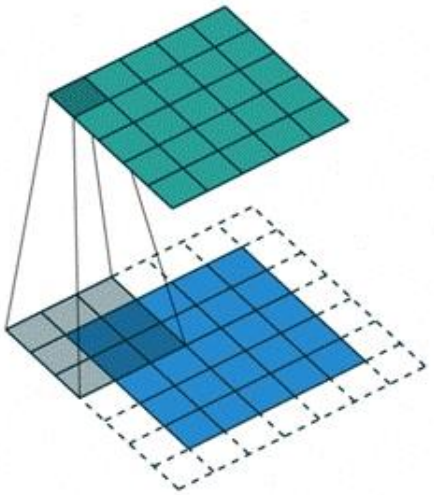


Рис. 1. Пример работы свёрточной нейронной сети

Свертка по вышеуказанному алгоритму позволяет YOLOv11 обнаруживать различные паттерны и характеристики в изображении. К ним относятся края, углы, текстуры и формы искомых в кадре объектов. Различные фильтры-матрицы могут выделять разные особенности на одном и том же образце, что позволяет использовать модель для эффективного распознавания практически любых объектов.

2. Активационная функция: ReLU (Rectified Linear Unit) (рисунок 2):

$$\mathrm{ReLU}(x) = \max(0, x), \quad (2)$$

где x – входной сигнал.

Функция ReLU также предотвращает проблему исчезающего градиента, закономерно возникающую при обучении глубоких нейронных сетей. Обновления параметров становятся невеликими, и сеть перестает учиться.

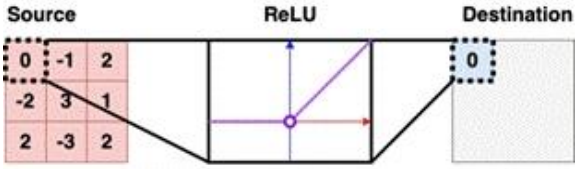


Рис. 2. Пример работы активационной функции ReLU

3. Функции потерь

3.1 *Bounding Box Loss*: *GIoU* Loss (Generalized Intersection over Union Loss) – расширение классической метрики Intersection over Union (рисунок 3)

$$GIoU = IoU - \frac{|C|(A \cup B|}{|C|}, \quad (3)$$

где *IoU* – Intersection over Union (пересечение поверх объединения) – это отношение площади пересечения областей, ограниченных предсказанными и реальными рамками к площади их объединения,

A – область, ограниченная предсказанными рамками.

B – область, ограниченная реальными рамками,

C – наименьший прямоугольник, содержащий обе компоненты ограничивающих рамок.

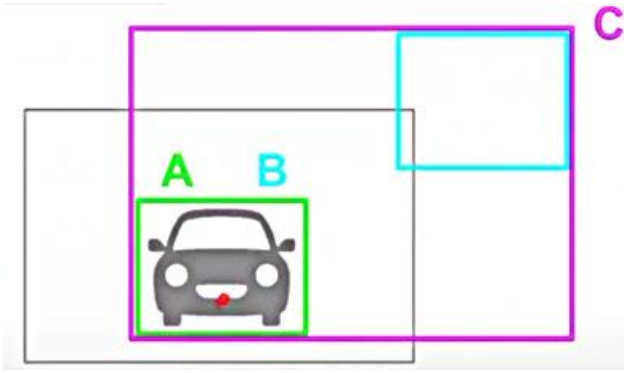


Рис. 3. Пример работы алгоритма GIoU Loss

3.2 *Confidence Loss* (Binary Cross-Entropy) – логарифмическая потеря (бинарная кросс-энтропия) –

функция потерь, используемая в машинном обучении для задач бинарной классификации (рисунок 4).

$$L_{conf} = -\sum_i (y_i \cdot \log(p_i) + (1 - y_i) \cdot \log(1 - p_i)), \quad (4)$$

где y_i – бинарная метка (0 или 1), указывающая, содержит ли *i*-й *bounding box* объект,

p_i – предсказанная вероятность того, что *i*-й *bounding box* содержит объект.

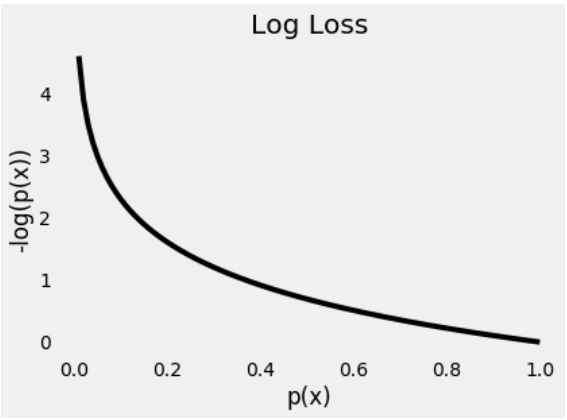


Рис. 4. Функция потерь Confidence Loss

Confidence Loss помогает YOLOv11 научиться отличать объекты от фона и уверенно предсказывать ограничивающие рамки только для реальных объектов на кадре.

3.3 *Classification Loss* (Cross-Entropy)

YOLOv11 использует кросс-энтропию для измерения ошибки классификации по формуле:

$$L_{cls} = -\sum_i \sum_c y_{ic} \cdot \log(p_{ic}), \quad (5)$$

где: y_{ic} – бинарная метка (0 или 1), указывающая, принадлежит ли *i*-й объект к классу *c*,

p_{ic} – предсказанная вероятность того, что *i*-й объект принадлежит к классу *c*.

Classification Loss измеряет ошибку классификации, то есть насколько хорошо модель предсказывает класс объекта. Он использует кросс-энтропию, которая сравнивает предсказанную вероятность принадлежности объекта к каждому классу (p_{ic}) с реальной меткой (y_{ic}), которая указывает, к какому классу принадлежит объект. *Classification Loss* помогает YOLOv11 научиться различать разные классы объектов и правильно классифицировать обнаруженные объекты.

Подробная архитектура модели представлена на рисунке 5.

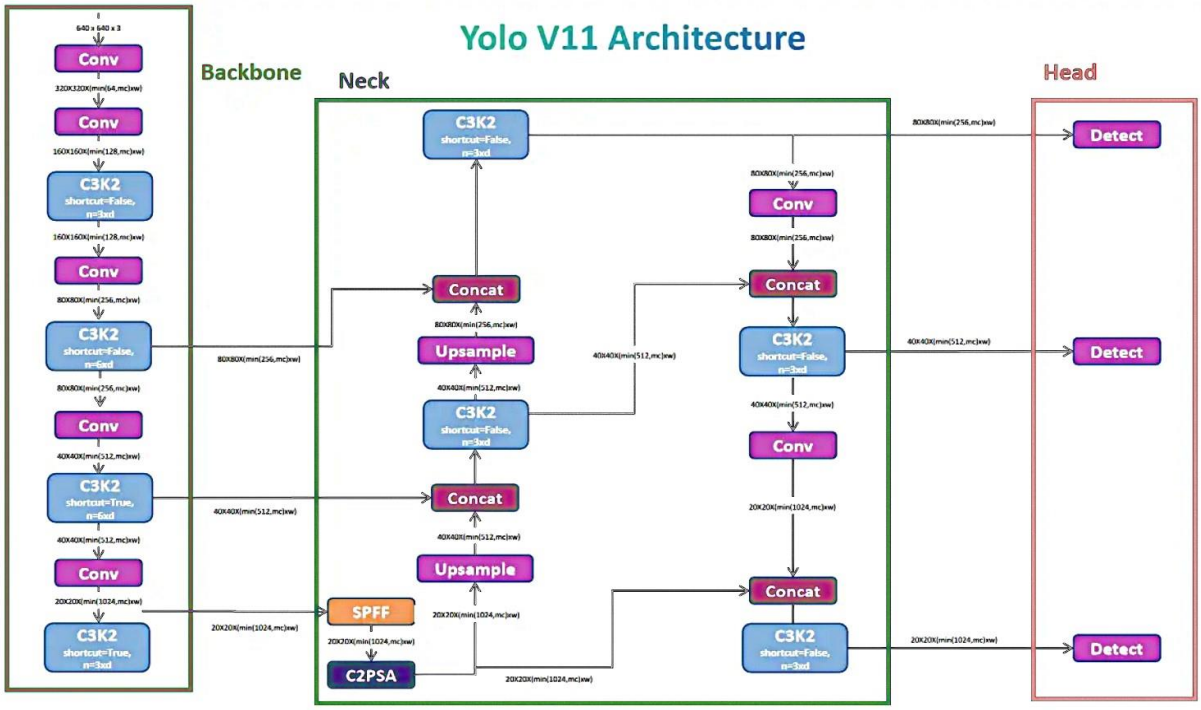


Рис. 5. Архитектура модели YOLOv11 [7]

Комплексность математического аппарата позволяет модели достигать высокой точности и скорости в обнаружении объектов. Алгоритм использует метод цикличного прямого распространения (forward pass) и обратного распространения (backward pass) через нейросеть для детекции искомых объектов на изображениях (кадрах видеопотока).

Параметры вычислительной сложности модели YOLO (11 поколение алгоритма, вышедшее в сентябре 2024 года) проведены в таблице 2.

Таблица 2. – Вычислительная сложность используемого алгоритма YOLOv11m на наборе данных COCO в задании обнаружения объектов

Модель	YOLOv11m
Размер (пикселей)	640
Ср. точн. mAPval 50-95	51,5
Скорость CPU ONNX (мс)	183,2 ± 2,0
Скорость T4 TensorRT10 (мс)	4,7 ± 0,1
Параметры (млн)	20,1
FLOPs (млрд)	68,0

4. Разработка системы распознавания

4.1. Структура системы

Система распознавания, в соответствии с рисунком 6, включает в себя четыре модуля:

1. Создание среды для обучения модели, сбор данных: настройка *PyCharm*, данные из открытых источников, IP-камеры уличного видеонаблюдения;

2. Подготовка собранных данных: разметка образов, аугментация данных;

3. Обучение модели: использование графического ускорителя;

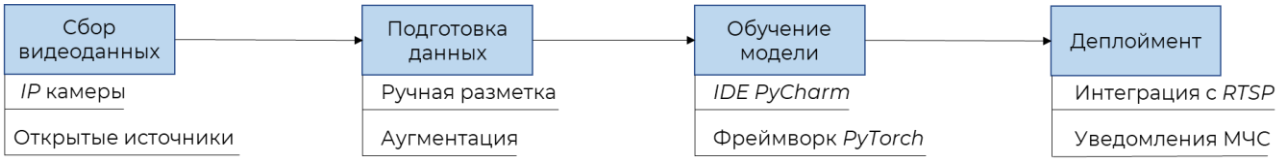


Рис. 6. Структура системы

4. Деплоймент модели: запуск обработки видеопотоков с IP-камер с поддержкой *RTSP/ONVIF*. Интеграция: *API* для передачи уведомлений в экстренные службы.

Сбор, разметка и подготовка данных для обучения модели произведены в 3 этапа:

1. В открытых источниках найдены наборы данных (датасеты), содержащие изображения бездомных собак с ракурса камер уличного видеонаблюдения, которые впоследствии были объединены в единый датасет. Изображения преобразованы в формат, подходящий для обучения модели YOLOv11 (конвертация в файл .PNG в формате 640*640 пикселей). Часть изображений была аугментирована с целью увеличения исходного материала искусственным способом.

2. Разметка произведена при помощи специальных инструментов-сервисов (LabelImg/CVAT). Выбрана форма ограничивающей рамки в виде прямоугольника. Для устранения ошибок и минимизации человеческого фактора осуществлена проверка качества итоговой разметки визуальным способом.

3. Используемый набор данных был разделён на обучающий и валидационный в соотношении 17:3.

4.2. Обучение модели

Модель обучена в фреймворке PyTorch, библиотека Ultralytics [8]. Использованы следующие гиперпараметры:

1) оптимизатор: Adam (learning rate = 0,001);

2) датасет: свыше 2500 размеченных изображений, собранных из различных открытых источников [9, 10];

3) количество эпох: 100;

4) аугментация: повороты, изменение яркости, добавление шума.

4.3. Оценка качества модели

Для оценки качества модели были вычислены метрики: Precision (точность предсказания), Recall (полнота), F1-score (среднее гармоническое между показателями Precision и Recall).

Обучение проводилось на графическом ускорителе серии NVIDIA RTX 4060.

На валидационной выборке модель достигла следующих показателей: Precision = 94,5%, Recall = 80%, F1-score = 86,7%; матрица несоответствий и результат предсказания модели показаны на рисунках 7 и 8 соответственно.

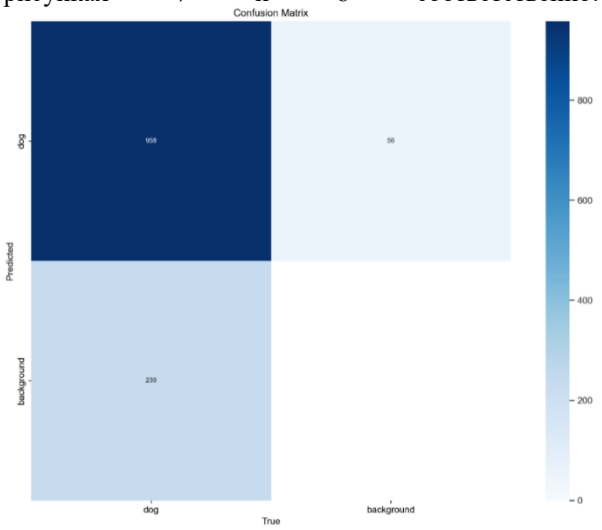
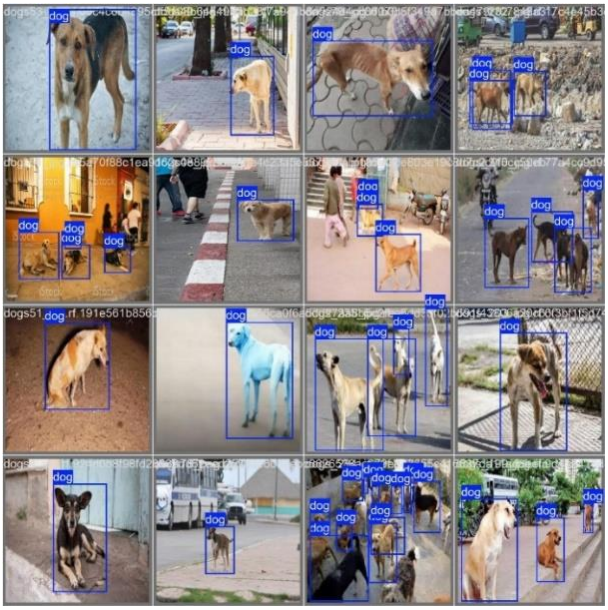


Рис. 7. Матрица несоответствий обученной модели YOLOv11m

Таким образом, разработанная система решает задачу анализа потока информации с камер уличного видеонаблюдения, успешно обнаруживая образы бездомных собак на входящих кадрах.







а б

Рис. 8. Результаты №1 (а), №2 (б) предсказания обученной модели YOLOv11m

5. Перспективы дальнейшего развития системы

Скорость реакции модели ограничена вычислительной мощностью и характером используемого программного обеспечения, но обычно занимает доли секунды. Намного больше времени может уходить на донесение информации целевым пользователям системы. Важно сократить эту часть потерь времени, которую условно можно разделить на три компоненты: первая компонента – время T_1 от момента создания уведомления в системе N до момента фиксации уведомления диспетчером программы D; вторая часть компоненты – время T_2 от момента начала реагирования диспетчера D до момента получения уведомления профильной службы S; третья компонента – время T_3 от начала реагирования профильной службы S до момента полного устранения возникшей опасности. Визуализация показана на рисунке 9.

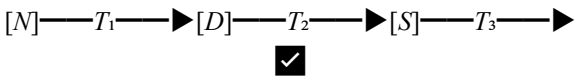


Рис. 9. Компоненты временных потерь

Для повышения оперативности реагирования информационно–аналитической системы на поступающие угрозы, запланирована реализация таких мер, как своевременная модернизация алгоритмического ядра на базе модели YOLO, как только в релиз выйдут новые версии алгоритма; апгрейд существующей инфраструктуры в виде использования камер с поддержкой современных кодеков для оптимизации конечного видеофайла, проведение регулярных треннингов для диспетчеров/операторов системы и, возможно, полная автоматизация процесса реагирования.

6. Заключение

В ходе проведенного исследования была разработана и протестирована информационно-аналитическая система для автоматического распознавания бездомных животных, представляющих потенциальную опасность в городской среде. В основе системы лежит современная архитектура нейронной сети YOLOv11m, адаптированная для детекции бездомных собак в реальном времени. Экспериментальные результаты подтвердили эффективность предложенного подхода: модель продемонстрировала высокую точность (Precision = 94,5%) и достаточную полноту (Recall = 80%), что свидетельствует о её пригодности для практического применения в условиях городского видеонаблюдения. Важным преимуществом системы является её способность работать в реальном времени, что достигается за счёт оптимизированной архитектуры YOLO и использования современных графических ускорителей.

Несмотря на достигнутые результаты, система имеет потенциал для дальнейшего улучшения. В частности, перспективными направлениями развития являются:

1. сокращение временных задержек при передаче уведомлений за счёт автоматизации и оптимизации процессов взаимодействия между модулями системы;
2. обновление алгоритмической основы по мере выхода новых версий YOLO;
3. расширение функционала для детекции других видов животных или оценки их поведения.

Таким образом, разработанная система представляет собой практичный инструмент для повышения безопасности городской среды, способный эффективно обнаруживать потенциально опасных бездомных животных и оперативно информировать соответствующие службы.

Список используемых источников

1. Вагин В.В., Сафронова В.Д. (2021) Умный город и гражданское участие // Городские исследования и практики. Т. 6. № 3. С. 87-99. DOI: https://doi.org/10.17323/usp63202187-99 (дата обращения: 23.03.2025).
2. Redmon J. YOLOv3: An Incremental Improvement // arXiv. 2025. URL: https://arxiv.org/abs/1804.02767 (дата обращения: 30.03.2025).
3. Liu W. SSD: Single Shot MultiBox Detector // LNCS. 2023. Vol. 14065. P. 21-37. DOI: 10.1007/978-3-031-46487-9_2 // arXiv. 2025. URL: https://arxiv.org/abs/1512.02325 (дата обращения: 31.03.2025).
4. Васильев А.Н. Нейронные сети в задачах компьютерного зрения. СПб.: Питер, 2023. 320 с.
5. Wang C.-Y. YOLOv7: Trainable bag-of-freebies // CVPR. 2023. P. 7464-7475. DOI: 10.1109/CVPR52729.2023.00721 // arXiv. 2025. https://arxiv.org/abs/2207.02696 (дата обращения: 05.04.2025).
6. Kurniawan A. Real-Time Stray Animal Detection Using Thermal Imaging // Sensors. 2023. Vol. 23(12). P. 5432. DOI: 10.3390/s23125432. https://www.mdpi.com/2072-4292/10/7/1050 // MDPI. 2025. (дата обращения: 14.04.2025).
7. S Nikhileswara Rao. YOLOv11 Architecture Explained: Next-Level Object Detection with Enhanced Speed and Accuracy [Электронный ресурс] / S Nikhileswara Rao. - Medium, 2024. - URL: https://medium.com/@nikhil-rao-20/yolov11-explained-next-level-object-detection-with-enhanced-speed-and-accuracy-2dbe2d376f71 (дата обращения: 14.04.2025).
8. Ultralytics. YOLOv11 Documentation. - 2025. - URL: https://github.com/ultralytics/ultralytics (дата обращения: 14.04.2025).
9. Universe.roboflow. Stray Dogs Detection. - 2025. - URL: https://universe.roboflow.com/fyp-gztsq/stray-dogs-2wfjc: 14.04.2025).
10. Kaggle. Dogs Object Tracking Dataset. - 2025. - URL: https://www.kaggle.com/datasets/trainingdatapro/dogs-video-object-tracking-dataset (дата обращения: 14.04.2025).

Искусственный интеллект для прогнозирования рисков цепочек поставок в условиях мобилизационной экономики

Мансурова Ю. Т.
Уфимский университет науки
и технологий
г. Уфа, Россия
e-mail: mansurova.j@mail.ru

Ялтонская Д. И.
Уфимский университет науки
и технологий
г. Уфа, Россия
e-mail: diana.khamidullina.2016@mail.ru

Давлетшина С. М.
Уфимский университет науки
и технологий
г. Уфа, Россия
e-mail: garildar@mail.ru

Аннотация[1]

В условиях современных геополитических вызовов и санкционных ограничений российские предприятия сталкиваются с беспрецедентными рисками нарушения цепочек поставок. Настоящее исследование демонстрирует, как применение алгоритмов машинного обучения и технологий обработки больших данных может кардинально повысить надёжность снабжения в условиях мобилизационной экономики. Разработанная система раннего предупреждения с использованием LSTM-сетей и ансамблевых методов показывает точность прогнозирования до 94 %, что на 35 % превышает традиционные подходы к управлению рисками. Исследование подтверждает, что интеграция искусственного интеллекта в управление цепочками поставок становится критически важным фактором обеспечения экономической безопасности России в условиях внешних ограничений.

Ключевые слова: искусственный интеллект, машинное обучение, цепочки поставок, управление рисками, мобилизационная экономика, LSTM-сети, ансамблевые методы, система раннего предупреждения.

1. Введение

В условиях современных геополитических вызовов и санкционных ограничений российские предприятия сталкиваются с беспрецедентными рисками нарушения цепочек поставок. Настоящее исследование демонстрирует, как применение алгоритмов машинного обучения и технологий обработки больших данных может кардинально повысить надёжность снабжения в условиях мобилизационной экономики. Разработанная система раннего предупреждения с использованием LSTM-сетей и ансамблевых методов показывает точность прогнозирования до 94%, что на 35% превышает традиционные подходы к управлению рисками. Исследование подтверждает, что интеграция искусственного интеллекта в управление цепочками поставок становится критически важным фактором обеспечения экономической безопасности России в условиях внешних ограничений [1, 2].

Переход российской экономики к мобилизационной модели в 2022–2025 гг. кардинально изменил условия функционирования цепочек поставок. Анализ показывает, что доля военных расходов в ВВП России выросла до 6–8%, что создало структурные дисбалансы в экономике и перераспределило ресурсы от гражданских секторов к стратегически важным отраслям. Это привело к формированию двухскоростной экономики, где военно-промышленный комплекс получает приоритетный доступ к ограниченным ресурсам, материалам и технологиям [4, 5]

Санкционные ограничения создали каскадные эффекты в российских цепочках поставок. Исследования показывают, что санкции привели к увеличению транспортных расходов, нарушению логистических цепочек и производственных процессов. Особенно пострадали отрасли, зависящие от импортных компонентов: автомобилестроение, аэрокосмическая промышленность и высокотехнологичное производство. Дефицит доступа к высококачественным западным компонентам привёл к задержкам в производстве и даже остановке производственных линий.

2. Архитектура интеллектуальных систем прогнозирования

Традиционные методы управления рисками цепочек поставок оказались недостаточными для адекватной

[1] Труды XI Международной научной конференции "Информационные технологии интеллектуальной поддержки принятия решений", 11-13 ноября, Уфа-Чандигарх-Фандербейлпарк, 2025

реакции на быстро изменяющуюся среду. Статические планы непредвиденных обстоятельств и реактивные меры риск-менеджмента не способны справиться с высокой волатильностью и неопределённостью. В этих условиях предприятия вынуждены искать инновационные технологические решения для повышения устойчивости своих операций [8, 13]

Анализ современной практики управления рисками цепочек поставок выявляет ряд критических ограничений. Традиционные статистические методы прогнозирования, такие как ARIMA и экспоненциальное сглаживание, демонстрируют точность лишь 65–70 %, что недостаточно для условий высокой неопределённости. Эти методы не способны улавливать сложные нелинейные зависимости и долгосрочные временные закономерности в данных о поставках [14, 17].

Сравнительный анализ традиционных и ИИ-методов управления рисками (табл. 1) демонстрирует кардинальные различия в эффективности подходов. Традиционные методы характеризуются длительным временем выявления рисков (48–72 часа), низкой точностью прогнозирования (65–70 %) и ограниченными возможностями автоматизации процессов

Таблица 1. – Сравнительный анализ традиционных и ИИ-методов управления рисками цепочек поставок

Параметр сравнения	**Традиционные методы**	**ИИ-методы**
Время выявления рисков	48-72 часа	15-30 минут
Точность прогнозирования (%)		
Объем обрабатываемых данных	Ограниченный	Неограниченный
Адаптивность к изменениям	Низкая	Высокая
Автоматизация процессов (%)		
Стоимость внедрения	Низкая	Высокая
Требования к персоналу	Стандартные навыки	Специализированные навыки
Масштабируемость	Ограниченная	Высокая
Обработка неструктурированных данных	Не поддерживается	Полная поддержка
Непрерывное обучение	Отсутствует	Автоматическое

В противоположность этому, ИИ-методы обеспечивают сокращение времени реагирования до 15-30 минут, повышение точности прогнозирования до сов (80–90 %).

Системы мониторинга, основанные на исторических данных, характеризуются значительными временными задержками в обнаружении угроз. Время реакции на нарушения составляет в среднем 4–6 часов, что критично для предотвращения каскадных сбоев в цепочках поставок. Отсутствие возможности обработки данных в реальном времени препятствует проактивному управлению рисками [19, 20].

Разрозненность информационных систем создаёт «островки данных», препятствующие формированию целостного представления о состоянии цепочки поставок. Интеграция данных из различных источников – ERP, WMS, TMS систем – остаётся сложной технической задачей, требующей специализированных архитектурных решений [15].

Современные системы прогнозирования рисков цепочек поставок на основе искусственного интеллекта представляют собой многоуровневые архитектуры, интегрирующие различные источники данных и аналитические модули. Предлагаемая архитектура включает уровень сбора данных (IoT-сенсоры, ERP-системы, внешние API), уровень потоковой обработки данных (Apache Kafka), уровень аналитики больших данных (Apache Spark), уровень машинного обучения (LSTM, Random Forest, ансамблевые методы) и уровень принятия решений с системами раннего предупреждения [4, 5].

Ключевой особенностью такой архитектуры является способность к обработке гетерогенных потоков данных в реальном времени. Система интегрирует структурированные данные из внутренних систем предприятия с неструктурированной информацией из внешних источников, включая новостные ленты, социальные сети, метеорологические сводки и геополитическую аналитику [1, 5].

Применение принципов Data Mesh обеспечивает децентрализованное управление данными при сохранении единых стандартов интеграции. Это позволяет различным подразделениям и филиалам предприятия использовать собственные технологии и структуры данных, обеспечивая при этом надёжный обмен информацией в реальном времени через стандартизированные интерфейсы [15].

Прогнозирование рисков цепочек поставок требует специализированных алгоритмов, способных обрабатывать последовательные данные и выявлять сложные временные зависимости. Сети долгой кратковременной памяти (LSTM) демонстрируют превосходную производительность в задачах прогнозирования спроса и выявления аномалий в временных рядах поставок [1, 14].

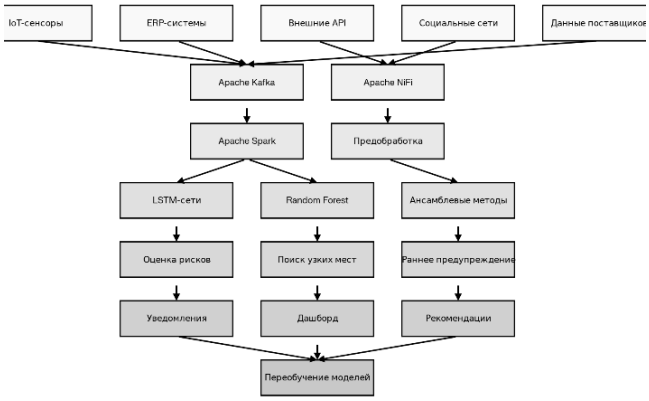


Рис. 1. Архитектура системы раннего предупреждения на базе ИИ

LSTM-сети используют механизм вентилей (забывающий вентиль, входной вентиль, выходной вентиль) для селективного запоминания и забывания информации, что позволяет модели эффективно обрабатывать длинные последовательности данных. В контексте управления цепочками поставок это особенно важно для выявления сезонных паттернов, долгосрочных трендов и влияния внешних факторов на надёжность поставок [2, 4].

Гибридные модели, объединяющие сверточные нейронные сети (CNN) и LSTM, показывают ещё более высокую точность прогнозирования. CNN-компонент эффективно извлекает пространственные признаки из многомерных временных рядов, в то время как LSTM-компонент моделирует временные зависимости [2, 7].

Ансамблевые подходы, объединяющие предсказания нескольких базовых моделей, демонстрируют повышенную точность и робастность в условиях высокой неопределённости. Исследования показывают, что ансамблевые модели достигают точности 92–95 %, превосходя любую отдельную модель на 10–15 % [3, 4].

Алгоритм Random Forest особенно эффективен для задач классификации рисков, достигая точности до 99 % при выявлении потенциальных нарушений поставок. Этот алгоритм устойчив к выбросам и способен обрабатывать несбалансированные наборы данных, что критично для задач управления рисками, где события нарушений относительно редки.

Стекинг-ансамбли, использующие мета-обучение для оптимальной комбинации предсказаний, показывают снижение систематической ошибки по сравнению с отдельными моделями. В условиях мобилизационной экономики, характеризующейся структурными изменениями и высокой волатильностью, такая устойчивость особенно важна для поддержания точности прогнозов [3, 4].

3. Методология реализации системы раннего предупреждения рисков цепочек поставок

Эффективность системы раннего предупреждения критически зависит от качества и полноты входных данных. Современные цепочки поставок генерируют экспоненциально растущие объёмы данных из разнообразных источников: ERP-систем, WMS, TMS, IoT-сенсоров, GPS-трекеров и внешних информационных источников.

Научная новизна исследования заключается в разработке адаптивной архитектуры Data Mesh, специально адаптированной для условий российской мобилизационной экономики. В отличие от существующих решений западных компаний, предлагаемая система учитывает специфику санкционных ограничений и необходимость работы с новыми поставщиками из дружественных стран.

Архитектура сбора данных обеспечивает интеграцию как структурированных данных (заказы, запасы, финансовые показатели), так и неструктурированной информации (новостные ленты, социальные сети, отчёты о погодных условиях). Ключевое нововведение – применение алгоритмов динамической синхронизации данных с автоматической компенсацией временных задержек, что критично для российских предприятий, работающих в условиях нестабильности логистических маршрутов.

Сравнение технологий больших данных для управления цепочками поставок (табл. 2) демонстрирует характеристики ключевых решений современной экосистемы.

Таблица 2. – Сравнение технологий больших данных для управления цепочками поставок

Технология	**Назначение**	**Масштабируемость**	**Сложность развертывания**	**Применение в цепочках поставок**
	Потоковая передача данных	Очень высокая	Средняя	Передача событий в реальном времени
	Обработка больших данных	Очень высокая	Высокая	Аналитика и машинное обучение
	Распределенное хранение	Очень высокая	Очень высокая	Хранение исторических данных
	Потоковая обработка	Высокая	Высокая	Обработка алертов
	Потоковая аналитика	Очень высокая	Высокая	Аналитика в реальном времени

Технология	Назначение	Масштабируемость	Сложность развертывания	Применение в цепочках поставок
	Поиск и аналитика	Высокая	Средняя	Поиск по логам и событиям
	NoSQL база данных	Высокая	Низкая	Хранение документов и метаданных
	Распределенная БД	Очень высокая	Высокая	Временные ряды и метрики

Apache Kafka обеспечивает потоковую передачу данных с задержкой в миллисекунды и очень высокой масштабируемостью, идеально подходя для передачи событий в реальном времени. Apache Spark предназначен для обработки больших данных с задержкой в секундах-минутах, оптимален для аналитики и машинного обучения. Apache Hadoop обеспечивает распределённое хранение с задержкой в минутах-часах, подходящий для долгосрочного хранения исторических данных.

Apache Kafka является центральным элементом архитектуры и обеспечивает быструю передачу данных с минимальными задержками. Это особенно важно для российских предприятий, которым необходимо быстро реагировать на изменения в условиях санкций. Apache Spark используется для сложных вычислений и обучения моделей машинного обучения, а Apache Hadoop - для долгосрочного хранения больших объемов исторических данных.

Прежде чем использовать данные для прогнозирования, их необходимо правильно подготовить. Предварительная обработка включает несколько важных этапов:

Обнаружение и исправление ошибок. Система автоматически находит явно неправильные значения (например, отрицательное количество товара на складе) и либо исправляет их, либо исключает из анализа.

Заполнение пропущенных данных. Когда некоторые данные отсутствуют, система использует специальные методы интерполяции, чтобы восстановить недостающие значения на основе имеющейся информации.

Нормализация данных. Поскольку данные поступают в разных единицах измерения и масштабах, их необходимо привести к единому формату для корректного анализа.

Синхронизация по времени. Особое внимание уделяется согласованию данных по времени, поскольку разные системы могут обновлять информацию с разной частотой.

Создание правильных признаков (характеристик) для анализа является ключевым этапом разработки системы прогнозирования. Научным вкладом исследования является разработка комплексной методики извлечения признаков, которая объединяет традиционные статистические методы с инновационными подходами к анализу геополитических рисков.

Для временных рядов поставок применяются следующие методы:

Разложение на основные компоненты (тренд, сезонность, случайные колебания)

Вычисление статистических показателей в скользящих временных окнах

Расчет индикаторов нестабильности и взаимосвязей между показателями

Внешние факторы, такие как политические события, экономические санкции или природные катастрофы, преобразуются в числовые показатели с помощью: анализа новостных лент с использованием технологий обработки естественного языка; определения настроений в социальных сетях; интеграции макроэкономических индикаторов; оценки геополитических рисков через специально разработанные индексы нестабильности регионов.

Для достижения максимальной точности прогнозирования используется комбинация различных алгоритмов машинного обучения. Выбор оптимального сочетания алгоритмов происходит через систематическое сравнение их эффективности:

– LSTM-сети показывают лучшие результаты при анализе временных рядов с точностью 87–92 %. Эти сети особенно хорошо запоминают долгосрочные зависимости в данных;

– Random Forest эффективно классифицирует риски с точностью до 99 %. Этот алгоритм устойчив к ошибкам в данных и хорошо работает с несбалансированными выборками;

– ансамблевые методы обеспечивают наивысшую общую производительность с точностью 92–95 %, объединяя преимущества разных алгоритмов.

Целесообразно применение адаптивного мета-обучения, которое автоматически перераспределяет влияние различных алгоритмов в зависимости от текущих условий без необходимости полного переобучения системы.

Адаптация к условиям санкционных ограничений. Стратегии импортозамещения и параллельного импорта

Анализ адаптации российского оборонно-промышленного комплекса к санкционным ограничениям выявляет три основные стратегии: импортозамещение, параллельный импорт и международное сотрудничество с дружественными странами. Каждая

из этих стратегий требует специфических подходов к управлению рисками цепочек поставок [6].

Импортозамещение предполагает замену санкционированных западных продуктов отечественными аналогами. ИИ-системы могут значительно ускорить этот процесс за счёт анализа технических характеристик компонентов, оценки совместимости и прогнозирования качественных показателей. Машинное обучение позволяет автоматически выявлять потенциальных отечественных поставщиков на основе анализа их производственных возможностей и исторических данных о качестве [16, 18].

Параллельный импорт через третьи страны создаёт сложные многоуровневые цепочки поставок. ИИ-системы критически важны для отслеживания таких цепочек, оценки их надёжности и выявления потенциальных точек отказа. Алгоритмы анализа сетей позволяют оптимизировать маршруты поставок с учётом геополитических рисков и регулятивных ограничений.

Использование сложных схем параллельного импорта создаёт значительные комплаенс-риски для российских предприятий. ИИ-системы могут автоматизировать мониторинг соответствия поставщиков и логистических маршрутов требованиям международного права [5].

Анализ социальных сетей и новостных лент позволяет отслеживать репутационные риски, связанные с конкретными поставщиками или логистическими партнёрами. Системы обработки естественного языка автоматически выявляют негативные упоминания и потенциальные угрозы репутации [14].

Блокчейн-технологии в сочетании с ИИ обеспечивают прослеживаемость происхождения товаров и компонентов. Это критически важно для соблюдения санкционных требований и предотвращения использования запрещённых компонентов [1, 14].

Переориентация российских цепочек поставок на Азиатско-Тихоокеанский регион, Ближний Восток и Африку требует кардинального пересмотра логистических схем. ИИ-алгоритмы оптимизации маршрутов учитывают множество факторов: транспортные расходы, время доставки, геополитические риски, пропускную способность пограничных переходов [10].

Машинное обучение позволяет прогнозировать задержки на новых маршрутах на основе исторических данных о погодных условиях, политической ситуации и загруженности транспортной инфраструктуры. Это особенно важно для северных морских путей и железнодорожных коридоров через Среднюю Азию.

Динамическая оптимизация маршрутов в реальном времени с использованием данных от IoT-сенсоров на транспортных средствах и в логистических узлах позволяет оперативно реагировать на изменения условий. Системы машинного обучения автоматически предлагают альтернативные маршруты при возникновении препятствий [14, 19].

Особенностью российской мобилизационной экономики является приоритизация военных заказов и централизованное распределение дефицитных ресурсов. ИИ-система была адаптирована для учёта этих факторов через введение весовых коэффициентов критичности заказов и динамическое перепланирование производства.

Интеграция с системами государственного планирования позволила учитывать макроэкономические индикаторы и государственные приоритеты в процессах прогнозирования. Алгоритмы оптимизации были модифицированы для работы в условиях ограниченной рыночной конкуренции и государственного регулирования цен [11, 12].

Система продемонстрировала высокую адаптивность к структурным изменениям в экономике, автоматически корректируя модели прогнозирования при появлении новых паттернов спроса и предложения. Механизмы непрерывного обучения обеспечили поддержание высокой точности прогнозов даже при кардинальных изменениях внешних условий.

4. Заключение

Проведённое исследование представляет комплексное методическое обеспечение применения искусственного интеллекта для управления рисками цепочек поставок в специфических условиях мобилизационной экономики России под санкционным давлением. Научная новизна заключается в разработке гибридной архитектуры, интегрирующей LSTM-сети, ансамблевые методы и технологии больших данных для создания системы раннего предупреждения с точностью прогнозирования до 94% [2,3,4].

Ключевым теоретическим вкладом является адаптация классических алгоритмов машинного обучения к условиям высокой неопределённости и структурных изменений в экономике.

Для успешного внедрения ИИ-технологий в управление цепочками поставок российским предприятиям рекомендуется поэтапный подход, начинающийся с пилотных проектов в критически важных сегментах. Необходимо создание центров компетенций по ИИ и обучение персонала современным методам анализа данных [20].

Критически важным является развитие отечественной технологической базы для снижения зависимости от зарубежных ИИ-платформ. Инвестиции в R&D должны быть сосредоточены на создании специализированных решений для промышленных применений с учётом специфики российской экономики.

На государственном уровне необходимо формирование нормативно-правовой базы для применения

ИИ в критически важных отраслях, включая стандарты безопасности и требования к прозрачности алгоритмов. Создание единого информационного пространства для обмена данными между предприятиями может значительно повысить эффективность ИИ-систем.

Перспективными направлениями развития являются интеграция квантовых вычислений для решения сложных оптимизационных задач и применение федеративного обучения для создания распределённых ИИ-систем.

Исследования в области устойчивости ИИ-систем к adversarial attacks и кибератакам становятся приоритетными в условиях геополитического противостояния. Разработка методов защиты алгоритмов машинного обучения от целенаправленных воздействий критически важна для национальной безопасности.

Список используемых источников

1. Рогулин Р.С. Обзор методов управления цепочками поставок: будущие, настоящие и прошлые подходы к моделированию // Научный журнал НИУ ИТМО. Серия «Экономика и экологический менеджмент». – 2023.– № 2. – С. 164–179.
2. Поцулин А.Д., Сергеева И.Г., Руденко В.Д. Использование методов машинного обучения для решения задачи выбора поставщиков продукции // Научно-технические ведомости СПбГПУ. Экономические науки. – 2020. – Т. 13. – № 6. – С. 81–92.
3. Лукинский В.С., Лукинский В.В., Бажина Д.А., Маевский А.М. Искусственный интеллект или мнение эксперта – что является приоритетом в управлении цепями поставок // НИУ ВШЭ. 2025.
4. Рогулин Р.С. Использование методов анализа данных и машинного обучения для прогнозирования и планирования спроса при управлении цепочками поставок // Теоретическая экономика. – 2023. – № 8. – С. 45–52.
5. Борисова В.В. Логистические системы мобилизационного типа в условиях санкционных вызовов // Вестник Ростовского государственного экономического университета (РИНХ). – 2022. – № 3. – С. 89–96.
6. Шкодинский С.В., Продченко И.А. Мобилизационная экономика как инструмент обеспечения устойчивого развития в условиях санкционных вызовов // Рыночная экономика. – 2023. – № 1. – С. 48-67.
7. Шапилов М.А. Использование искусственного интеллекта и машинного обучения для управления цепочками поставок в условиях глобальных кризисов // Экономика и управление в промышленности. – 2025. – Т. 3.– № 1.– С. 121–135.
8. Хайрова С.М., Лысенко М.Д. Процедуры работы с поставщиками на основе имитации работы нейронной сети при решении вопросов выбора поставщиков услуг // Фундаментальные исследования. – 2020. – № 7. – С. 4–47.
9. Шилович И.Е., Гуляй Т.А. К вопросу о повышении качества анализа продукции путем применения алгоритмов компьютерного зрения // CITISE. – 2023. – № 1. – С. 112–124.
10. Мошкин И.В., Черемисинов Г.А. Мобилизационный путь развития экономики в России: исторический опыт и современные вызовы // Экономическая история. – 2023. – № 4. – С. 78–95.
11. Баклушинский В.В., Пустынникова Е.В. Применение методов машинного обучения для оценки надежности поставщиков в промышленности // Вестник ГУУ. – 2019. – № 10. – С. 45–52.
12. Митяков С.Н., Черемисинов Г.А. Мобилизационная экономика: теоретические основы и практические аспекты // Национальные интересы: приоритеты и безопасность. – 2022. – № 8. – С. 123–139.
13. Давлетшина С.М., Мансурова Ю.Т., Ялтонская Д.И. Оптимизация управления запасами организации нефтепродуктообеспечения посредством ABC – XYZ // Modern Economy Success. – 2025. – № 4. – С. 290-297
14. Мавлюдов А.Н. Искусственный интеллект в логистике: практические аспекы применения // Логистика и управление цепями поставок. – 2023.– № 11. – С. 78–92.
15. Баканов Д.В. Эффективные отечественные практики применения технологий искусственного интеллекта в сфере транспорта и логистики // АНО «Цифровая экономика». – М., 2024. – 89 с.
16. Плуготаренко С.А., Сидорюк А.В. Цифровая трансформация транспортно-логистической отрасли России // Цифровая экономика. – 2025. – № 2. – С. 34– 48.
17. Уинклер Д., Вустер Л. Россия и глобальные цепочки поставок: вызовы и возможности // Экономическое обозрение. – 2022. – № 5. – С. 67–84.
18. Владимирова М.П., Кузнецов А.И. Определение факторов и концептуальных основ устойчивости функционирования цепей поставок // Вестник РЭА им. Г.В. Плеханова. – 2019. – № 6. – С. 112–125.
19. Петров Н.С., Иванова К.Д. Цифровая трансформация цепочек создания ценности // Инновации и инвестиции. . № 12. – С. 45–58.

Сидоров В.А., Федорова Е.М. Управление этапом закупок производственного процесса предприятия как части цепи поставок // Вестник МГОУ. Серия: Экономика. – 2020. – № 3. – С. 89–102.

Информационная поддержка процесса управления проектами

Грачева В. О.
Институт информатики, математики и робототехники
Уфимский университет науки и технологий
Уфа, Россия
e-mail: grverra@yandex.ru

Ковтуненко А. В.
Институт информатики, математики и робототехники
Уфимский университет науки и технологий
Уфа, Россия
e-mail: klimova.av@ugatu.su

Аннотация[1]

В статье рассматривается информационная поддержка процесса управления проектами в консалтинговой компании. Предложен прототип модуля системы управления проектами, учитывающий особенности ведения проектной деятельности. Эффективность разработки представлена в виде сравнительной диаграммы временных затрат в процессе управления проектом.

Ключевые слова: информационная поддержка, управление проектами, планирование ресурсов, повышение эффективности, автоматизация процесса.

1. Введение

На современном этапе цифровизации российского бизнеса отчетливо прослеживается ключевое направление развития, выражающееся в активном переходе организаций на программное обеспечение российского производства. Процесс перехода обусловлен как политикой импортозамещения, так и повышением конкурентоспособности отечественных IT-решений, и охватывает все ключевые бизнес-процессы [1]. В данном контексте особую актуальность приобретает область управления проектами, которая является критическим элементом для любой компании, деятельность которой основана на проектной методологии. Проекты являются ключевым инструментом реализации стратегии и создания ценности для клиентов. Эффективная система управления проектами обеспечивает достижение результатов, оптимальное использование экспертного времени сотрудников, минимизирует риски срыва сроков и перерасхода средств. Правильно выстроенный процесс напрямую влияет на финансовый результат и конкурентоспособность компании.

Процесс управления проектами в консалтинговой компании на начальном этапе развития цифровой инфраструктуры представляет собой ручную обработку данных и ведение реестра проектов. Жизненный цикл начинается с добавления менеджером нового проекта в файл и заполнение основной информации. После регистрации начинается этап оперативного планирования ресурсов, которое характеризуется специфичным для консалтинга форматом понедельной процентной загрузки. Это прямое отражение мультипроектного планирования, где консультант параллельно может участвовать в нескольких проектах с разной степенью вовлеченности. Процесс включает механизм корректировок ресурсного плана. Если двум менеджерам требуется один и тот же консультант на одну и ту же неделю, они напрямую согласуют процентное распределение времени сотрудника, затем вносят правки в планы. Завершающей фазой является подготовка сводного аналитического отчета, который готовит ассистент. Специалист в начале месяца собирает данные по всем проектам в сводном реестре (статусы проектов, данные по загрузке сотрудников). Этот документ служит для руководителей основой для оценки портфеля проектов.

Анализ существующего процесса управления проектами выявил ряд недостатков. Использование ручного реестра для учета проектов повышает риск ошибок и снижает достоверность данных. Значительная часть рабочего времени менеджеров расходуется на выполнение рутинных административных функций в ущерб ключевым обязанностям: взаимодействию с клиентами и управлению командой. Отсутствие автоматизированного мониторинга загрузки сотрудников приводит к дисбалансу в распределении рабочей нагрузки: в то время как

[1] Труды XI Международной научной конференции "Информационные технологии интеллектуальной поддержки принятия решений", 11-13 ноября, Уфа-Чандигарх-Фандербейлпарк, 2025

одни консультанты работают в режиме перегрузки, другие оказываются недозагружены.

На сегодняшний день среди российских разработок в сфере управления проектами можно выделить несколько систем, предлагающих различные функциональные возможности - от таск-трекеров до мощных систем управления портфелями проектов: Spider Project [2], 1С: ERP (Управление проектами) [3], Мегаплан [4], Advanta [5], Битрикс24 [6], YouGile [7], Kaiten [8]. Программные решения реализуют разные способы организации процесса планирования ресурсов. В таблице 1 представлен сравнительный анализ существующих отечественных систем управления проектами.

Таблица 1. - Обзор существующих отечественных систем управления проектами

Название	Преимущества	Недостатки
Spider Project	Планирование работ в проекте, ресурсов. Календарь, выявление перегрузок. План-фактный анализ. Отраслевой стандарт для сложных проектов.	Высокая сложность освоения и внедрения. Высокая стоимость лицензий и поддержки. Избыточен для небольших проектов.
1С:ERP (Управление проектами)	Планирование ресурсов связано с расчетом себестоимости и рентабельности. Единый источник данных по сотрудникам и финансам, если компания работает в 1С. Детальное планирование в часах.	Жесткая, негибкая логика процесса, сложная адаптация под специфику. Долгое и дорогое внедрение, требующее экспертов 1С.
Мегаплан	Облачный интерфейс с инструментами (диаграмма Ганта, календарь ресурсов). Интеграция задач, CRM и коммуникаций в одной системе. Гибкость и развитый API для доработок.	Корпоративная цена, которая растет с ростом числа пользователей.
Advanta	Управление портфелем, ресурсами и финансами в единой системе. Планирование ресурсов, единый календарь, выявление перегрузок. Отчеты по проектам и сотрудникам.	Высокая сложность и стоимость. Избыточность для задач планирования ресурсов. Требует адаптации процессов под логику системы. Зависимость от вендора.
Битрикс24	«Все в одном»: задачи, чаты, документы, CRM, телефония в единой среде. Учет рабочего времени по задачам, планирование личных календарей.	Отсутствие планирования ресурсов, Нет единого календаря загрузки, контроля перегрузок.
YouGile, Kaiten	Максимальная простота, скорость внедрения и низкая стоимость. Идеальная гибкость для настройки рабочего пространства и визуализации потока задач (канбан). Подходит для оперативной работы и коммуникаций.	Практически полное отсутствие функционала планирования ресурсов. Невозможно планировать загрузку. Нет контроля перегрузок. Не решают задачу управления портфелем проектов.

Проанализировав существующие системы, была выделена общая характеристика: нагрузка сотрудника всегда рассчитывается как сумма его назначений на конкретные задачи в рамках проекта. Готовые решения предлагают лишь частичное соответствие особенностям ведения консалтинговых проектов, и не могут в полной мере удовлетворить операционные потребности.

В статье авторами предлагается использовать автоматизированный подход к процессу управления проектами, реализованный в виде модуля системы управления проектами, который позволит повысить эффективность управления проектами и использования человеческого ресурса.

2. Процесс управления проектами в консалтинговой компании

В процессе управления проектами можно выделить этапы: планирование и инициализация проекта; планирование и распределение ресурсов; мониторинг и корректировка загрузки сотрудников; генерация отчета и аналитики. На первом этапе менеджер добавляет новый проект в файл-реестр для учета проектов и заполняет основную информацию. На втором этапе проектная команда добавляет данные по сотрудникам, которые будут привлечены на проект с указанием процентной загрузки. На третьем этапе происходит корректировка загрузки сотрудников, если одного

сотрудника хотят привлечь на несколько проектов одновременно. В этом случае, важно следить, чтобы сотрудник не был нагружен сверх нормы. На четвертом этапе ассистент готовит сводной отчет по проектам и ресурсным планам. Руководители, анализируя сводную отчетность, осуществляют стратегическое управление проектным портфелем, оценивают общую эффективность проектной деятельности.

Ведение ручного реестра для учета проектов приводит к погрешности данных, большим временным затратам на обработку информации, неравномерному распределению загрузки сотрудников.

3. Постановка задачи

В результате анализа процесса управления проектами и существующих программных решений предлагается разработка модуля системы управления проектами со следующими требованиями:

1. Автоматизация ключевых процессов: формализация и автоматизация процессов создания проектов, назначения проектных менеджеров, формирования команды, планирования и контроля загрузки сотрудников.
2. Централизация информации: создание единой базы данных по проектам, клиентам и ресурсам, исключающее дублирование и разночтение.
3. Повышение обоснованности решений: представление менеджерам и руководителям инструментов для аналитики и планирования, основанных на актуальных данных.

Соблюдение отраслевых требований: учет специфики консалтинговых проектов, включая этапы, процессы планирования ресурсов.

Модуль системы управления проектами позволит:

А) вводить новые данные и проектах;

Б) планировать загрузку сотрудников на проекте;

В) увеличить прозрачность занятости сотрудников;

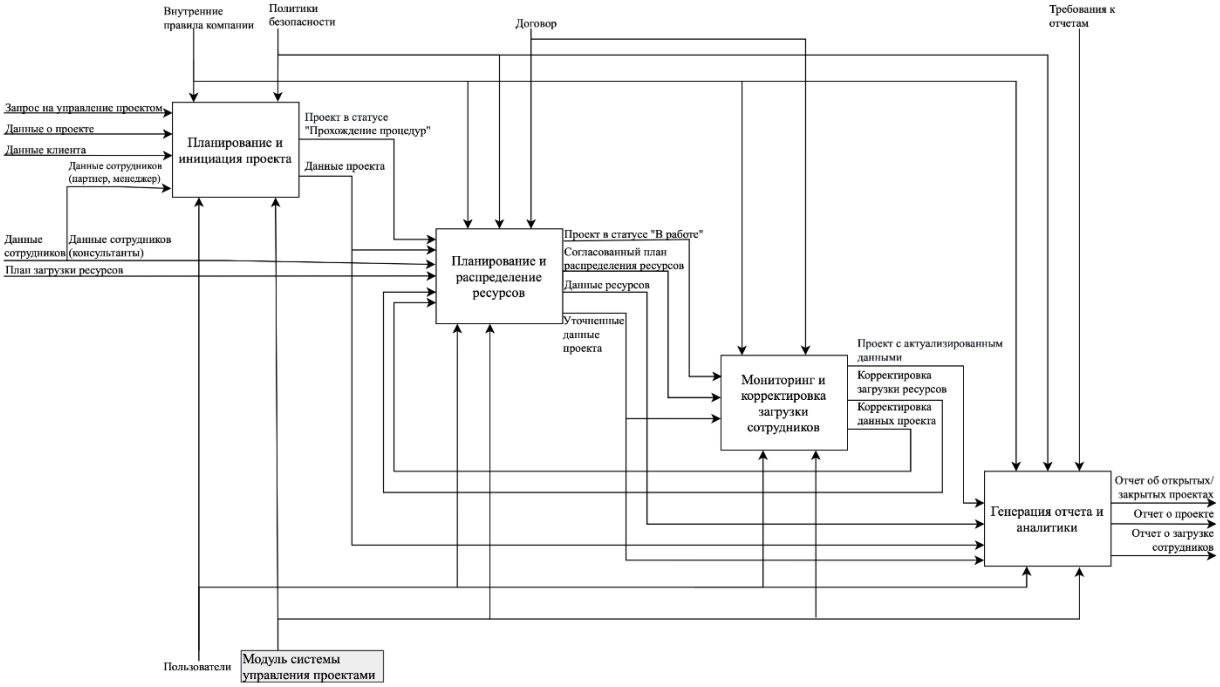
Г) оперативно выявлять конфликты ресурсов и риски перегрузки сотрудников;

Д) автоматизировать формирование отчетной документации и аналитических сводок.

4. Информационное и алгоритмическое обеспечение процесса управления проектами

Для детализации основного процесса представим декомпозицию функциональной модели модуля системы управления проектами в нотации IDEF0 (рисунок 1), выделив основные блоки:

1. Планирование и инициализация проектов - начальная стадия жизненного цикла проекта, на которой происходит создание проекта в системе и внесение его параметров.
2. Планирование и распределение ресурсов - детальное распределение загрузки членов проектной команды с указанием процентного соотношения их занятости на проекте во временные периоды (недели).
3. Мониторинг и корректировка загрузки сотрудников - отслеживание загрузки ресурсов.
4. Генерация отчета и аналитики - отчеты, представляющие состояние проектного портфеля, в том числе данные по каждому проекту.



Информационная поддержка процесса управления проектами

Рис. 1. Декомпозиция контекстной диаграммы «Управление проектами»

Основываясь на UML-диаграммы, был разработан алгоритм управления проектом, отображающий последовательностью действий в работе с проектом (рисунок 2).



Рис. 2. Алгоритм управления проектом

5. Архитектура модуля системы управления проектами

Модуль системы управления проектами состоит из четырех основных элементов: аутентификация; управление проектом; планирование ресурсов; отчетность (рисунок 3). Такая структура позволяет полностью учесть требования к модулю.



Рис. 3. Алгоритм управления проектом

Пользователь взаимодействует с клиентской частью через пользовательский интерфейс (UI), который отображается в веб-браузере. Интерфейс построен с использованием HTML, CSS и JavaScript, который загружается с веб-сервера и выполняется в браузере. Клиентская часть обрабатывает действия пользователя и отправляет запросы на серверную часть через REST API для выполнения операций с данными.

Работа с данными происходит с помощью взаимодействия с базой данных, которая служит основным хранилищем информации. Все данные сохраняются в реляционной базе данных, использующей ORM- Object-Relational Mapping. В качестве базы данных используется ProstgreSQL [9]. Использование SQLAlchemy позволяет эффективно работать с базой данных, упрощая создание, обновление и удаление записей.

6. Используемые технологии

Для разработки клиентской части модуля системы управления проектами был использован классический стек технологий: HTML, CSS и JavaScript [10]. В основе лежит фреймворк Flask, который обеспечивает серверный рендеринг страниц с помощью шаблонизатора Jinja2. Такой подход позволяет создавать динамические веб-страницы, которые формируются на сервере и затем отправляются в браузер пользователя, что обеспечивает хорошую SEO-оптимизацию и быструю первоначальную загрузку.

Для асинхронного взаимодействия с сервером применяется Fetch API, который позволяет выполнять HTTP-запросы к REST API, серверного приложения без перезагрузки страниц. Обмен данными происходит в формате JSON, обеспечивающем эффективную передачу структурированной информации.

Для работы с данными используется SQLAlchemy-инструмент объектно-реляционного отображения, который позволяет работать с ProstgreSQL как с коллекцией Python-объектов. Подключение к ProstgreSQL обеспечивается драйвером psycopg2, гарантирующим надежное и эффективное взаимодействие с базой данных.

Выбранный технологический стек позволяет эффективно реализовать бизнес-логику системы управления проектами.

7. Программная реализация

В прототипе модуля системы управления проектами реализован комплексный набор функций, охватывающий основные бизнес процессы управления проектами. Особое внимание было уделено функционалу планирования ресурсов, который составляет основу системы.

Система автоматически анализирует временные рамки проекта и генерирует календарную сетку с учетом рабочих дней. Алгоритм учитывает различные сценарии: проекты, начинающиеся или заканчивающиеся в середине недели, пересекающиеся границы месяцев; полные и неполные рабочие недели. Для каждого проекта создается уникальная временная структура, точно соответствующая его календарным рамкам. Основным инструментом планирования является интерактивная таблица, где отображаются сотрудники, назначенные на проект, и недельные интервалы в рамках проекта. Каждая ячейка на пересечении сотрудника и не недели представляет собой поле для ввода процента загрузки.

Пользователь может назначать сотрудников на проект, указывать процент загрузки для каждой недели, корректировать распределение ресурсов по мере использования планов, удалять сотрудников из проекта.

Пример модуля планирования загрузки на проекте представлен на рисунке 4.

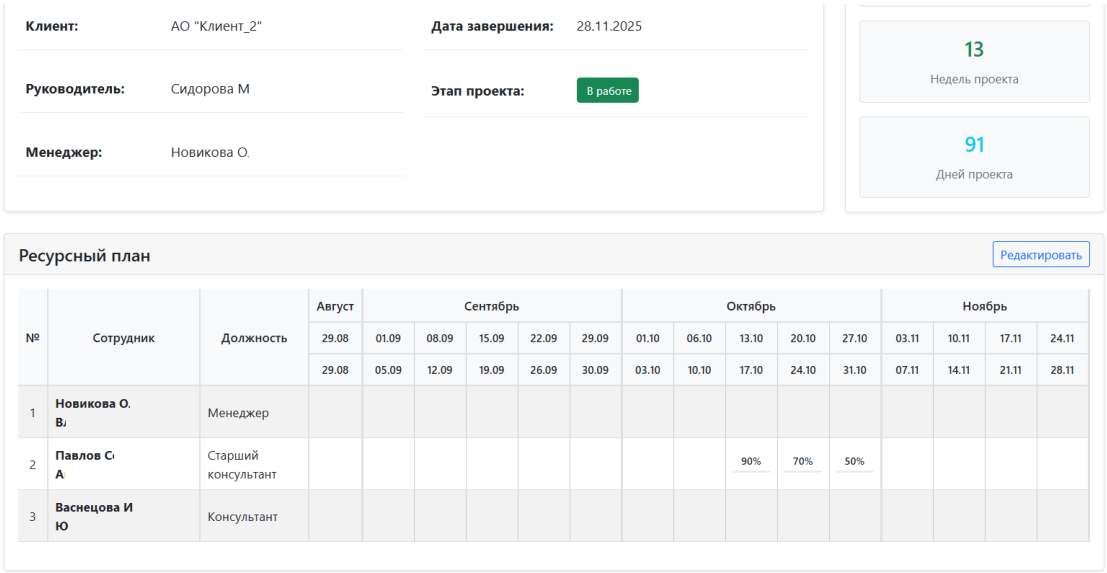


Рис. 4. Таблица «Ресурсный план» на детальной странице проекта

Эффективность управления проектами с использованием разработанного модуля предлагается рассматривать по основным типам работ: создание проекта, планирование ресурсов, корректировка планов загрузки, подготовка отчетности. На рисунке 5 представлена диаграмма сравнения временных затрат на управление проектом «До» и «После» внедрения модуля.



Рис. 5. Диаграмма сравнения временных затрат

Проведя анализ полученных данных, приходим к выводу, что, используя модуль системы управления проектами, удалось снизить временные затраты на процесс ведения проекта на 59%. Процесс планирования ресурсов и корректировки сократился на 44%.

8. Заключение

Разработанный программный модуль системы управления проектами позволяет автоматизировать ключевые процессы: создание проектов, назначение менеджеров, формирование команды, планирование и контроль загрузки сотрудников. Особое внимание в модуле уделено планированию ресурсов, которое представляет сложный детализированный инструмент для распределения человеческих ресурсов по проектам. Описаны технологии, позволяющие обеспечить эффективность внедрения модуля.

Проведенные экспериментальные исследования показали, что внедрение модуля системы управления проектами не только решает текущие операционные задачи, но и создает основу для долгосрочных улучшений в управлении проектной деятельностью. Сокращение времени на рутинные операции улучшает качество планирования, повышает общую эффективность использования человеческих ресурсов компании.

Список используемых источников

1. Тронина И. А. Политика импортозамещения информационных технологий и ее место в стратегическом развитии государственных компаний / И. А. Тронина, С. С. Лопухов // Комплексное развитие территориальных систем и повышение эффективности регионального управления в условиях цифровизации экономики: Материалы V Национальной (всероссийской) научно-практической конференции, Орел, 22 декабря 2022 года. – Орел: Орловский государственный университет имени И.С. Тургенева, 2023. - С. 270-277; URL:https://elibrary.ru/jpmumo?ysclid=mj1xtvrly f196133886 (дата обращения: 05.09.2025) (электронный ресурс).
2. Система управления проектами Spider Project. URL:https://www.spiderproject.ru/ru/publications/ (дата обращения: 07.09.2025). (электронный ресурс).
3. Модуль 1С:ERP (Управление проектами). URL:https://solutions.1c.ru/catalog/erp-pm-modul/ features?ysclid=mj1zan043j705183147 (дата обращения: 07.09.2025). (электронный ресурс).
4. Мегаплан. URL:https://megaplan.ru/crm-capabilities/projects-management/ (дата обращения: 07.09.2025). (электронный ресурс).
5. Информационная система управления проектами Advanta. URL:https://www.advanta-group.ru/about-system/vozmoznosti/ (дата обращения: 07.09.2025). (электронный ресурс).
6. Битрикс24.URL:https://www.bitrix24.ru/features/tasks/ (дата обращения: 07.09.2025). (электронный ресурс).
7. YouGile. URL:https://ru.yougile.com/product (дата обращения: 07.09.2025). (электронный ресурс).
8. Kaiten. URL:https://faq-ru.kaiten.site/ced4d01a-9da6-4fa0-8800-5de58b8859ad (дата обращения: 07.09.2025). (электронный ресурс).
9. PostgreSQL. URL:https://www.postgresql.org/ (дата обращения: 02.09.2025). (электронный ресурс).
10. Ковтуненко А. В., Орашанов М. С. Модуль поддержки принятия решений при составлении расписаний в университете // Информационные технологии интеллектуальной поддержки принятия решений (памяти проф. Н.И. Юсуповой) ITIDS'2024: Труды X Международной научной конференции. В 2-х томах, Уфа, 12-14 ноября 2024 года. - Уфа: Уфимский университет науки и технологий, 2024. - С. 41-47. URL:https://elibrary.ru/item.asp?id=80347176&pf f=1 (дата обращения: 02.09.2025). (электронный ресурс)

Оптимизация распределения оборотного капитала селлера на электронных маркетплейсах: алгоритмический подход на основе интегрального показателя

Бастанов А. Э.
ООО «Инвестиционное агентство «Градиент»
г. Уфа, Россия
gradinv@gmail.com

Ахметов И. В.
Уфимский государственный нефтяной технический университет,
г. Уфа, Россия
ilnurakhmetov@gmail.com

Аннотация[1]

В условиях интенсивного развития электронной коммерции и обострения конкуренции критически важным для продавцов (селлеров) становится эффективное управление оборотным капиталом при формировании бюджета закупок. Целью исследования является разработка алгоритма, позволяющего оптимизировать распределение ограниченных финансовых ресурсов между товарными позициями в портфеле селлера. В основе алгоритма лежит интегральный показатель – индекс доходности, позволяющий комплексно учитывать рентабельность и оборачиваемость товаров. В статье представлена пошаговая процедура алгоритма, проанализированы его ограничения, связанные с гипотезами о стабильности рыночных параметров, и предложены направления для дальнейшего совершенствования.

Ключевые слова: управление ассортиментом, оборотный капитал, маркетплейс, бюджет закупок, индекс доходности.

1. Введение

Современный этап развития розничной торговли характеризуется стремительной цифровизацией и ростом доли онлайн-каналов, особенно электронных маркетплейсов. Это привело к демократизации доступа на рынок для множества продавцов, что, в свою очередь, вызвало резкое обострение конкуренции. В таких условиях критически важным для выживания и развития бизнеса селлера становится эффективное управление внутренними ресурсами, среди которых наиболее значимым является оборотный капитал, направляемый на формирование товарных запасов.

Актуальность настоящего исследования обусловлена потребностью селлеров в практических инструментах, позволяющих не просто реагировать на рыночные колебания, но и проактивно распределять финансовые ресурсы для максимизации совокупной доходности товарного портфеля. Традиционные методы управления запасами зачастую не учитывают специфику высокодинамичной среды маркетплейсов, где ключевыми драйверами финансовых результатов выступают как рентабельность товара, так и его оборачиваемость. Возникает задача разработки специализированного алгоритма, который бы интегрировал эти два фактора в единую систему принятия решений при формировании бюджета закупок.

Проблематика управления товарным ассортиментом и запасами является классической для теории маркетинга и финансового менеджмента. Фундаментальные основы заложены в работах по ABC- и XYZ-анализу, которые позволяют классифицировать товары по вкладу в выручку и стабильности спроса [1]. Однако данные методы, будучи полезными для стратегического анализа, не предоставляют готового алгоритма для тактического распределения финансовых ресурсов в условиях их ограниченности.

Особое место в ряду методов стратегического управления ассортиментом занимают подходы, основанные на матричном анализе портфеля. Наиболее известной из них является матрица Boston Consulting Group (BCG), предлагающая категоризацию продуктов по четырем квадрантам на основе сочетания темпов роста рынка и относительной доли продукта [2]. Более совершенной с методической точки зрения представляется матрица «Привлекательность рынка/Конкурентоспособность» (McKinsey/ General

[1] Труды XI Международной научной конференции "Информационные технологии интеллектуальной поддержки принятия решений", 11-13 ноября, Уфа-Чандигарх-Фандербейлпарк, 2025

Electric), использующая многокритериальную оценку по обеим осям [3].

Впрочем, применение методологии матричного анализа в ее классическом виде сопряжено с рядом ограничений, наиболее существенными из которых в контексте управления оборотным капиталом селлера являются статичность и ретроспективность, сложность расчета отдельных показателей, а также слабая применимость для принятия тактических финансовых решений по управлению оборотным капиталом.

В практике управления производственными предприятиями известны следующие подходы к распределению оборотного капитала [4,5]:

1) метод на основе маржинальной рентабельности: приоритет отдается товарам с наивысшей маржой. Недостаток – игнорирование оборачиваемости товаров, что может привести к замораживанию капитала в товарах с низкой оборачиваемость;

2) метод на основе оборачиваемости: ресурсы в первую очередь направляются на самые «ходовые» товары. Риск – инвестиции в низкомаржинальные позиции, что не обеспечивает роста общей прибыли;

3) ABC–анализ в разрезе прибыли: товары категории А (приносящие основную прибыль) получают первоочередное финансирование. Однако внутри группы А также возникает задача выбора, для решения которой необходим более тонкий инструмент;

4) оптимизационные модели (например, метод множителей Лагранжа [6], модели математического программирования [7]): позволяют найти точку максимума общей прибыли при заданном бюджетном ограничении. Несмотря на математическую строгость, их практическое применение часто затруднено из-за сложности оценки точных функциональных зависимостей спроса для широкого ассортимента.

Таким образом, в существующих подходах наблюдается дефицит практичных, ориентированных на оперативное управление алгоритмов, которые бы синтезировали сильные стороны различных методов и были адаптированы к условиям работы селлеров на маркетплейсах.

2. Цель и задачи исследования

Цель исследования – разработка практического алгоритма для оптимального распределения ограниченного объема оборотного капитала селлера между товарными позициями при формировании бюджета закупок товаров для торговли на электронных маркетплейсах.

Задачи исследования:

1) предложить и обосновать использование интегрального критерия для оценки качества товарного портфеля селлера;

2) разработать пошаговый итеративный алгоритм распределения оборотных средств;

3) визуализировать предложенный алгоритм в виде блок–схемы;

4) провести анализ ограничений предложенного подхода и определить границы применимости предложенного алгоритма.

3. Научная новизна

Научная новизна работы заключается в следующем:

1) предложен алгоритм, ориентированный на решение задачи оперативного планирования закупок;

2) в качестве ключевого критерия приоритизации предложено использовать сводный критерий – индекс доходности;

3) алгоритм интегрирует в себе элементы портфельного анализа и итеративной процедуры распределения оборотных средств.

Практическая значимость работы состоит в создании формализованного инструмента для селлеров на электронных маркетплейсах, позволяющего:

1) оптимизировать структуру товарных запасов и бюджет закупок;

2) повысить эффективность использования оборотного капитала;

3) снизить риски замораживания средств в товарах с низкой оборачиваемостью;

4) принимать обоснованные управленческие решения на основе объективных расчетных показателей.

Алгоритм адаптирован к специфике электронной торговли и может быть внедрен без использования сложного программного обеспечения.

2. Алгоритм распределения оборотного капитала селлера на маркетплейсах

Портфель селлера на маркетплейсах характеризуется значительной разнородностью товарных позиций, каждая из которых демонстрирует различные значения ключевых операционных метрик и вносит неравномерный вклад в формирование конечных финансовых результатов. В условиях ограниченности оборотного капитала эффективное управление таким ассортиментом требует применения инструментов, позволяющих осуществлять сравнительный анализ товаров с учетом множества критериев.

В этом контексте для комплексной оценки товарных позиций предлагается использовать интегральный показатель – индекс доходности [8], рассчитываемый как произведение коэффициента

оборачиваемости на показатель маржинальной рентабельности. Данный сводный показатель позволяет преодолеть ограничения однокритериального подхода, обеспечивая сбалансированную оценку, учитывающую как прибыльность, так и скорость возврата вложенных средств.

Особую практическую ценность индекс доходности представляет при анализе товаров с разнонаправленными характеристиками – когда высокая маржинальная рентабельность компенсирует относительно низкую оборачиваемость, и наоборот. Такой подход идентифицировать товары со скрытым потенциалом доходности, которые могут остаться невыявленными при изолированной оценке показателей, и соответствует принципу Парето в управлении ассортиментом, где меньшая часть товарных позиций генерирует основную долю доходности. Это представляется особенно актуальным для селлеров, сталкивающихся с необходимостью оптимизации финансовых потоков в условиях высокой конкурентной среды маркетплейсов, где неоптимальное распределение ресурсов между товарными позициями может привести к существенным упущенным выгодам.

Предлагается следующий алгоритм распределения финансовых ресурсов селлера с использованием показателя индекса доходности:

I этап. Определяется объем свободных F_0 финансовых ресурсов селлера, который необходимо распределить на закуп товара.

Заполняется табл. 1, которая включает в себя показатели по каждому виду товаров из портфеля селлера, в том числе: показатели оборачиваемости (T_i), маржинальной рентабельности (M_i), прогнозный объем продаж (S_i), индекс доходности (R_i).

Таблица 1. – Исходные данные для анализа.

Наименование товара	T_i	M_i	S_i	R_i
Товар$_i$				

II этап. Товары упорядочиваются по критерию индекса доходности R_i, данные заносятся в табл. 2, проставляются приоритеты по мере убывания индекса доходности R_i.

Таблица 2. – Приоритетные товары по критерию индекса доходности.

Приоритет	Наименование товара	T_i	M_i	S_i	R_i
1					
…					
A					

III этап. На основании упорядоченного списка товаров, полученного в табл. 2, заполняется табл. 3, с указанием по каждому товару минимального запаса в натуральных показателях $N_{i\,min}$, закупочной стоимости единицы товара F_i и расчетного показателя стоимости минимального запаса по каждому товару $N_{i\,min} \cdot F_i$.

Таблица 3. – Расчет стоимости минимальных запасов по приоритетным товарам.

Приоритет	Наименование товара	$N_{i\,min}$	F_i	$N_{i\,min} \cdot F_i$
1				
…				
A				

Рассчитывается суммарная потребность Fmin в финансовых ресурсах на закупку минимального запаса товаров по формуле:

$$F_{min} = \sum_{i=1}^{A} N_{i\,min} \cdot F_i \qquad (1)$$

Рассчитывается величина F превышения объема свободных финансовых ресурсов F_0 над величиной F_{min} по формуле:

$$F = F0 - Fmin \qquad (2)$$

Индексу i присваивается значение 1.

IV этап. В табл. 3 выбирается товар с i-м приоритетом и определяется количество единиц для закупки при заданном объеме финансовых ресурсов F по формуле:

$$w = \text{округление}\left(\frac{F}{F_i}\right) \qquad (3)$$

Если $w < 1$, то переход на VIII этап, иначе переход на V этап.

V этап. Полученное значение w сравнивается с значением S_i и определяется объем закупки i-го товара N_i по следующему правилу:

если $N_{i\,min} + w \le S_i$, то $N_i = N_{i\,min} + w$, иначе $N_i = S_i$.

В табл. 4 вносятся значения величин N_i, $N_i \cdot F_i$.

Таблица 4. – Распределение финансовых ресурсов на закуп товаров.

Приоритет	Наименование товара	N_i	F_i	$N_i \cdot F_i$.
1				
…				
A				

VI этап. Определяется значение F по формуле:

$$F = F - F_i \cdot (N_i - N_{i\,min}) \qquad (4)$$

VII этап. Индексу i присваивается значение i+1, переход на IV этап.

VIII этап. Завершение распределения свободных финансовых ресурсов F_0 по критерию индекса доходности и заполнения табл. 4, на основании которой формируется бюджета закупок товаров.

Блок–схема алгоритма представлена на рисунке 1.

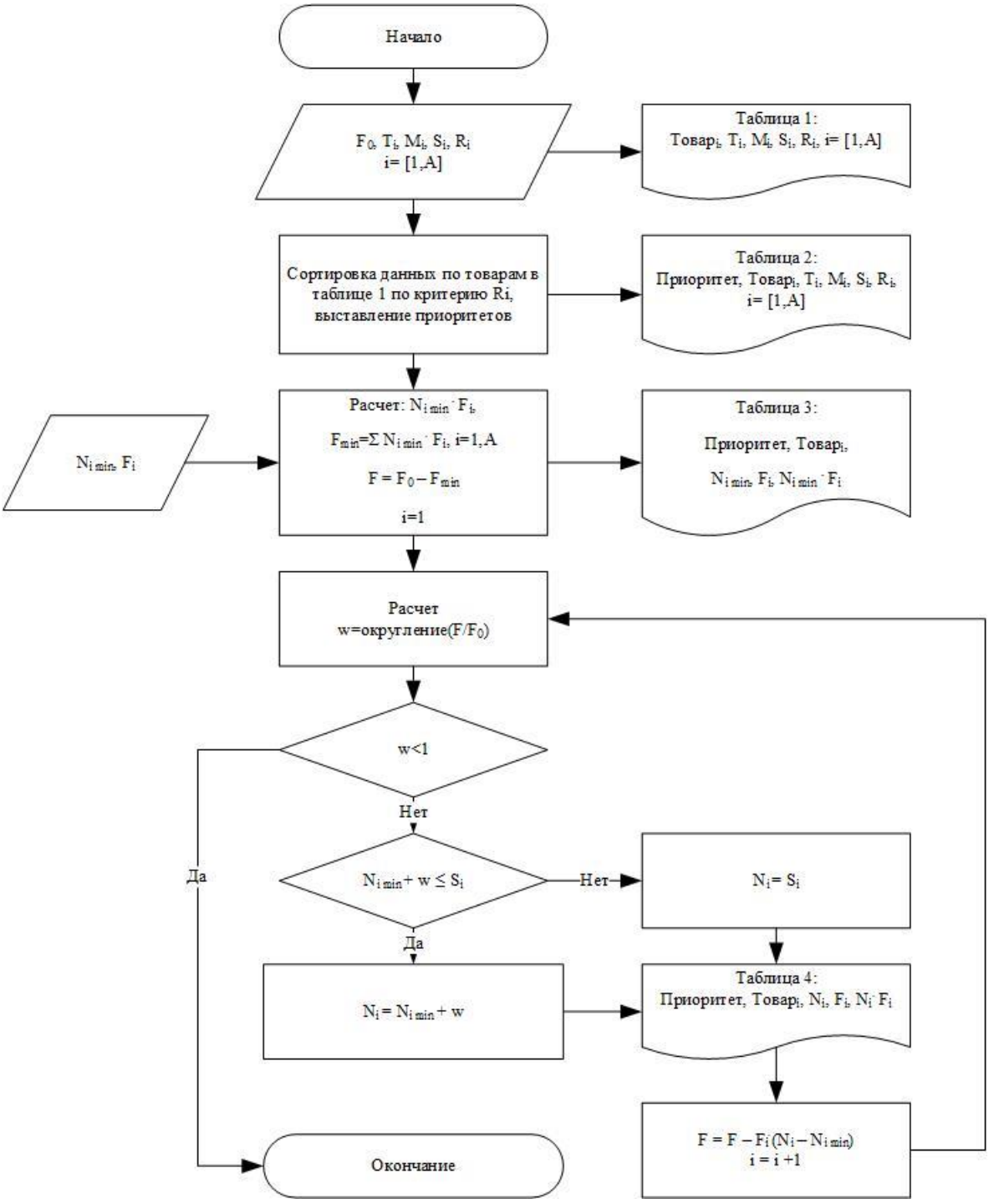


Рис. 1. Блок–схема алгоритма оптимального распределения ресурсов селлера.

3. Ограничения исследования и перспективы дальнейших работ

Несмотря на теоретическую обоснованность и практическую ценность, предложенный алгоритм обладает рядом ограничений, которые необходимо учитывать при его применении:

1) статичность параметров модели. Алгоритм оперирует детерминированными значениями показателей (T_i, M_i, S_i) на момент планирования. В реальных условиях эти параметры подвержены значительной волатильности под влиянием сезонности, изменения покупательского поведения, конкурентной среды и алгоритмов ранжирования маркетплейсов. Отсутствие вероятностной составляющей может привести к субоптимальным решениям при резких изменениях рыночной конъюнктуры;

2) одноэтапный характер модели. Критическим ограничением является одношаговая природа алгоритма, который оптимизирует распределение капитала только на один предстоящий период. Модель не учитывает:

– динамику изменения спроса в последующих периодах;
– влияние текущих решений по распределению капитала на будущие состояния товарного портфеля;
– накопительный эффект от реинвестирования полученной прибыли. Это может приводить к «эффекту запаздывания», когда решения, оптимальные в краткосрочной перспективе, оказываются неэффективными в долгосрочном горизонте планирования;
3) упрощенная модель потребительского спроса. В алгоритме предполагается, что спрос абсолютно неэластичен относительно уровня запасов и маркетинговой активности. Не учитывается возможность стимулирования продаж через рекламные бюджеты, а также эффект потерь от неудовлетворенного спроса при недостатке товарных запасов;
3) игнорирование логистических ограничений и издержек. Модель не учитывает:
– стоимость хранения на фулфилменте маркетплейса;
– объемно-весовые характеристики товаров;
– ограничения складских мощностей;
– транспортные издержки и минимальные партии поставок.
Для преодоления выявленных ограничений и повышения практической эффективности алгоритма предлагаются следующие направления исследований:
1) разработка многоэтапной динамической модели:
– создание многошаговой оптимизационной модели с конечным или бесконечным горизонтом планирования;
– применение методов динамического программирования и стохастической оптимизации;
– учет межвременных эффектов и возможности реинвестирования прибыли;
2) разработка адаптивных механизмов управления:
– интеграция с системами прогнозирования спроса на основе машинного обучения;
– реализация механизма регулярного пересчета параметров модели;
– введение корректирующих коэффициентов для учета сезонности;
3) учет логистических аспектов управления запасами:
– введение в модель ограничений на складские мощности;
– учет стоимости фулфилмента и логистики в расчете эффективности;
– оптимизация с учетом объемно-весовых характеристик товаров;
4) совершенствование модели спроса:
– учет эластичности спроса от уровня запасов и ценовой политики;
– интеграция с системой управления рекламными бюджетами;
– моделирование потерь от неудовлетворенного спроса;
5) верификация и апробация алгоритма:
– проведение A/B тестирования на реальных данных;
– сравнительный анализ с альтернативными методами распределения капитала;
– разработка программного обеспечения для автоматизации расчетов;
Перспективным направлением представляется создание комплексной системы поддержки принятия решений, объединяющей предложенный алгоритм с методами искусственного интеллекта для прогнозирования спроса и управления рисками в режиме, близком к реальному времени, с учетом многоэтапной природы процесса управления оборотным капиталом. Особое внимание следует уделить разработке моделей, учитывающих взаимовлияние решений в последовательные периоды и долгосрочные последствия текущих распределений капитала.
Представленные ограничения не снижают ценность полученных результатов, но определяют границы применимости предложенного алгоритма и формируют содержательную повестку для дальнейших исследований. Разработанная модель представляет значимость как концептуальная основа и отправная точка для создания более сложных и адаптивных систем управления оборотным капиталом.

4. Заключение

Проведенное исследование позволило разработать и теоретически обосновать алгоритм распределения оборотного капитала селлера на электронных

маркетплейсах, основанный на использовании индекса доходности.

Представленное решение вносит существенный вклад в развитие методологии интеллектуальной поддержки принятия решений в сфере электронной коммерции и имеет значительный потенциал для практического внедрения.

Ключевые преимущества разработанного алгоритма включают:

1) простоту реализации и низкие требования к вычислительным ресурсам, что делает его доступным для массового применения малыми предприятиями, составляющими основу продавцов на электронных маркетплейсах;

2) возможность интеграции в существующие корпоративные информационные системы без необходимости масштабной модернизации IT-инфраструктуры;

3) адаптивность к условиям высокой волатильности спроса и ограниченности финансовых ресурсов, характерных для малого бизнеса в электронной торговле.

Основные научно-практические результаты работы демонстрируют, что предложенный подход позволяет достичь баланса между методологической строгостью и практической применимостью. Выявленные ограничения, включая статичность параметров модели и ее одноэтапный характер, не снижают ценность полученных результатов, но определяют направления для дальнейшего совершенствования.

Перспективы практического внедрения и развития алгоритма связаны с:

– созданием специализированных SaaS-решений для малого и среднего бизнеса;

– разработкой мобильных приложений для оперативного управления бюджетом закупок;

– интеграцией с API маркетплейсов для автоматизации сбора исходных данных.

Эмпирическая верификация алгоритма на реальных данных и его сравнение с альтернативными методами распределения капитала составят предмет дальнейших исследований автора. Учитывая растущую конкуренцию на электронных маркетплейсах, представленная методика обладает значительным потенциалом для практического применения и может стать основой для создания стандартизированных решений в области управления оборотным капиталом селлеров.

Список используемых источников

1. Друри К. Управленческий и производственный учет: Пер. с англ.; Учебник. – М.: ЮНИТИ–ДАНА, 2002. – 1071 с.
2. Стратегии, которые работают: Подход BCG: сб. ст.: пер. с англ. / Сост. Карл Штерн и Джордж Сток-мл.; – 2-е изд.- М.: Манн, Иванов и Фербер, 2007. – 496 с.
3. Грант Р.М. Современный стратегический анализ. 5-е изд./ Пер. с англ. Под ред. В.Н. Фунтова. – СПб.: Питер, 2011. – 560 с.
4. Хомяченко О. В. Технология маркетинга на фирме. Экономический аспект. – М.: МЦФЭР, 1996.
5. Балашов В.Г., Ириков В.А. Технологии повышения финансового результата: практика и методы. – М.: МЦФЭР, 2006. – 672 с.
6. Интрилигатор М. Математические методы оптимизации и экономическая теория / Пер. с англ. Г.И. Жукововой, Ф.Я. Кельмана. – М.: Айрис-пресс, 2002. – 576 с.
7. Таха, Хемди А. Введение в исследование операций, 7–е издание.: Пер. с англ. – М.: Издательский дом «Вильямс», 2005. – 912 с.
8. Эффективное управление запасами / Джон Шрайбфедер; Пер с англ. – М.: Альпина Паблишер, 2016. – 304 с.

Анализ методов автоматизированного сбора пространственных данных телекоммуникационной инфраструктуры

Трубин В. Д.
УУНиТ, каф. ГИС
Уфа, Россия
e-mail: trubin.vd@ugatu.su

Христодуло О. И.
УУНиТ, каф. ГИС
Уфа, Россия
e-mail: o-hristodulo@mail.ru

Аннотация[1]

Современные телекоммуникационные сети сталкиваются с критическим несоответствием между высокой автоматизацией логической топологии и отсутствием системных методов реконструкции физической инфраструктуры. Отсутствие актуальной, структурированной документации, фрагментарность данных и нестандартные условия эксплуатации затрудняют диагностику, модернизацию и создание цифровых двойников.

Исследование направлено на разработку методологии автоматизированного формирования физических карт телекоммуникационных сетей в существующих зданиях, с акцентом на преодоление ограничений традиционного ручного аудита и интеграцию разнородных данных для создания валидной пространственной модели.

Предложен гибридный подход, объединяющий четыре метода сбора данных: использование BIM/CAD-моделей (как опорной основы), лазерное сканирование (для точной геометрии), фотограмметрию (для доступного и текстурированного сбора) и радиолокационное сканирование (для обнаружения скрытых кабельных каналов). Полученные данные обрабатываются с применением алгоритмов семантической сегментации и математической оптимизации, включая формирование cost field и генерацию топологических кандидатов для построения оптимальных маршрутов.

Доказана возможность эффективной интеграции данных из различных источников, что позволяет создавать полную, пространственно привязанную физическую карту сети. Выявлены сильные стороны каждого метода: лазерное сканирование — точность, фотограмметрия — доступность, UWB — проникновение сквозь конструкции, а оптимизация — автоматизация топологического проектирования.

Предложенный гибридный подход обеспечивает масштабируемое, точное и экономически обоснованное решение для формирования физических карт в реальных условиях, открывая путь к созданию цифровых двойников телекоммуникационной инфраструктуры. **Ключевые слова**: физическая карта сети; телекоммуникационные сети; цифровой двойник инфраструктуры; сбор пространственных данных.

1. Введение

Телекоммуникации играют важнейшую роль в процессах обмена информацией в любой информационной системе. Без физической связи между узлами невозможна передача данных. И несмотря на то, что современные среды передачи цифровой информации проектируются в соответствии с рекомендациями, предлагаемых управленческими, процессными и системными стандартами проектирования, таких как ITIL [1] или ГОСТ Р ИСО/МЭК 27001-2021 [2], все существующие стандарты носят рекомендательный характер. Нередко проектирование и реализация телекоммуникационной инфраструктуры выполняется исполнителем работ с учётом большого количества допущений, связанных с особенностями архитектуры пространства для инфраструктуры. К тому же, актуальной проблемой остаётся системное отсутствие полной, актуальной и структурированной документации по уже эксплуатируемым сетям, что обусловлено несистемным ведением технической документации, отсутствием требований к её обновлению в процессе модернизации и фрагментарностью данных, хранящихся в разрозненных форматах. Существующий ГОСТ Р 70444-2022 [3],

[1] Труды XI Международной научной конференции "Информационные технологии интеллектуальной поддержки принятия решений", 11-13 ноября, Уфа-Чандигарх-Фандербейлпарк, 2025

стандартизирующий ведение кабельного журнала слаботочных систем, регламентирует документацию маркировки каждого кабеля для их отслеживания, однако такой подход не даёт достаточно полной и наглядной картины для понимания мест проходов коммуникаций внутри сооружений. Эти факторы приводят к снижению надёжности диагностики, затрудняют эксплуатацию, модернизацию и восстановление сетей, а также ограничивают возможность моделирования цифровых двойников и автоматизированного управления инфраструктурой.

Стоит отметить, что следует проводить четкое разграничение между логическими и физическими связями в телекоммуникационных сетях. Для обозначения растровых или векторных изображений, визуализирующих эти связи, используется понятие карты сетей (network maps) [4]. В свою очередь карту сетей можно классифицировать как физическую (визуализирующую характеристики устройств на 1 уровне модели OSI) или логическую (визуализирующую характеристики устройств на 2-3 уровнях модели OSI). В данной работой под физической картой сети будет подразумеваться взаимное расположение устройств сети и каналов связи между ними с пространственной привязкой к системе координат. Под логической картой будет подразумеваться абстрактное представление взаимодействия сетевых устройств, отражающее топологию передачи данных и логические соединения.

Логические связи, определяемые протокольными взаимодействиями и топологией передачи данных, могут быть эффективно реконструированы на основе анализа трафика и идентификации оконечных устройств (MAC-адреса [5], IP-маршруты [6], SNMP [7]). В то же время построение точных физических карт требует детализированного учета реальных кабельных трасс, соединений в патч-панелях, распределительных шкафах и топографических особенностей инфраструктуры, что существенно усложняется из-за отсутствия стандартизированного учета физической топологии, несистемного обновления схем и физической недоступности ключевых узлов. Таким образом, автоматизация формирования логических карт реализуема на уровне сетевых протоколов, тогда как формирование физической топологии остаётся трудоёмкой, ручной задачей, требующей инвентаризации и визуального аудита инфраструктуры.

2. Актуальность задачи формирования физических карт телекоммуникационной инфраструктуры

Актуальность задачи формирования физических карт телекоммуникационной инфраструктуры обусловлена критическим несоответствием между высокой степенью автоматизации логической топологии и практически полным отсутствием системных методов реконструкции физической топологии существующих сетей. Разводки кабельных трасс на уровне поэтажных планов помещений остаются не доступны для автоматизированных систем из-за отсутствия единой методологии документирования, несогласованности данных между подразделениями организаций или низкой дисциплины обновления схем при модернизации. Это приводит к увеличению времени диагностики сбоев, риску ошибочного отключения критических каналов, снижению устойчивости сетей к авариям. В условиях роста сложности сетей и требований к бесперебойной работе критических инфраструктур, разработка методов автоматизированного, точного и масштабируемого формирования физических карт становится важной технической задачей для обеспечения надёжности, безопасности и управляемости цифровой среды передачи данных.

Следует отметить, что формирование физических карт существующих телекоммуникационных сетей и разработка проектной документации для проектируемых сетей – это не просто разные этапы жизненного цикла инфраструктуры, а принципиально различающиеся по методологии, исходным данным и целевым ограничениям задачи. Проектные планы строятся на основе заданных спецификаций, стандартов и конфигураций, тогда как реконструкция физической топологии существующих сетей требует восстановления неформализованных, часто не задокументированных или устаревших телекоммуникационных планов, характеризующихся несоответствиями между проектом и фактом, фрагментарностью данных и отсутствием централизованного учёта. В контексте формирования цифрового двойника инфраструктуры точность физической модели определяет её валидность для симуляции и прогнозирования отказов. В настоящем исследовании авторы сосредоточатся именно на задаче автоматизированного формирования физических карт существующих телекоммуникационных сетей.

3. Методы формирования логической карты телекоммуникационных сетей

В рамках решения задачи формирования логической топологии телекоммуникационной сети необходимо определить минимальный набор данных, достаточных для её корректного восстановления. При условии функциональной исправности сети и корректного обмена IP- и L2-трафиком между узлами, логические связи могут быть идентифицированы на основе наблюдаемых взаимодействий — без прямого доступа к физической инфраструктуре. Задача автоматического формирования логической карты, хотя и формально разрешима, представляет собой сложную многофакторную проблему, допускающую множество подходов, выбирающихся в зависимости от условий сбора данных, а также состояния и архитектуры телекоммуникационных сетей. Современные методы включают в себя данные из следующих источников:

- Протоколы обнаружения соседей (LLDP [8], CDP [9]) — для восстановления L2-связей между коммутаторами и шлюзами;
- Протоколы маршрутизации (OSPF [10], BGP [11]) — для построения L3-топологии и выявления маршрутизирующих узлов;
- Анализ трафика (NetFlow [12], IPFIX [13]) — для идентификации логических потоков и зависимостей между хостами;
- Конфигурационные данные (SNMP [14], NETCONF [15], RESTCONF [16]) — для извлечения интерфейсов, VLAN, ACL и статических маршрутов;
- Системы управления конфигурацией (CMDB) — для агрегации метаданных и привязки логических связей к инвентарным объектам;
- OpenFlow Discovery Protocol — доступен в сетях, построенных по принципам SDN, однако обеспечивающий централизованному контроллеру детальную, актуальную и унифицированную информацию о топологии на уровне портов и связей, в реальном времени и без зависимости от проприетарных протоколов [17,18]. Использование OFDP позволяет достичь максимальной полноты и точности реконструкции, особенно в динамических и виртуализированных средах.

Выбор источников данных будет определяться архитектурой сети, доступностью протоколов и требованиями к точности. В качестве примера решения задачи, авторами в работе [19] описан вариант использования систем мониторинга для сбора и хранения логических карт телекоммуникационных сетей в базе пространственных данных.

4. Методы формирования физической карты телекоммуникационных сетей

Также, для решения поставленной задачи требуется набор пространственных данных, включающий актуальные планы помещений и координаты размещения телекоммуникационных инфраструктур. Ключевые трудности сбора координат объектов инфраструктур обусловлены динамичностью пространственной информации, гетерогенностью условий сбора (методы, масштабы), разнообразием архитектурных решений, а также отсутствием стандартизированных инструментов и протоколов для систематизированного сбора, интеграции и долгосрочного хранения данных [20]. Вопросы хранения описаны авторами в [21], поэтому в данной работе рассмотрим основные методы автоматизированного сбора данных для формирования физических карт.

- Использование существующих CAD-файлов, BIM-моделей и архитектурных чертежей поэтажных планов. Прямых нормативных требований к формату хранения или обязательной периодической актуализации архитектурных планов внутри организаций в общем виде законодательством не установлено. Регулирование в основном касается этапов проектирования и сдачи объектов в эксплуатацию, поэтому использование этого метода сбора данных ограничено относительно новыми сооружениями. С 2024 года в РФ введено обязательное применение BIM-технологий для капитального строительства [22], что упростит процесс формирования базы пространственных данных для новых сооружений.
- Метод лазерного сканирования использует лазерные лучи для создания трёхмерного облака точек и обеспечивает миллиметровую точность [23]. Технология широко используется для создания BIM-моделей и позволяет неинвазивно, без доступа к внутренним конструкциям, зафиксировать геометрию объектов реального мира даже в труднодоступных зонах, что особенно актуально для телекоммуникационных сетей, размещаемых за фальшпотолками или в коробах [24].
- Метод фотограмметрии подразумевает построение трёхмерной модели объекта из серии фотографий, снятых с разных углов. Фотографии обрабатываются программными алгоритмами (SIFT, SFM, MVS) для реконструкции 3D-геометрии пространства [25]. Оборудование более доступно в сравнении с лазерами, что обеспечивает низкую стоимость и высокую масштабируемость сбора данных.
- Метод радиолокационного сканирования использует сигналы Ultra-Wideband или Wi-Fi Round-Trip Time для триангуляции позиций устройств и активно применяется в сферах

интернета вещей, медицины и автомобилестроения [26]. Алгоритмы обработки сигнала позволяют формировать 2D/3D-карты пространства с высокой точностью позиционирования, не требуя прямой видимости. Метод особенно ценен для скрытого обследования, динамических сред и объектов с ограниченным доступом, обладает точностью в несколько сантиметров.

Сравнительный анализ и области применения описанных методов приведены в таблице 1.

Для формирования точных цифровых планов помещений и сетевых топологий необходимо интегрировать данные из разнородных источников, однако их сбор и систематизация сталкиваются с фундаментальными вызовами: нестабильностью данных, отсутствием унифицированных стандартов и инструментов. Перспективным подходом является **гибридная стратегия**, сочетающая существующие BIM-данные с автоматизированным сбором данных с помощью вышеперечисленных методов, что обеспечит полноту, актуальность и масштабируемость цифровых моделей помещений, как показано на Рис. 1. При наличии CAD/BIM-моделей их возможно использовать в качестве опорных моделей для регистрации дополнительных данных, а также как источник семантики. Лазерное сканирование обеспечивает получение плотных трёхмерных облаков точек с миллиметровой точностью, что делает его основным источником геометрической информации для формирования точной пространственной модели внутренних помещений. Фотограмметрия позволяет быстро собрать данные и стать первичным источником данных семантической сегментации после обработки снимков методами машинного обучения. Радиолокационное сканирование позволяет обнаруживать скрытые металлические структуры — такие как кабельные лотки и трубы — за непрозрачными конструкциями (бетон, гипсокартон), обеспечивая точную локализацию телекоммуникационной инфраструктуры, недоступной для оптических методов.

Следующей задачей, которую необходимо решить для построения физических карт сети на поэтажных планах помещений, является обработка полученных данных.

Таблица 1. – Методы автоматизированного формирования физических карт

Критерий	Существующие CAD/BIM-модели	Лазерное сканирование
Краткое описание метода	Использование готовых цифровых моделей, созданных на этапе проектирования	Использование лазерных лучей в помещении для формирования плотного трёхмерного облака точек
Точность геометрии выходных данных	Зависит от исходного проектирования модели	Миллиметровая. Не зависит от внешних факторов
Затраты	Высокие в случае создания моделей с нуля, низкие если они предоставлены	Высокие. Дорогостоящее оборудование и оператор
Скорость сбора данных	Мгновенно при наличии моделей, при создании с нуля низкая	Быстро. 1-5 минут на помещение
Привязка к системе координат	Да	Да. Применяются алгоритмы ICP [27] и SLAM [28]
Семантические данные	Элементы размечены вручную	Нет. Требуется постобработка для классификации
Доступность скрытых элементов	Зависит от проектирования	Нет

Продолжение таблицы 1

Критерий	Фотограмметрия	Радиолокационное сканирование
Краткое описание метода	Построение трёхмерной модели на основе анализа серии перекрывающихся изображений	Использование радиоволн UWB для обнаружения отражений от металлических и диэлектрических структур
Точность геометрии выходных данных	Сантиметровая. Зависит от внешних факторов	Сантиметровая. Не зависит от внешних факторов
Затраты	Низкие, доступное оборудование	Умеренные, оборудование не такое дорогое
Скорость сбора данных	Умеренно. 10-30 минут съёмка и обработка	Умеренно. 5-20 минут сканирование и обработка
Привязка к системе координат	Да. Применяется алгоритм SFM [29]	Да. Применяются анкерные точки [30]
Семантические данные	Текстура позволяет идентифицировать объект, желательна постобработка	Нет. Требуется интерпретация по характеристикам сигнала
Доступность скрытых элементов	Нет	Да. Позволяет обнаруживать скрытые элементы

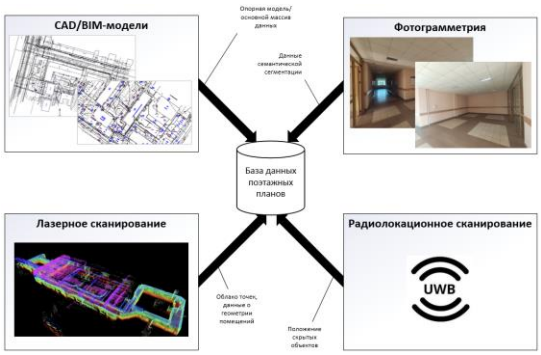


Рис. 1. Схема сбора данных гибридным подходом

5. Анализ методов автоматизированного формирования физических карт в смежных областях

В работе [31] предложен метод извлечения информации о телекоммуникационных сетях из двухмерных чертежей в области судостроения. Используется предобученная нейронная сеть для обнаружения путей кабелей, решается задача классификации изображения. Помимо линейных объектов, алгоритм считывает текстовую документацию и формирует массив атрибутивных данных без пространственной привязки. Этот метод возможно адаптировать для решения поставленной задачи, при наличии существующих чертежей планов разводки кабельной системы организации.

В работе [32] авторами предлагается метод автоматического создания путей прокладки кабелей в автомобильной технике на сеточной модели с введением полей стоимости, полностью основанный на математическом аппарате. Предложенный подход можно адаптировать для задачи определения физического расположения телекоммуникационных сетей. Геометрия, полученная на этапе сбора данных с помощью любого обозначенного метода, преобразуется в дискретную сетку, где стоимость каждой ячейки определяется вероятностью её принадлежности к кабельному каналу, а также наличием препятствий и требований к минимальным радиусам изгиба. Затем, с использованием метода множителей Лагранжа и субградиентной оптимизации генерируется множество топологических кандидатов, которые затем анализируются на соответствие инженерным ограничениям. Также, алгоритм позволит сгенерировать множество топологически различных кандидатов, что позволяет учесть неопределённость в инженерных требованиях.

В работе [33] рассматривается применение технологий глубокого обучения по отношению к облакам точек для решения разного рода задач в области строительства. В работе [34] приводится решение задачи семантической сегментации облака точек с использованием свёрточных нейронных сетей. Оба метода подходят для семантической сегментации полученных с помощью лазерного сканирования и фотограмметрии данных, что автоматизирует ввод физической карты сети в базу данных, однако потребует дополнительной ручной проверки и внесения корректировок.

Исследование [35] демонстрирует использование методов машинного зрения для распознавания образов на цифровых планах электрических коммуникаций, что в последствии используется для анализа сети в целом. Предложенный метод может быть адаптирован для обучения свёрточной нейронной сети на существующих физических картах телекоммуникационных сетей, а полученная модель использована для предсказания пролегания

кабелей сети на планах помещений, для которых это требуется.

Также, необходима информационная система, которую возможно будет использовать для интеграции собранных данных о логических и физических связях телекоммуникационной сети. Возможно использование CAD-систем, однако это ограничивает дальнейшего потребителя полученных данных в гибкости взаимодействия с ними. Поэтому, предлагается использовать геоинформационные системы, предоставляющие широкий набор инструментов для пространственного и атрибутивного анализа собранных данных, а также их корректировки, чтоб подробно описано авторами в работе [36]. Схема используемых данных для интеграции и визуализации карт телекоммуникационных сетей в геоинформационной системе представлена на Рис. 2.

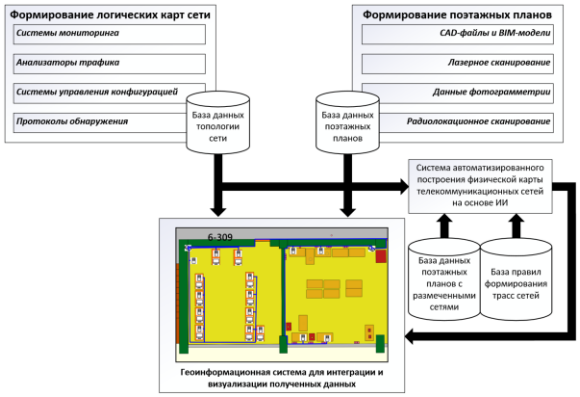


Рис. 2. Схема используемых данных для интеграции карт телекоммуникационных сетей в ГИС

Список используемых источников

1. Axelos. ITIL 4 Foundation. 2nd изд. 2019.
2. ГОСТ Р ИСО/МЭК 27001-2021 Информационная технология (ИТ). Методы и средства обеспечения безопасности. Системы менеджмента информационной безопасности. Требования (Переиздание) от 30 ноября 2021 - docs.cntd.ru [Электронный ресурс]. URL: https://docs.cntd.ru/document/1200181890 (дата обращения: 09.10.2025).
3. ГОСТ Р 70444-2022. Слаботочные системы. Кабельные системы. Кабельный журнал. Требования к составу и содержанию: Стандарт 70444–2022. Москва: Российский институт стандартизации, 2023.
4. CNSSI No. 4009. Committee on National Security Systems, 2015.
5. Institute of Electrical and Electronics Engineers. IEEE standard for local and metropolitan area networks: overview and architecture. New York: Institute of Electrical and Electronics Engineers, 2002. 1 с.
6. Internet Protocol: Request for Comments RFC 791. Internet Engineering Task Force, 1981. 51 с.
7. Fedor M. и др. Simple Network Management Protocol (SNMP): Request for Comments RFC 1157. Internet Engineering Task Force, 1990. 36 с.
8. Krishnan S. и др. Link-layer event notifications for detecting network attachments: 4957. RFC Editor, 2007.
9. Cisco Systems. Cisco discovery protocol configuration guide, cisco IOS XE... [Электронный ресурс]. 2015. URL: https://www.cisco.com/c/en/us/td/docs/ios-xml/ios/cdp/configuration/xe-16-12/cdp-xe-16-12-book.pdf (дата обращения: 14.10.2025).
10. Moy J. OSPF version 2: 2328. RFC Editor, 1998.
11. Rekhter Y., Hares S., Li T. A border gateway protocol 4 (BGP-4): 4271. RFC Editor, 2006.
12. Claise B. Cisco systems NetFlow services export version 9: 3954. RFC Editor, 2004.
13. Leinen S. IP flow information export (IPFIX) over TCP: Internet-Draft draft-leinen-ipfix-tcp-01. Internet Engineering Task Force / Internet Engineering Task Force, 2004.
14. User-based security model (USM) for version 3 of the simple network management protocol (snmpv3): 2274. RFC Editor, 1998.
15. Enns R. и др. Network configuration protocol (NETCONF): 6241. RFC Editor, 2011.
16. Bierman A., Björklund M., Watsen K. RESTCONF protocol: 8040. RFC Editor, 2017.
17. Alhanani R.A., Abouchabaka J. An overview of different techniques and algorithms for network topology discovery // 2014 Second World Conference on Complex Systems (WCCS). 2014. С. 530–535.
18. Zacharis A. и др. Performance evaluation of topology discovery protocols in software defined networks // 2021 IEEE Conference on Network Function Virtualization and Software Defined Networks (NFV-SDN). Heraklion, Greece: IEEE, 2021. С. 135–140.
19. Проектирование базы данных системы мониторинга вычислительной сети с применением ГИС-технологий / Р. В. Иликбаева, В. Д. Трубин, Н. А. Мартымянов, О. И. Христодуло // Наукосфера. – 2024. – № 5-2. – С. 6-13. – DOI 10.5281/zenodo.11261084. – EDN RJNQUM.

20. Zlatanova S. и др. Problems In Indoor Mapping and Modelling // Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. 2013. Т. XL-4/W4. С. 63–68.

21. Информационная поддержка управления компьютерной инфраструктурой сложных распределенных систем / Е. С. Брекоткина, А. С. Павлов, С. В. Павлов [и др.] // Вестник компьютерных и информационных технологий. – 2022. – Т. 19, № 3(213). – С. 23-32. – DOI 10.14489/vkit.2022.03.pp.023-032. – EDN CKZTLG.

22. Постановление Правительства Российской Федерации от 05.03.2021 № 331 «Об установлении случая, при котором застройщиком, техническим заказчиком, лицом, обеспечивающим или осуществляющим подготовку обоснования инвестиций, и (или) лицом, ответственным за эксплуатацию объекта капитального строительства, обеспечиваются формирование и ведение информационной модели объекта капитального строительства».

23. Середович В.А. Наземное лазерное сканирование. Новосибирск: Сибирский государственный университет геосистем и технологий, 2009. 261 с.

24. ГОСТ Р 53245-2008. Информационные технологии. Системы кабельные структурированные. Монтаж основных узлов системы. Методы испытания. Москва: Стандартинформ, 2008.

25. Pepe M., Costantino D. Techniques, Tools, Platforms and Algorithms in Close Range Photogrammetry in Building 3D Model and 2D Representation of Objects and Complex Architectures // Comput.-Aided Des. Appl. 2020. Т. 18, № 1. С. 42–65.

26. Cheraghinia M. и др. A Comprehensive Overview on UWB Radar: Applications, Standards, Signal Processing Techniques, Datasets, Radio Chips, Trends and Future Research Directions // IEEE Commun. Surv. Tutor. 2025. Т. 27, № 4. С. 2283–2324.

27. Besl P.J., McKay N.D. A method for registration of 3-D shapes // IEEE Trans. Pattern Anal. Mach. Intell. 1992. Т. 14, № 2. С. 239–256.

28. Durrant-Whyte H., Bailey T. Simultaneous localization and mapping: part I // IEEE Robot. Autom. Mag. 2006. Т. 13, № 2. С. 99–110.

29. Lowe D.G. Distinctive Image Features from Scale-Invariant Keypoints // Int. J. Comput. Vis. 2004. Т. 60, № 2. С. 91–110.

30. Cho J., Lee B. Optimal layout of four anchors to improve accuracy of Ultra-Wide band based indoor positioning // Expert Syst. Appl. 2024. Т. 255. С. 124808.

31. Putra A.R., Ha S., Park K.-P. Automatic extraction of cable connection information from 2D drawings for electrical outfittings design in shipyards // Int. J. Nav. Archit. Ocean Eng. 2024. Т. 16. С. 100630.

32. Karlsson T. и др. Automatic Cable Harness Layout Routing in a Customizable 3D Environment // Comput.-Aided Des. 2024. Т. 169. С. 103671.

33. Yue H. и др. Deep learning applications for point clouds in the construction industry // Autom. Constr. 2024. Т. 168. С. 105769.

34. Babacan K., Chen L., Sohn G. SEMANTIC SEGMENTATION OF INDOOR POINT CLOUDSUSING CONVOLUTIONAL NEURAL NETWORK // ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci. 2017. Т. IV-4/W4. С. 101–108.

35. Dosi D. и др. SkeySpot: Automating Service Key Detection for Digital Electrical Layout Plans in the Construction Industry: arXiv:2508.10449. arXiv, 2025.

36. Христодуло О.И., Павлов С.В., Трубин В.Д. Обеспечение полноты и актуальности пространственной информации в распределенных вычислительных системах больших организаций // Программные Продукты И Системы. 2024. № 2. С. 221–229.

Playwright как универсальный инструмент тестирования веб-приложений: особенности и преимущества

Пильник А. С.
Уфимский университет науки
и технологий
Институт информатики, математики
и робототехники
Уфа, Россия
e-mail: katkov2010anton@mail.ru

Ризванов Д. А.
Уфимский университет науки
и технологий
Институт информатики, математики
и робототехники
Уфа, Россия
ridmi@mail.ru

Аннотация[1]

Современные практики разработки ПО всё чаще требуют автоматизированных решений для обеспечения качества — особенно на фоне ускоряющихся циклов релизов и усложнения веб-интерфейсов. Ручное тестирование в таких условиях уже не справляется с объёмом и скоростью необходимых проверок. В статье рассматривается фреймворк Playwright, созданный Microsoft, как эффективное средство для сквозной автоматизации тестирования веб-приложений. Автор анализирует архитектурные особенности инструмента, его ключевые функции и преимущества перед такими решениями, как Selenium, Cypress и Puppeteer. Отдельное внимание уделено интеграции Playwright в DevOps-процессы, поддержке кроссбраузерного и мобильного тестирования, а также возможностям параллельного запуска и детальной трассировки сценариев. Показано, что применение Playwright позволяет не только ускорить тестовые циклы, но и повысить надёжность выявления ошибок даже в динамически меняющихся интерфейсах. В совокупности эти качества делают Playwright важным элементом современной инфраструктуры контроля качества программного обеспечения.

Ключевые слова: автоматизация тестирования; Playwright; Cypress; Selenium; Puppeteer; e2e-тестирование; CI/CD

1. Введение

Современная разработка программного обеспечения характеризуется высокой динамикой изменений и сокращением времени между релизами. В этих условиях автоматизация тестирования становится важнейшим элементом обеспечения качества. Традиционные ручные методы тестирования, несмотря на гибкость, не позволяют поддерживать требуемую скорость и масштаб проверок. В результате компании переходят к построению систем автоматизированного контроля качества на всех стадиях жизненного цикла продукта [1].

Появление инструментов автоматизации изменило подход к QA. Сначала преобладали решения с графическим интерфейсом (например, QTP, TestComplete), затем внимание специалистов сместилось к фреймворкам с программными API. Одним из первых таких инструментов стал Selenium WebDriver, заложивший стандарты взаимодействия с браузером через драйвер [2]. Однако со временем появились новые вызовы — необходимость кроссбраузерности, поддержки мобильных устройств, асинхронных сценариев и интеграции в CI/CD.

На этом фоне компания Microsoft представила Playwright — инструмент нового поколения, объединивший лучшие практики автоматизации. Его архитектура ориентирована на производительность, параллельность и стабильность выполнения тестов. Благодаря универсальности и поддержке языков программирования (JavaScript, Python, Java, C#), Playwright быстро стал стандартом де-факто в современной автоматизации тестирования [3].

2. Современные тенденции автоматизации тестирования

Рост числа веб-приложений, микросервисов и облачных решений обострил потребность в системах, способных обеспечивать проверку функционала в условиях высокой изменчивости интерфейса и данных. Традиционные инструменты часто не справлялись с асинхронностью современных приложений (SPA, React, Angular), а тесты становились хрупкими и трудными в сопровождении [4].

[1] Труды XI Международной научной конференции "Информационные технологии интеллектуальной поддержки принятия решений", 11-13 ноября, Уфа-Чандигарх-Фандербейлпарк, 2025

В настоящее время можно выделить несколько тенденций:

1. Переход к “shift-left testing” — раннее тестирование на стадии разработки.
2. Интеграция тестов в DevOps-процессы (CI/CD).
3. Использование облачных платформ тестирования.
4. Применение паттернов проектирования — Page Object, Data-Driven, BDD.
5. Рост роли e2e-тестирования как средства контроля пользовательского опыта [1,4].

Playwright органично вписывается в эти тенденции, позволяя писать стабильные тесты даже для динамически обновляемых интерфейсов, не требуя внешних драйверов и легко интегрируясь с DevOps-практиками [5,3].

3. Эволюция инструментов: от Selenium до Playwright

Исторически автоматизация тестирования начиналась с Selenium, предложившего унифицированный API для работы с браузерами [2]. Однако его архитектура, основанная на взаимодействии с WebDriver через HTTP, ограничивала скорость и гибкость.

Позднее появился Cypress, который предложил выполнение тестов внутри браузера, устраняя часть задержек. Однако Cypress ограничен экосистемой JavaScript и не обеспечивает кроссбраузерной совместимости [4,6].

Puppeteer, разработанный Google, стал первым инструментом на базе DevTools-протокола, что обеспечило высокую скорость и точность выполнения. Но Puppeteer ориентирован только на браузеры Chromium, что снижает его универсальность [6].

Playwright объединил лучшие идеи предшественников — DevTools-протокол, кроссбраузерную поддержку, изоляцию контекста, простую настройку и поддержку CI/CD. В результате он стал универсальным решением, сочетающим стабильность Selenium, скорость Puppeteer и удобство Cypress [5,6,3].

4. Архитектура и принципы работы Playwright

В основе Playwright лежит концепция изолированных браузерных контекстов. Каждый тестовый сценарий запускается в собственной «песочнице» — отдельной сессии браузера с независимыми cookies, localStorage и кэшем. Такой подход исключает пересечение состояний между тестами и особенно полезен при параллельном выполнении [3].

Ключевые компоненты архитектуры включают:

- BrowserContext — изолированная среда сессии;
- Page — программное представление веб-страницы;
- Locator API — механизм поиска элементов, учитывающий динамику интерфейса;
- Fixtures — вспомогательные конструкции для настройки окружения перед запуском теста;
- Trace Viewer — инструмент для визуального анализа хода выполнения сценария;
- Встроенный тест-раннер — средство для запуска, сбора результатов и формирования отчётов.

Среди преимуществ архитектуры — встроенная поддержка асинхронных операций, автоматические ожидания элементов (без необходимости ручного добавления таймаутов), утверждения, ориентированные на веб-поведение (web-first assertions), а также встроенная возможность параллельного выполнения и шардирования тестов. Кроме того, Playwright позволяет эмулировать мобильные устройства и различные сетевые условия, что делает его гибким инструментом для реалистичного тестирования пользовательского опыта.

5. Проектирование тестов и шаблоны

Одним из устоявшихся подходов к организации автоматизированных тестов является шаблон Page Object Model (POM). Он позволяет отделить логику взаимодействия с интерфейсом от самих тестовых сценариев, упрощая поддержку и масштабирование кодовой базы [4]. Playwright поддерживает этот паттерн «из коробки» — благодаря чистому и лаконичному API реализация POM получается интуитивно понятной даже в крупных проектах.

Например, страница авторизации может быть оформлена как отдельный класс в соответствии с Рисунком 1:

```python
class LoginPage:
    def __init__(self, page: Page):
        self.page = page

    def enter_username(self, username: str):
        self.page.fill('input[name="username"]', username)

    def enter_password(self, password: str):
        self.page.fill('input[name="password"]', password)

    def click_login_button(self):
        self.page.click('button[type="submit"]')

    def is_error_message_displayed(self):
        return self.page.is_visible('.error-message')
```

Рис. 1. Страница авторизации как класс

Кроме того, Playwright поддерживает Data-Driven Testing и Behavior-Driven Development (BDD) через интеграцию с pytest-bdd и playwright-bdd, что

позволяет описывать сценарии на языке Gherkin [5,3].

6. Сравнительный анализ производительности

Согласно исследованиям [6], Playwright демонстрирует лучшие результаты среди четырёх фреймворков (Selenium, Cypress, Puppeteer, Playwright). В трёх сценариях — от простого входа до комплексного набора e2e-тестов — Playwright показал среднее время выполнения на 25–40 % ниже, чем Cypress, и на 15 % ниже, чем Puppeteer.

Для простых сценариев (логин, форма) Playwright оказался в 3–4 раза быстрее Cypress, а при сложных сценариях сохранял преимущество в стабильности. Selenium уступал по скорости и требовал больше ресурсов [6,2].

Playwright также выделяется отсутствием внешних зависимостей, упрощённым CI/CD-внедрением и эффективным использованием headless-режима [3].

7. Интеграция Playwright с DevOps и CI/CD

Playwright изначально разработан для интеграции с конвейерами DevOps [5,3]. Он поддерживает GitHub Actions, Jenkins, GitLab CI и Azure DevOps. Запуск тестов осуществляется одной командой в соответствии с Рисунком 2:

```
npx playwright test --reporter=html
```

Рис. 2. Запуск тестов

Playwright создаёт детализированные отчёты, включая trace-файлы, которые можно открыть в Trace Viewer для отладки [5].

Это делает Playwright частью полноценного DevOps-конвейера, обеспечивая непрерывное тестирование (Continuous Testing) и быстрое выявление дефектов.

8. Реинжиниринг QA-процессов с Playwright

Реинжиниринг QA-процессов предполагает переход от разрозненных проверок к системному, автоматизированному подходу [1]. Playwright способствует этому благодаря:

- стандартизации тестовых сценариев;
- повышению стабильности и прозрачности отчётности;
- снижению времени обратной связи;
- уменьшению затрат на сопровождение тестов;
- интеграции с CI/CD и BDD.

По данным [1,5], внедрение Playwright сокращает длительность тестового цикла на 30–40 %, снижает количество ручных проверок и повышает точность детектирования дефектов.

9. Практические рекомендации

Для эффективного внедрения Playwright рекомендуется:

1. Использовать POM и выделять бизнес-логику.
2. Применять параметризацию окружений.
3. Использовать фикстуры для подготовки данных.
4. Вести централизованный лог тестов.
5. Интегрировать тесты в CI/CD.
6. Применять визуальное тестирование и скриншоты.
7. Контролировать стабильность метриками флаки-рейта [4,5,3].

10. Перспективы развития

Современные тенденции указывают на интеграцию AI-инструментов для автоматического создания и восстановления тестов (self-healing tests) [3]. В ближайшем будущем Playwright получит поддержку генеративного ИИ для автоматического написания сценариев, анализа логов и предсказания сбоев.

Также ожидается усиление направлений:

- облачного тестирования;
- интеграции с контейнерами;
- анализа покрытия кода и оптимизации наборов тестов [3,2].

11. Заключение

Playwright — это мощный инструмент нового поколения, сочетающий скорость, гибкость и стабильность. Он объединяет лучшие решения Selenium, Cypress и Puppeteer, устраняя их ограничения. Использование Playwright способствует построению современных QA-процессов, ориентированных на DevOps и непрерывное качество [5,6,3,2].

Реинжиниринг тестирования с применением Playwright позволяет повысить прозрачность процессов, уменьшить ручной труд и ускорить релизы, делая его ключевым элементом цифровой трансформации QA.

Список используемых источников

1. Коледачкин А. А. Преимущества и недостатки автоматизации тестирования: анализ экономической эффективности и качества // Universum: технические науки. — 2024. — № 9(126). — URL:

https://7universum.com/ru/tech/archive/item/18271 (дата обращения: 08.10.2025). (электронный научный журнал)

2. Selenium.dev. Selenium Documentation. — URL: https://www.selenium.dev/documentation (дата обращения: 13.10.2025). (электронный ресурс)
3. Playwright.dev. Introduction to Playwright. — URL: https://playwright.dev/docs/intro (дата обращения: 13.10.2025). (электронный ресурс)
4. Tproger.ru. Современный подход к автоматизации тестирования веб-приложений с помощью Python Playwright. — 2024. — URL: https://tproger.ru/articles/sovremennyj-podhod-k-avtomatizacii-testirovaniya-veb-prilozhenij-s-pomoshhyu-python-playwright-242777 (дата обращения: 15.10.2025). (статья)
5. CodiLime. What is Playwright: All you need to know. — 2024. — URL: https://codilime.com/blog/what-is-playwright-all-you-need-to-know (дата обращения: 14.10.2025). (статья)
6. TestEngineer.ru. Сравнение скорости Cypress, Selenium, Playwright и Puppeteer. — 2023. — URL: https://testengineer.ru/speed-comparison-cypress-selenium-playwright-puppeteer (дата обращения: 17.10.2025). (статья)

Об одном подходе к решению задачи построения развертки многогранников

Верхотуров М. А.
Институт информатики, математики и робототехники
Уфимский университет науки и технологий
Уфа, Россия
e-mail: verhotur_m@rambler.ru

Верхотурова Г. Н.
Институт информатики, математики и робототехники
Уфимский университет науки и технологий
Уфа, Россия
e-mail: verhoturova.gn@yandex.ru

Иванов А. Ю.
ООО «РН-ТЕХНОЛОГИИ»
Отдел развития технологий разработки, геомеханики и ГРП
Управление по физико-математическому моделированию в цифровых системам
Уфа, Россия
e-mail: kinologprofe2003d@gmail.com

Аннотация[1]

В работе рассматривается проблема построения развертки многогранников, которая возникает при решении многих задач геометрического моделирования и проектирования. Для решения данной задачи в работе использовался подход на базе проецирования моделей на плоскость путем разрезания исходной поверхности вдоль ребер, построения множества остовных деревьев и выбора из них дерева с минимальным количеством компонент связности. В статье описаны основные шаги получения развертки, используемые алгоритмы и особенности их реализации, приведены примеры работы программного обеспечения, а также результаты вычислительного эксперимента.

Ключевые слова: развертка; многогранник; остовное дерево; пересечения; минимальное покрытие множества

1. Введение

При решении многих задач геометрического моделирования и проектирования, виртуальной и дополненной реальности, а также компьютерной графики, используются полигональные аппроксимации (например, треугольные). С другой стороны, при изготовлении изделий до сих пор очень широко применяются плоские материалы, из которых изготавливаются соответствующие поверхности. Одним из возможных вариантов построения поверхности является проектирование развёртки полигонально саппроксимированной поверхности (рисунок 1).

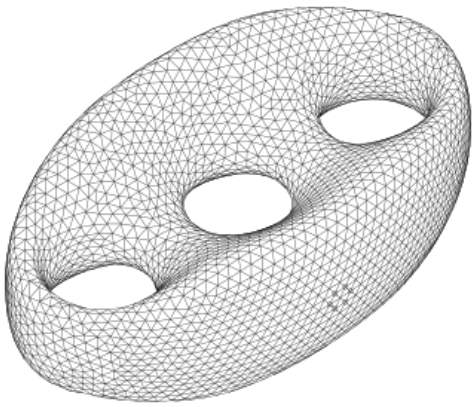


Рис. 1. Пример полигонально саппроксимированной поверхности

При интеграции этих подходов для изготовления изделий соответствующего класса получается задача, состоящая из двух связанных этапов: построение развертки (которая может иметь множество вариантов и часто состоит из нескольких частей) и плотное размещение этих частей разверток множества деталей на плоском материале. При этом следует иметь ввиду, что при получении разных вариантов разверток и их частей, получаемые раскройные карты могут иметь разные коэффициенты раскроя. Каждая из этих задач относится к задаче оптимизационного геометрического моделирования/проектирования

[1] Труды XI Международной научной конференции "Информационные технологии интеллектуальной поддержки принятия решений", 11-13 ноября, Уфа-Чандигарх-Фандербейлпарк, 2025

[1], в которой также подробно описан второй этап. В данной работе подробно рассматривается первый этап – построение развертки.

2. Развертка многогранников

Полигональная сетка – это совокупность вершин, рёбер и граней, которые определяют форму многогранного объекта в геометрическом моделировании и компьютерной графике. Гранями обычно являются треугольники, четырёхугольники или многоугольники, при этом у каждой грани есть нормаль, которая перпендикулярна поверхности грани. Развёртка многогранника – это совокупность непересекающихся многоугольников, расположенных на плоскости, соединенных ребрами и соответственно равных граням многогранника, которые могут быть сложены вдоль ребер для получения изначального многогранника.

Таким образом задачу развертки полигональной сетки можно свести к задаче нахождения развертки многогранника.

Существуют два основных вида разверток многогранников [2]:

— Реберная развертка;

— Общая развертка.

Реберная развертка – метод, при котором многогранник разрезается вдоль своих ребер для формирования цельной развертки без самопересечений.

Общая развертка – метод, который позволяет делать разрезы через внутреннюю часть граней, не ограничиваясь только ребрами. Чтобы получить развертку многогранника необходимо развернуть модель на плоскость путем разрезания исходной поверхности вдоль ребер. Данные ребра будут называться разрезаемыми ребрами, все остальные – сгибаемыми ребрами.

Исходную полигональную поверхность M можно представить в виде графа G(M) = (V, E),

где V – множество вершин поверхности M,

E – множество ребер поверхности M.

Тогда двойственным графом для исходной поверхности M является граф $D(M) = (V_d, E_d)$,

где V_d – множество полигонов поверхности M,

E_d – множество пар (f, g), где f и g грани имеющее общее ребро.

Любое остовное дерево графа D(M) определяет развертку поверхности M, так как оно содержит только сгибаемые ребра. Все ребра, не включенные в остовное дерево, определяют множество разрезаемых ребер. Таким образом, чтобы получить развертку поверхности M, требуется найти остовное дерево графа D(M).

Чтобы упростить и ускорить процесс сборки модели, необходимо избегать самопересекающиеся развертки, так как они требует дополнительного разрезания развертки на самонепересекающиеся части.

Единственный известный способ найти самонепересекающуюся развертку – перебор всех возможных вариантов остовных деревьев. Согласно [3] количество остовных деревьев для связанного графа составляет не менее 2^t, где t рассчитывается по формуле (1):

$$t = \left\lceil \frac{-1+\sqrt{1+8(E-V+1)}}{2} \right\rceil, \quad (1)$$

где E – количество ребер в графе;

V – количество вершин.

Таким образом, для модели, состоящей, например, преимущественно из четырехугольных граней,

$V_d \approx 3v$, $E_d \approx 2v$,

где v – количество вершин в исходной модели (выводится через формулу Эйлера).

Для случая, когда v = 500, количество остовных деревьев будет приблизительно $\geq 2^{32}$. Перебор такого количества вариантов требует продолжительных вычислений. Следовательно, возникает проблема уменьшения числа возможных остовных деревьев.

Эту задачу можно сформулировать следующим образом.

Пусть дано:

M - исходный многогранник,

G = (V, E)- граф, у которого

V = {f | ∀ f ∈ множество граней в M},

E = {(f, g) | ∀ f, g ∈ V, имеющие общее ребро}.

Каждая грань $f = (a, b, c)$ – треугольник, заданный вершинами $a, b, c \in R^3$.

Представлением треугольника f на плоскости является равный ему треугольник

$p(f) = (l, k, m)$, $l, k, m \in R^2$.

Разверткой многогранника M назовем такой граф U = (V', E'), где V' = {p(f) | ∀ f ∈ V}, $E' \subset E$.

При этом ∀i ≠ j:

$int(p(f_i)) \cap int(p(f_j)) = \emptyset$, где int(f) – внутренняя область треугольника f,

∀ (fi, fj) ∈ E': $p(f_i)$ и $p(f_j)$ имеют общее ребро в развертке.

Найти: преобразование φ: G → U, k(U) → min, где k(U) – количество компонент связности в графе U.

2.1. Нахождение дерева развёртки

Определим функцию веса $w : E_d \to \mathbb{R}^+$. Функция w присваивает вес каждому ребру в двойственном графе, где ребро с большим весом считается более предпочтительным для разрезания.

Raphael Straub и Hartmut Prautzsch в своей работе «Creating Optimized Cut-Out Sheets for Paper Models from Meshes» [4] предложили различные эвристики для расчета весов ребер в исходном графе модели для нахождения минимального остовного дерева. В ходе данной работы были рассмотрены две эвристики для расчета весов, позволяющие минимизировать количество пересечений в развертке, тем самым сокращая количество сегментов развертки.

Эвристика минимального периметра (1):

$$m(e) = 1 - \frac{l - l_{min}}{l_{max} - l_{min}} \in [0, 1], \qquad (2)$$

где l – длина ребра e;

l_{min}, l_{max} – минимальная и максимальная длина в графе.

Эвристика плоского дерева (2):

$$f(e) = \frac{|c^t(p-q)|}{\|p-q\|}, \qquad (3)$$

где c – некоторый вектор;

p, q – вектора координат вершин концов ребра e.

Для устранения пересечений авторы статьи не стали перебирать возможные варианты разверток, а с помощью решения задачи о покрытии множества находят минимальное множество разрезаемых ребер, необходимое для устранения пересечений.

В результате анализа для вычисления весов ребер была выбрана эвристика минимального периметра.

Так как периметр развертки равен удвоенной длине разрезаемых ребер, данная эвристика отвечает за минимизацию периметра развертки, присваивая больший вес коротким ребрам.

Тогда поиск развертки сводится к задаче нахождения минимального остовного дерева, для которой известны эффективные алгоритмы нахождения решения, например, алгоритмы Прима, Крускала, Борувки.

В данной работе минимальное остовное дерево формировалось с помощью алгоритма Крускала следующим образом.

Нахождение дерева начинается с леса из деревьев, содержащих по одной вершине исходного графа G.

Алгоритм последовательно ищет ребро для добавления в расширяющийся лес путем поиска ребра с минимальным весом среди всех ребер, соединяющих два дерева в лесу.

Если обе вершины рассматриваемого ребра принадлежат одному и тому же связному компоненту, то такое ребро отбрасывается – в противном случае образуется цикл.

Если вершины находится в разных компонентах связности, ребро принимается и связывает две компоненты.

2.2. Нахождение плоской развертки многогранника

Пусть $T = (A, B, C)$ – треугольник, заданный вершинами $A(x_A, y_A, z_A)$, $B(x_B, y_B, z_B)$, $C(x_c, y_c, z_c)$,

где x_N, y_N, z_N, $N \in \{A, B, C\}$ – координаты вершин.

Пусть $T_p = (a, b, c)$ – плоское представление треугольника, заданного вершинами $a(x_a, y_a)$, $b(x_b, y_b)$, $c(x_c, y_c)$, $n \in \{a, b, c\}$ – координаты вершин на плоскости.

Рассмотрим нахождение плоского представления для первой грани, треугольника (A, B, C). Вершину a установим в точку $(0, 0)$.

Для определения положения точки B на плоскости, рассчитаем длину ребра AB, равную длине ребра ab на плоскости по формуле (4):

$$\overline{AB} = \overline{ab} = \sqrt{(x_A - x_B)^2 + (y_A - \mathrm{y})^2 + (\mathrm{z} - z_B)^2}, \qquad (4)$$

Тогда вершину b можно установить в точку $(\overline{AB}, 0)$.

Для последней вершины c возможно два положения относительно ребра ab: $c_1 = (c_{1_x}, c_{1_y})$ или $c_2 = (c_{2_x}, c_{2_y})$, рассчитываемые по формулам (5 – 8).

$$s = \frac{\|(B-A)\times(C-A)\|}{(\overline{AB})^2} \qquad (5)$$

$$d = \frac{(B-A)\cdot(C-A)}{(\overline{AB})^2} \qquad (6)$$

$$\begin{matrix} c_{1_x} = a_x + d(b_x - a_x) - s(b_y - a_y) \\ c_{1_y} = a_y + d(b_y - a_y) + s(b_x - a_x) \end{matrix} \qquad (7)$$

$$\begin{matrix} c_{2_x} = a_x + d(b_x - a_x) + s(b_y - a_y) \\ c_{2_y} = a_y + d(b_y - a_y) - s(b_x - a_x) \end{matrix} \qquad (8)$$

Для первой грани можно выбрать любой из двух вариантов, от него будет зависеть направление построения развертки. Для всех же последующих граней необходимо проверять, чтобы для очередного разворачиваемого треугольника вершина с не была размещена с той же стороны, что и для предыдущего треугольника.

Для этого зададим $c_0 = (a_{0_x}, a_{0_y})$ – точка предыдущего треугольника, которая не принадлежит текущему треугольнику.

Введем функцию (9) для определения стороны размещения вершины c для текущего треугольника:

$$f = (c_x - a_x) \cdot (b_y - a_y) - (c_y - a_y) \cdot (b_x - a_x) \qquad (9)$$

Возможны три случая:

$f < 0$ для случая расположения точки с одной стороны;

$f > 0$ для случая расположения точки с противоположной стороны; вырожденный случай $f = 0$, когда вершины лежат на одной прямой.

В нашей задаче вырожденного случая нет.

Если $f_0 < 0$ и $f_1 > 0$ или $f_0 > 0$ и $f_1 < 0$,

тогда $c = c_1$, в противном случае $c = c_2$.

Для определения порядка формирования развертки в работе был использован алгоритм поиска в глубину и список смежности граней, полученный из найденного минимального остовного дерева.

2.3. Обнаружение и устранение пересечений

Как уже упоминалось, развертка может содержать пересечения. Чтобы устранить эти пересечения, разделим полученную развертку на несколько частей, вводя новые разрезы. Поскольку в данной работе стремимся к минимальному числу отдельных элементов, следует минимизировать количество этих разрезов. Прежде всего необходимо найти все пересечения.

Существуют различные методы для проверки пересечения двух многоугольников [5]:

a) ***Перебор ребер***. Простой и надежный способ, проверяющих пересечения пар каждого ребра из двух многоугольников. Для треугольника потребуется проверить 9 пар. Пересечение отрезков можно проверить через нахождение ориентированной площади треугольника или векторное произведение.

b) ***Separating Axis Theorem (SAT)***. Данный метод основан на теореме, говорящей, что если есть ось, на которой проекции двух выпуклых многоугольников не пересекаются, то и сами многоугольники не перекрываются. В качестве осей рассматривается нормали ко всем ребрам.

c) ***Нахождение пересечения многоугольников***. Данный метод заключается в нахождении многоугольника пересечения, если результат не пустой, многоугольник пересекаются. Примером такого алгоритма является алгоритм Сазерленда-Ходжмана. Мощный алгоритм, но избыточен для треугольников.

d) ***Вхождение вершины одного треугольника в другой***. Проверить вхождение хотя бы одной вершины одного треугольника в другой можно с помощью барицентрических координат. Если точка лежит внутри треугольника, то все ее барицентрические координаты неотрицательны.

e) ***Ограничивающие прямоугольники***. Проверка пересечения ориентированных относительно осей координат ограничивающих прямоугольников каждого многоугольника. Не дает точного результата, но позволяют сократить множество проверяемых многоугольников.

Так как в нашей задаче участвуют только треугольники, то оптимальным вариантом будем использование комбинации методов ограничивающих прямоугольников и перебора ребер. Для проверки пересечения отрезков будем использовать ориентированную площадь треугольника (10):

$$Area(A, B, C) = (x_B - x_A) \cdot (y_C - y_A) - (y_B - y_A) \cdot (x_C - x_A) \quad (10)$$

Так как каждая грань исходной модели соответствует определенной вершине в найденном минимальном остовном дереве, для каждой пары пересекающихся граней существует единственный соединяющий их путь в дереве. Чтобы устранить пересечения, необходимо разрезать каждый из этих путей хотя бы в одном месте.

Данная задача сводится к решению задачи о минимальном покрытии множества.

Пусть $T = (V'_d, E'_d)$, где $E'_d \subseteq E_d$, дерево развертки,

$P = \{P_1, \dots, P_n\}$ множество всех путей $P_i \subseteq E'_d$ между парами пересекающихся граней. Необходимо найти минимальное множество ребер S, покрывающее множество путей

$P = \{P_1, \dots, P_n\}$, такое что $S \subseteq P_1 \cup \dots \cup P_n$,

S содержит ребро e каждого из путей P, при этом должно выполняться условие (11) для минимизации стоимости разрезаемых ребер (12), которая имеет обратно значение относительно выбранной функции веса:

$$S = \underset{\substack{S' \subseteq P_1 \cup \dots \cup P_n \\ \forall i=1,\dots n:\, S' \cap P_i = \emptyset}}{\operatorname{argmin}} \sum_{e \in S'} c(e) \quad (11)$$

$$c(e) = 1 - m(e) \quad (12)$$

Условие выбора ребра - покрывающее наибольшее количество путей за минимальную стоимость.

Для решения этой задачи в работе реализован жадный алгоритм.

2.4. Результаты экспериментальных исследований и их анализ

Программный продукт был разработан в среде разработки Qt Creator, версия Qt 5.15, на языке программирования C++. В качестве графической библиотеки была выбрана OpenGL.

Тестирование проводилось на трех видах моделей, для каждой модели были подготовлены по пять уровней детализации от 100 до 1000 полигонов. На рисунке 1 представлены примеры используемых моделей кролика с уровнями детализации 100, 400 и 1000 полигонов.

Для каждого уровня детализации было проведено по 10 экспериментов и взято среднее значение их

результатов. Каждый пример представляет из себя полигональную модель, состоящую из треугольников.

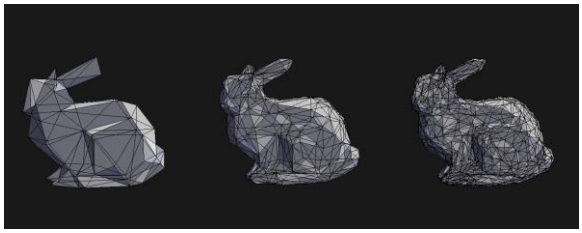


Рис. 2. Уровни детализации модели кролика 100, 400 и 1000 полигонов

В качестве примера с большим количеством полигонов была выбрана модель стэндфордского кролика (1014 полигонов). Построение развертки кролика заняло 2,094 секунды. Результат представлен на рисунке 2.

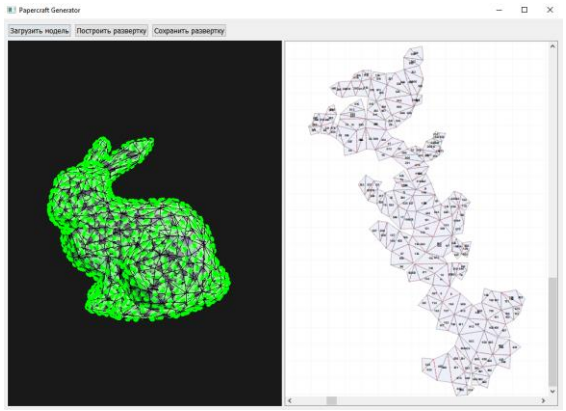


Рис. 3. Модель и развертка стэндфордского кролика

Результаты проведенных экспериментов показали, что время построения развертки зависит от количества полигонов, при этом зависимость близка к линейной, что является хорошим результатом для решаемой задачи.

3. Заключение

Таким образом, в работе разработан и реализован подход для формирования развертки моделей на плоскость.

В качестве метода представления модели выбрано представление в виде списка граней. Разработаны и реализованы алгоритмы нахождения минимального остовного дерева с эвристикой минимального периметра для определения весов ребер, поиска пар пересекающихся граней, устранения пересечений, разбиения развертки на части.

Результаты показали, что время построения развертки линейно зависит от количества полигонов, что является вполне удовлетворительным для решаемой задачи.

Список используемых источников

1. Верхотуров М.А. Задача нерегулярного раскроя фигурных заготовок: оптимизация размещения и пути режущего инструмента // Вестник УГАТУ—2007. — т. 9, №. 2 (20), с. 106-118.
2. Demaine E. D. Geometric Folding Algorithms. Linkages, Origami, Polyhedra / E. D. Demaine, J. O'Rourke // Cambridge: Cambridge University Press, 2007. – 486 p. – ISBN: 978-0-5218-5757-4.
3. Read, R. C. Bounds on backtrack algorithms for listing cycles, paths, and spanning trees / R. C. Read, R.E. Tarjan // Networks – 1975. – vol. 5. – pp. 237-252.
4. Straub, R. Creating optimized cut-out sheets for paper models from meshes / R. Straub, H. Prautzsch// Karlsruhe Reports in Informatics. – 2011. – vol. 22. – pp. 1-15.
5. Mark de Berg. Chapter 2: Line Segment Intersection / Mark de Berg, Marc van Kreveld, Mark Overmars, Otfried Schwarzkopf. // Computational Geometry. – 2nd. – Springer. – 2000. – pp. 19-44. – ISBN 3-540-65620-0.

Классификация сердечно-сосудистых заболеваний на основе энтропийного анализа ЭКГ с использованием GPU

Вафин Р. Р.
ФГБОУ ВО УУНиТ, ИИМиРТ
Уфа, Россия
e-mail: vafrus74@yandex.ru

Насыров Р. В.
ФГБОУ ВО УУНиТ, ИИМиРТ
Уфа, Россия
e-mail: nrash@yandex.ru

Аннотация[1]

Анализ временных последовательностей процессов для решения задач распознавания, классификации и принятия решений является актуально задачей. Это связано с вариативностью ситуаций получения таких данных, сложным поведением объекта исследования, наличием структурной сложности полученных данных. В связи с этим рассматривается проблема использования высокотехнологичных методов анализа данных для обработки данных временных последовательностей со сложной вариативной внутренней структурой с целью улучшения уровня их распознавания и классификации. В качестве примера рассматриваются результаты измерений электрокардиограмм в различных форматах, которые отражают особенности состояния сердечно-сосудистой системы. В качестве метода анализа и выявления особенностей предлагается использовать показатели нелинейной динамики временных последовательностей, в частности показатели энтропий как мер хаотичности временных последовательностей. Исследуются возможности вычисления таких оценок на центральном процессоре и на графическом процессоре, на основании чего оценивается применимость вариантов способов реализации предлагаемых алгоритмов. Проводится сравнительное исследование затрат времени на проведение вычислений показателей энтропии для последовательностей разных размеров, на основании которых можно оценить применимость методов на практике. Результаты проведенных исследований показывают применимость указанных методов для решения задач распознавания и классификации временных последовательностей.

Ключевые слова: временная последовательность, энтропия, GPU, OpenCL, классификация, сердечно-сосудистые заболевания.

1. Введение

Сердечно-сосудистые заболевания (ССЗ) являются одной из основных причин смертности во всем мире. Ранняя и точная диагностика ССЗ по ЭКГ остается актуальной задачей. Традиционные методы анализа часто оказываются недостаточно чувствительными к слабовыраженным и скрытым патологиям. Энтропийные метрики, позволяющие количественно оценить сложность и нерегулярность сердечного ритма, показали свою эффективность в выявлении таких аномалий [14]. Однако их широкому применению в клинической практике мешает высокая вычислительная сложность, особенно при обработке длительных записей, таких как данные холтеровского мониторинга. Кроме этого, с точки зрения практических задач необходимость проведения исследования обусловлена определением работоспособности методов на ограниченных вычислительных ресурсах с оценкой возможности использования в задачах диагностики заболеваний и организации и планирования процесса их лечения.

Целью работы является исследование методики сравнительной оценка эффективности ресурсоемких вычислений различных показателей средствами CPU и GPU на примере данных ЭКГ.

В связи с этим задачи данный работы:

1. Разработка и реализация высокопроизводительного алгоритма вычисления набора энтропийных метрик (ApEn, SampEn, MFE, PE, Спектральная энтропия) для анализа ЭКГ на графических процессорах с использованием OpenCL.
2. Оценка эффективности предложенной реализации путем сравнения времени вычислений с одно- и многопоточными версиями на ЦП.
3. Проверка диагностической ценности вычисленных признаков путем построения и оценки модели классификации ССЗ на основе машины опорных векторов.

[1] Труды XI Международной научной конференции "Информационные технологии интеллектуальной поддержки принятия решений", 11-13 ноября, Уфа-Чандигарх-Фандербейлпарк, 2025

4. Определение основных сценариев использования рассматриваемых в статье алгоритмов.

2. Показатели энтропии для анализа ЭКГ как временных рядов

Энтропийный анализ временных рядов, к которым относится сигнал ЭКГ, является мощным инструментом для выявления скрытых паттернов и аномалий, который позволяет количественно оценить сложность и нерегулярность сердечного ритма, что является важным диагностическим признаком для некоторых ССЗ [1].

В исследованиях данных временных последовательностей используются следующие типы энтропий:

- Аппроксимированная энтропия (ApEn): Мера регулярности временных рядов. Высокое значение ApEn указывает на меньшую регулярность и большую сложность сигнала, что может коррелировать с патологиями [2].
- Энтропия выборки (SampEn): Усовершенствованная версия ApEn, лишенная смещения, связанного с само сравнением, что делает ее более точной и стабильной, особенно на коротких записях ЭКГ [3].
- Мультифрактальная энтропия (MFE): Расширяет анализ, учитывая мультифрактальную природу сигнала, что позволяет детально изучать нерегулярные структуры на различных масштабах [4].
- Пермутационная энтропия (PE): Оценивает сложность через анализ паттернов перестановок в последовательности данных. Устойчива к шуму и эффективна для захвата динамики сигнала ЭКГ [5].
- Спектральная энтропия: Основана на преобразовании Фурье и измеряет хаотичность сигнала в частотной области, что полезно для выявления аритмий [6].

Наиболее простым для реализации и последующей интерпретации являются первые два показателя.

2.1. Алгоритмы вычисления энтропий

Известно несколько реализаций алгоритмов вычисления энтропий, которые отличаются скоростью и качеством результатов, например, такие как [8-11]. При этом различия обусловлены используемым языковым средством, качеством реализации алгоритмов, а так же реализацией на базе GPU. Последний фактор в настоящее время является наиболее значимым в решении задачи ускорения вычислений. Среди рассмотренных наибольший интерес представляет библиотека для вычисления энтропий, предложенная коллективом авторов [7]

2.2. Выбор набора исходных данных для примера классификации заболеваний

Проверка эффективности проводится на тестовой выборке из реальных данных пациентов с известными диагнозами. Для проверки работоспособности модуля анализа ЭКГ используется датасет Icentia11k содержащий ЭКГ 11000 пациентов и датасет MIT-BIH arrhythmia database [12, 13]. Для каждого ЭКГ производится расчет различных оценок энтропии для исходного сигнала, производной сигнала 1 и 2 степени. Сами ЭКГ разрезаются на короткие отрезки, содержащие 4 R-R интервала. Длина данных отрезков динамическая и зависит от частоты сердечного ритма пациента в каждый контекстный промежуток времени.

2.3. Проблемы эффективности вычислений энтропий на ЦП

Для задачи анализа денных непосредственно на компьютере врача, достаточно вычисление на обычном процессоре. Объем данных для ЭКГ в 12 отведениях с частотой 250Гц будет 3Кбайт\с. Такой объем данных можно обрабатывать за примемлимое количество времени, в зависимости от модели процессора, обработка одной секунды данных может занимать в среднем от 1 до 10 секунд.

Для обучения и дообучения моделей классификации, потребуется обрабатывать огромное количество данных, что даже на кластере из компьютеров, имеющих только ЦП будет занимать большое количество времени. Например, датасет Icentia11k (около 1.1Тб) на одном компьютере с ЦП ryzen 5 4600h будет обрабатываться около 450 дней.

Для задачи анализа данных холтер мониторинга объем данных превосходит на 3 порядка, чем анализ коротких интервалов обычного ЭКГ. Для этих целей необходимо применение более мощных вычислительных устройств.

3. Предлагаемый подход к реализация алгоритмов анализа на графическом процессоре

Вследствие ограничений используемого оборудования, а именно использования высокопроизводительных видеокарт AMD, было решено использовать OpenCL для выполнения вычислений на видеокарте. Использование OpenCL позволит использовать данный код с любой видеокартой, поддерживающей OpenCL версии как минимум 1.2.

Управляющая программа написана на python и использование библиотеки puopencl для запуска вычислений на видеокарте. Основная вычислительная нагрузка (попарное сравнение

векторов для ApEn/SampEn) была вынесена в кернелы OpenCL. В алгоритме было распараллелено вычисление сумм векторов которые меньше указанной метрике в каждом из алгоритмов вычисления энтропии. Это позволило полностью загрузить вычислительные блоки видеокарты.

4. Результаты вычислительных экспериментов

4.1. Сравнительная таблица различных режимов

Для проверки работы были выбраны из числа имеющихся компьютеры с графическими ускорителями GTX 1650 ti, RX 6950 XT. А также, для проверки работоспособности алгоритмов на центральных процессорах, процессоры Ryzen 5 4600H и Ryzen 5 7500f. При этом сравнивались вычисления, проводимые в режимах однопоточного вычисления на CPU, многопоточного вычисления на CPU, на разных вариантах бюджетных графических процессоров.

Графики времени вычисления в зависимости от длины интервала для каждого устройства

Графики для многопотока на ГПУ и ЦПУ для алгоритмов ApEn и SampEn для 52 ЭКГ. Считалось общее время обработки датасета.

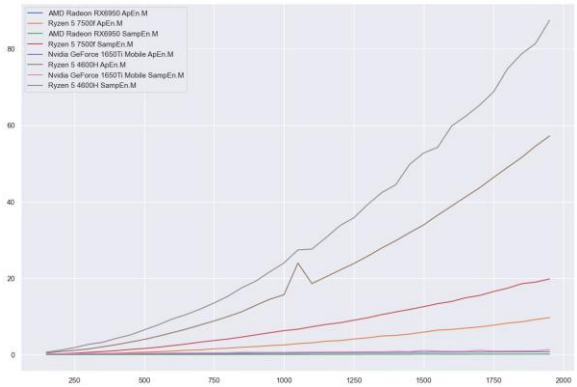


Рис. 1. Линейный масштаб времени

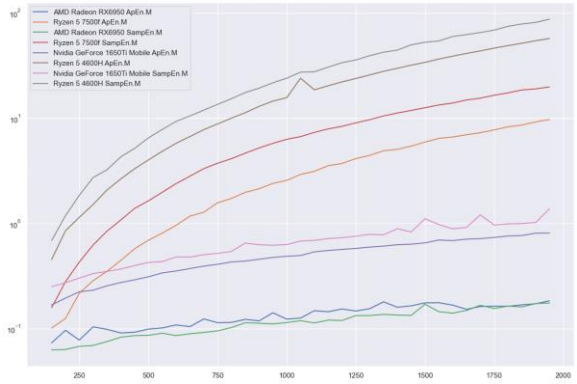


Рис. 2. в логарифмической шкале



Рис. 3. Характерные времена расчетов нескольких показателей энтропий приведены в таблице 1.

Таблица 1. Сравнение времени вычисления (в секундах) для набора энтропий

Устройство	4 сек 100гц 12 каналов	4 сек 250гц 12 каналов	4 сек 500гц 12 каналов	Холтер 24 часа 100гц1 канал (отрезки по 10 сек)
GTX 1650 Ti	0.14c	0.22c	0.31c	58,8м
RX 6950 XT	0.05c	0.07c	0.09c	24м
Ryzen 5 4600H	0.22c	1.14c	4.01c	31,2ч
Ryzen 5 7500f	0.10c	0.21c	0.69c	5,1ч

Полученные результаты позволяют рассмотреть разные способы практического использования разработанных программ в составе программных комплексов СППР. Например, они позволяют сформировать минимальный набор сценариев применения разработанных программ в рамках предлагаемого подхода для разных по сложности задач. Предлагаемый набор сценариев представленные в таблице 2.

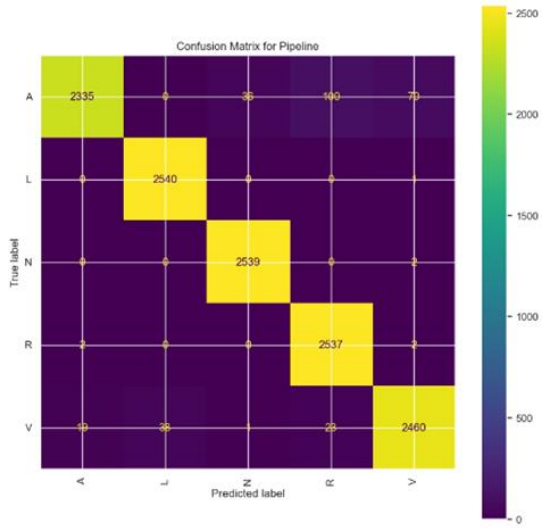
Таблица 2. Типовые параметры времени вычисления для различных задач

Задача	Доступное оборудование	Время
Работа врача	Простой компьютер (CPU до 4 ядер, без GPU)	0-5 секунд для одного пациента
Исследование внутри клиники (пациенты клиники)	Рабочая станция (многоядерный CPU, средняя GPU)	Для 1000 пациентов: Несколько часов
Разработка и\или доработка модели предсказания	Высокопроизводительный сервер (мощная GPU)	Для 10000 пациентов: Несколько часов

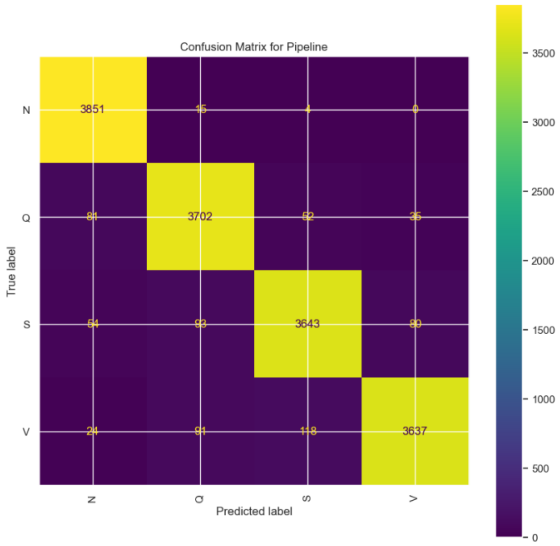
Отметим, что такие сценарии являются доступными даже на той технической базе, которая есть в большинстве медицинских организаций.

5. Результаты классификации ССЗ на основе показателей энтропии

С практической точки зрения типовым объектом исследования с использованием рассмотренных методов и алгоритмов являются временные последовательности в форме ЭКГ. При этом рассчитанные показатели энтропии для исходного сигнала и его производных используются для обучения моделей классификации сердечно-сосудистых заболеваний (ССЗ). В данном исследовании предлагается использовать машину опорных векторов (описание или ссылку). Исходные датасеты были сбалансированы по классам заболеваний.



а)



б)

Рис. 4 . Результаты классификации ССЗ на тестовых выборках а) Icentia11k и б) MIT-BIH arrhythmia database

Результаты классификации для данных, полученных из разных наборов, представлены на рисунке 1.

Показатели: precision: 0.95 f1-score: 0.96

6. Заключение

Таким образом, в статье проведено исследование методики сравнительной оценка эффективности ресурсоемких вычислений различных показателей средствами CPU и GPU на примере данных ЭКГ и получены следующие результаты.

- Разработана и реализована на базе OpenCL высокопроизводительная методика вычисления набора энтропийных метрик (ApEn, SampEn, MFE, PE, Спектральная энтропия) для анализа ЭКГ на графических процессорах.
- Проведена сравнительная оценка эффективности реализации. Показано, что использование GPU позволяет ускорить вычисления в 10-100 (для различных комбинаций CPU и GPE) раз по сравнению с многопоточным CPU для длительных записей, таких как данные холтеровского мониторинга.
- На основе вычисленных признаков построена модель классификации ССЗ. Полученные метрики (precision: 0.95 f1-score: 0.96) на тестовых выборках подтверждают диагностическую ценность энтропийных метрик для выявления патологий.
- Определены сценарии применения методики в зависимости от вычислительных ресурсов: для оперативной работы врача достаточно CPU, для научных исследований и обработки больших данных необходимо использование GPU.

Список используемых источников

1. Schuckers S. A. C. Use of approximate entropy measurements to classify ventricular tachycardia and fibrillation //Journal of electrocardiology. – 1998. – Т. 31. – С. 101-105.
2. Pincus S. M. Approximate entropy as a measure of system complexity //Proceedings of the national academy of sciences. – 1991. – Т. 88. – №. 6. – С. 2297-2301.
3. Richman J. S., Lake D. E., Moorman J. R. Sample entropy //Methods in enzymology. – Academic Press, 2004. – Т. 384. – С. 172-184.
4. Gospodinova E., Lebamovski P., Gospodinov M. Automatic analysis of ECG signals based on their fractal and multifractal properties //Proceedings of

the 22nd International Conference on Computer Systems and Technologies. – 2021. – С. 136-140.

5. Bandt C., Pompe B. Permutation entropy: a natural complexity measure for time series //Physical review letters. – 2002. – Т. 88. – №. 17. – С. 174102.
6. Zhang A., Yang B., Huang L. Feature extraction of EEG signals using power spectral entropy //2008 international conference on BioMedical engineering and informatics. – IEEE, 2008. – Т. 2. – С. 435-439.
7. Flood M. W., Grimm B. EntropyHub: An open-source toolkit for entropic time series analysis //PloS one. – 2021. – Т. 16. – №. 11. – С. e0259448.
8. Программный модуль вычисления SamPen. URL: https://github.com/kirajcg/pysampen (дата обращения: 01.10.2025)
9. Программный модуль вычисления SamPen на ГПУ. URL: https://pypi.org/project/sampen-gpu/ (дата обращения: 01.10.2025)
10. Программный модуль вычисления SamPen в среде Matlab. URL: https://nl.mathworks.com/ matlabcentral/fileexchan ge/35784-sample-entropy (дата обращения: 01.10.2025)
11. JiřTomčala, New Fast ApEn and SampEn Entropy Algorithms Implementation and Their Application to Supercomputer Power Consumption , http:// dx.doi.org/10.3390/e22080863
12. Tan S. et al. Icentia11k: An unsupervised representation learning dataset for arrhythmia subtype discovery //arXiv preprint arXiv:1910.09570. – 2019.
13. Moody G. B., Mark R. G. The MIT-BIH arrhythmia database on CD-ROM and software for use with it //[1990] Proceedings Computers in Cardiology. – IEEE, 1990. – С. 185-188.
14. Зулкарнеев Р.Х. Кардиореспираторная вариабельность при заболеваниях органов дыхания// Автореферат дисс. Д-ра мед. Наук, 2007 г.

Компьютерный анализ эмоционально окрашенной информации в задачах управления проектами: обзор современных методов и моделей

Богданова Д. Р.
Уфимский университет науки и технологий
Уфа, Россия
e-mail: dianochka7bog@mail.ru

Мавлиханов Ш. М.
Уфимский университет науки и технологий
Уфа, Россия
e-mail: mavlix95@gmail.com

Насыров И. Ю.
Уфимский университет науки и технологий
Уфа, Россия
e-mail: ilasn909@gmail.com

Аннотация[1]

В статье рассматриваются современные методы компьютерного анализа эмоционально окрашенной информации в задачах управления проектами. Проведён обзор существующих подходов к анализу текстовой, аудио и видеоинформации, а также методов мультимодального объединения признаков. Особое внимание уделяется применению трансформерных архитектур и методов глубокого обучения для оценки эмоционального состояния участников проектной команды.

В работе предложена формальная постановка задачи мультимодального анализа эмоций в проектных коммуникациях, а также разработана концептуальная архитектура системы, включающая этапы извлечения признаков, мультимодального слияния и оценки эмоциональной динамики. Предложенный подход позволяет выявлять отклонения эмоционального фона команды и использовать полученные данные для поддержки управленческих решений.

Ключевые слова: эмоциональный интеллект, искусственный интеллект, управление проектами, сентимент-анализ, мультимодальный анализ, распознавание эмоций.

1. Введение

Эффективность управления проектами во многом определяется качеством коммуникации между участниками команды. Эмоциональная окраска сообщений, интонация речи и невербальные проявления оказывают существенное влияние на принятие решений, уровень вовлечённости сотрудников и вероятность возникновения конфликтов.

Современные цифровые среды управления проектами (корпоративные мессенджеры, системы видеоконференций, трекеры задач) генерируют значительные объёмы данных, содержащих эмоциональную составляющую. Это создаёт предпосылки для применения методов компьютерного анализа с целью мониторинга эмоционального климата и раннего выявления рисков.

Несмотря на наличие большого количества исследований в области анализа эмоций, вопрос их комплексного применения в контексте управления проектами остаётся недостаточно формализованным. В связи с этим в работе не только проводится обзор современных методов и моделей, но и предлагается формальная постановка задачи мультимодального анализа эмоционально окрашенной информации, а также концептуальная архитектура её решения.

2. Обзорная часть

Фазульянов Д.В., Гусева А.И. в статье «Разработка мультимодального метода сентимент-анализа для поддержки принятия решений в организациях» [1] предлагают комплексный подход к интеграции текстовых, аудио и видеоданных для повышения точности определения эмоциональной окраски коммуникаций. Авторы обосновывают необходимость перехода от одномодальных методов к мультимодальным, поскольку текст не

[1] Труды XI Международной научной конференции "Информационные технологии интеллектуальной поддержки принятия решений", 11-13 ноября, Уфа-Чандигарх-Фандербейлпарк, 2025

всегда передаёт нюансы, выражаемые интонацией, мимикой или жестами. В работе представлены архитектуры нейронных сетей для каждой модальности: ResNet-50 для видеоданных, LSTM для аудиосигналов и BERT для текстовой обработки. Эксперименты на датасете eNTERFACE показали, что мультимодальная модель на базе SVM достигает точности 91%, превосходя одномодальные подходы. Особое внимание уделено практическому применению метода в организационных системах: раннее выявление стресса в командах, мониторинг удовлетворённости клиентов и поддержка стратегического планирования. Авторы подчёркивают, что интеграция мультимодального сентимент-анализа в процессы принятия решений повышает обоснованность управленческих действий и способствует адаптивности организации в условиях цифровой трансформации.

В статье «Обзор подходов к анализу эмоциональной окраски текстов СМИ» [2] проводится систематизация современных методов обработки естественного языка для выявления тональности в медиаконтенте. Рассматриваются как традиционные подходы (лексические словари, машинное обучение на признаках n-gram), так и современные трансформерные модели (BERT, RoBERTa, GPT). Особый акцент сделан на специфике русскоязычных текстов: проблемах морфологической вариативности, иронии, контекстуальной зависимости эмоциональных маркеров. Авторы демонстрируют, что комбинация лингвистических правил и нейросетевых классификаторов позволяет достичь точности до 89% на новостных корпусах. В контексте управления проектами отмечается потенциал применения таких методов для мониторинга репутации проекта в СМИ, анализа обратной связи от стейкхолдеров и выявления ранних сигналов о рисках, связанных с общественным восприятием проекта.

Jiwat Ram в материале «Sentiment analysis for project stakeholder management» [3] на платформе IPMA World аргументирует целесообразность использования NLP-технологий для работы со стейкхолдерами проектов. Автор выделяет три ключевых направления применения сентимент-анализа:

1) Проактивное выявление и адресация обеспокоенностей стейкхолдеров через анализ электронной переписки и протоколов встреч.

2) Мониторинг социальных медиа для оперативного реагирования на публичную обратную связь.

3) Управление кризисными ситуациями на основе анализа эмоциональной динамики в коммуникациях.

Подчёркивается, что сентимент-анализ не заменяет, а дополняет компетенции руководителя проекта, предоставляя инструмент для более тонкой настройки взаимодействия с людми. В заключении отмечается необходимость дальнейших исследований по интеграции таких решений в стандарты проектного менеджмента и адаптации алгоритмов под культурные особенности команд.

Мажит А.Н., Кожамкулова Ж.Ж. в работе «Детектирование лиц и распознавание эмоций с использованием алгоритмов компьютерного зрения» [4] представляют программный комплекс для анализа визуальных данных в реальном времени. Система интегрирует каскадные классификаторы Хаара для детекции лиц и свёрточные нейронные сети для классификации базовых эмоций по модели Экмана. Экспериментальная валидация показала высокую точность распознавания в контролируемых условиях. Авторы обсуждают перспективы применения технологии в проектном менеджменте: анализ вовлечённости участников видеоконференций, оценка эмоционального климата на стендап-встречах, автоматизированный мониторинг невербальных сигналов в распределённых командах. Вместе с тем отмечается необходимость учёта этических аспектов и культурной специфики выражения эмоций при внедрении подобных систем.

В статье, опубликованной в International Research Journal [5], описывают разработку веб-сервиса для распознавания эмоций на основе TensorFlow.js и OpenCV. Система поддерживает три режима работы: анализ статических изображений, обработка загруженных видео и распознавание в реальном времени через веб-камеру. Архитектура построена на паттерне MVT (Django), что обеспечивает модульность и масштабируемость. Тестирование на наборе FER-2013 показало точность 69,4%. Авторы подчёркивают практическую ценность решения для задач маркетинга, образования и психологического консультирования, а также указывают на направления развития: расширение спектра распознаваемых эмоций, интеграция аудиомодальности, адаптация под мобильные платформы. В контексте управления проектами такой инструмент может использоваться для ретроспективного анализа коммуникаций команды и формирования рекомендаций по улучшению взаимодействия.

Savchenko A.V. в статье [6] представляет результаты команды HSEmotion на конкурсе ABAW-8. Предложен мультимодальный пайплайн, объединяющий признаки из трёх каналов: визуальные дескрипторы (EmotiEffLib), акустические эмбеддинги (wav2vec 2.0, HuBERT) и текстовые представления (RoBERTa, Whisper). Ключевая инновация — стратегия фильтрации надёжных кадров для повышения качества классификации и использование позднего слияния предсказаний. Эксперименты подтвердили

значительное улучшение метрик по трём задачам: распознавание выражений лица, оценка интенсивности эмоциональной мимикрии и детекция амбивалентности. Хотя работа ориентирована на фундаментальные задачи аффективных вычислений, авторы отмечают потенциал применения в системах адаптивного обучения и автоматизированной поддержки принятия решений, что релевантно для цифровизации проектных процессов.

Исаева Е.В. и соавт. в статье «Семиотический анализ текстов и интерпретация знаковых систем в цифровую эпоху» [7] предлагают оригинальный подход к сентимент-анализу, сочетающий семиотическую теорию с методами машинного обучения на платформе KNIME. Авторы интерпретируют тексты отзывов как знаковые системы, выделяя лексические, синтаксические и прагматические маркеры эмоциональной окраски. Использование алгоритмов Decision Tree и XGBoost на корпусе IMDB обеспечило точность классификации до 92,0%. Новизна работы заключается в возможности содержательной интерпретации результатов ML-моделей через призму семиотики, что повышает доверие к автоматизированным решениям. В управлении проектами такой подход может применяться для глубокого анализа обратной связи от заказчиков, выявления скрытых ожиданий и формализации качественных данных для принятия стратегических решений.

Фадел В. в магистерской диссертации «Нейросетевые алгоритмы распознавания эмоций на основе сигналов ЭЭГ» [8] исследует возможности психофизиологического мониторинга эмоционального состояния. В работе использованы вейвлет-преобразования для извлечения признаков из сигналов электроэнцефалографии и архитектуры GRU для классификации эмоций. Хотя применение ЭЭГ в проектном менеджменте ограничено этическими и техническими барьерами, автор указывает на перспективность направления для исследований в области человеко-ориентированного проектирования и оценки когнитивной нагрузки в высокоинтенсивных проектах. Результаты демонстрируют принципиальную возможность объективной регистрации эмоциональных реакций, что может дополнять субъективные методы оценки в будущем.

Nelsen A. Rahul в статье «Strategic Alignment of Emotional Intelligence with Project Management» [9] развивает тезис о стратегической интеграции эмоционального интеллекта в процессы управления проектами. Автор предлагает фреймворк, связывающий компетенции эмоционального интеллекта (далее — ЭИ) — самосознание, саморегуляцию, эмпатию и социальные навыки — с этапами жизненного цикла проекта. На примерах NASA и AT&T показано, что команды с высоким уровнем ЭИ демонстрируют лучшую коммуникацию, лидерство и продуктивность. Особое внимание уделено практическим шагам развития ЭИ через обучение, коучинг и наставничество. Работа подчёркивает, что технологии ИИ должны не заменять, а усиливать человеческие компетенции, создавая синергию для устойчивого успеха проектов.

Лукашевич Н.В. в статье «Автоматический анализ тональности текстов: проблемы и методы» [10] представляет систематизированный обзор подходов к решению задачи сентимент-анализа с акцентом на обработку русскоязычных текстов. Автор выделяет три исторических этапа развития методов: классические алгоритмы машинного обучения (Naive Bayes, SVM, логистическая регрессия), нейросетевые архитектуры (LSTM, CNN) и современные трансформерные модели (BERT, RuBERT). На примере датасета SentiRuEval-2015 демонстрируется прогресс качества классификации: от F1=48.80 для SVM на мешке слов до результатов, близких к человеческой разметке, при использовании RuBERT с аугментацией данных. Особое внимание уделено проблемам, специфичным для русского языка: морфологической вариативности, контекстуальной зависимости тональности и дефициту размеченных корпусов. В контексте управления проектами автор подчёркивает практическую ценность методов для мониторинга обратной связи в корпоративных чатах, анализа отчётов о статусе задач и раннего выявления рисков, связанных с эмоциональным состоянием команды. Статья также рассматривает стратегии аугментации и псевдоразметки данных как способ преодоления ограниченности обучающих выборок, что особенно актуально для российских реалий проектного менеджмента.

Vishnu V.M., Kumar P., Prajwal T.H. и соавт. в работе «Facial Emotion Recognition System» [11] описывают архитектуру системы, объединяющей биометрическую идентификацию личности и распознавание эмоционального состояния по выражению лица. Предложенный пайплайн включает этапы детекции лиц (MTCNN, YOLO), извлечения признаков (CNN-эмбеддинги VGGFace, FaceNet) и классификации эмоций с учётом временной динамики (LSTM для видео). Авторы опираются на модель FACS Пола Экмана для интерпретации микровыражений и демонстрируют работу системы на датасетах FER2013, CK+, AffectNet. Эксперименты подтверждают устойчивость алгоритма к изменениям освещения, ракурса и наличию окклюзий. В прикладном аспекте обсуждаются сценарии использования в проектном менеджменте: оценка вовлечённости участников видеоконференций, мониторинг эмоционального климата на стендап-встречах, адаптивное управление нагрузкой на основе невербальных сигналов. Вместе с тем авторы указывают на этические вызовы: необходимость

информированного согласия, риски культурной предвзятости моделей и требования к защите персональных данных. Статья завершается обзором перспективных направлений: мультимодальная интеграция (лицо + голос + физиологические сигналы), edge-вычисления для работы в реальном времени и применение трансформеров для контекстуального анализа эмоций.

Павлюченко К.А. в статье «Анализ методов выявления эмоционального контекста в текстах» [12] проводит систематическое исследование проблемы автоматизированного определения эмоциональной окраски текстовых документов на естественном языке. Автор выделяет два основных направления в решении данной задачи: математическое (алгоритмическое) и лингвистическое. Математические методы, включая векторный анализ, наивный байесовский классификатор и рекуррентные нейронные сети, обеспечивают высокую скорость обработки больших массивов данных, но демонстрируют ограниченную устойчивость к контекстуальным нюансам и иронии. Лингвистический подход, напротив, позволяет учитывать стилистические, синтаксические и морфологические особенности текста, однако требует значительных временных затрат и экспертной квалификации. В работе предложена шестиступенчатая схема эмоционального анализа: от первичного ознакомления с текстом до обобщения результатов с учётом авторской позиции и жанровой специфики. Особое внимание уделено перспективам интеграции лингвистических и математических методов: автоматизация лексико-синтаксического анализа с последующей интерпретацией результатов в численном формате открывает возможности для расчёта корреляций, дисперсии и других статистических метрик эмоциональной динамики. В контексте управления проектами автор подчёркивает практическую ценность комбинированного подхода для мониторинга коммуникаций в корпоративных чатах, анализа обратной связи от стейкхолдеров и раннего выявления рисков, связанных с эмоциональным состоянием команды.

Михненко П.А. в статье «Анализ мультимодальных данных в управлении проектами: перспективы использования машинного обучения» [13] исследует возможности применения технологий искусственного интеллекта для обработки динамических массивов разнородных данных в процессах управления изменениями проектов. Автор обосновывает гипотезу о том, что современные аналитические процедуры в проектном менеджменте требуют решения задачи слияния и интерпретации мультимодальной информации (текст, аудио, видео, метаданные). Методологическую основу работы составили методы дескриптивного моделирования процессов УИП и экспертных оценок: в исследовании приняли участие 42 специалиста, оценивавших перспективность внедрения ИИ-технологий до 2030 года. Результаты экспертизы показали, что наиболее востребованными направлениями являются обработка естественного языка (NLP) и интеллектуальная поддержка принятия решений. В первом случае речь идёт об автоматическом извлечении сущностей, фактов и онтологий из текстовой документации проекта, во втором — о сверхкраткосрочном прогнозировании, имитационном моделировании поведения стейкхолдеров и анализе потоков данных в реальном времени. Автор подчёркивает, что технологии ИИ не должны заменять традиционные подходы к управлению проектами, а служить дополнительным инструментом, расширяющим аналитические возможности команды. Вместе с тем отмечается необходимость учёта этических и организационных рисков: снижение ответственности исполнителей, формирование зависимости от алгоритмических рекомендаций и потенциальное вытеснение человеческого труда. Статья завершается выводом о целесообразности поэтапной интеграции ИИ-решений в проектную практику с обязательным сохранением человеческого фактора в принятии стратегических решений.

Cheng B., Arora C., Liu X., Hoang T., Wang Y., Grundy J. в статье «Multi-Modal Emotion Recognition for Enhanced Requirements Engineering: A Novel Approach» [14] представляют платформу MEmoRE для мультимодального распознавания эмоций в процессе инженерии требований. Авторы обосновывают актуальность учёта эмоциональных сигналов стейкхолдеров на этапах сбора, валидации и уточнения требований, поскольку игнорирование эмоционального контекста может приводить к неверной интерпретации потребностей и снижению качества программного продукта. Платформа интегрирует три модальности: анализ мимики (на основе 3D-CNN), интонации голоса (2D-CNN с временным вниманием) и текстового сентимента (трансформерные модели), объединяя их через архитектуру VAANet (Visual Audio Attention Network). Экспериментальная оценка на датасете MELD показала, что оптимальная точность достигается при сегментации видеопотока на 10-секундные интервалы, при этом положительные эмоции (радость, удивление) распознаются точнее отрицательных (грусть, гнев). В контексте управления проектами авторы подчёркивают практическую ценность MEmoRE для:

1) Приоритизации требований на основе эмоциональной реакции стейкхолдеров.

2) Адаптации стратегий элиситации при выявлении фрустрации или замешательства.

3) Автоматизированного мониторинга вовлечённости в usability-тестировании.

Отмечаются ограничения: индивидуальная вариативность мимики, культурные различия в

выражении эмоций и необходимость информированного согласия участников. Перспективные направления включают интеграцию с VR-средами и поддержку специальных групп пользователей.

Nandwani P., Verma R. в обзорной статье «A review on sentiment analysis and emotion detection from text» [15] систематизируют современные методы анализа тональности и детекции эмоций в текстовых данных. Авторы чётко разграничивают понятия: sentiment analysis определяет полярность (позитив/негатив/нейтраль), тогда как emotion detection идентифицирует конкретные эмоциональные состояния (радость, гнев, страх и др.) на основе психологических моделей (Экмана, Плутчика, Рассела). В работе детально рассмотрены три уровня анализа: документный, предложный и аспект-ориентированный, а также три класса методов: лексикон-базированные, машинное обучение (SVM, Naive Bayes) и глубокое обучение (LSTM, BERT, CNN). Особое внимание уделено этапам предобработки текста (токенизация, лемматизация, удаление шума) и техникам векторного представления слов (Word2Vec, GloVe, BERT-эмбеддинги). Авторы подчёркивают, что для задач управления проектами наиболее релевантны аспект-ориентированные подходы, позволяющие выявлять эмоциональную окраску конкретных элементов: сроков, ресурсов, коммуникаций. Ключевые вызовы, отмеченные в обзоре: обработка иронии и сарказма, контекстуальная зависимость тональности, дефицит размеченных корпусов для профессиональных доменов. Статья завершается рекомендациями по выбору методов в зависимости от типа данных и предметной области, что делает её ценным ориентиром для внедрения эмоционального анализа в проектные процессы.

Kratzwald B., Feuerriegel S., Prendinger H. и соавт. в систематическом картографировании «Sentiment Analysis in Software Engineering: A Systematic Mapping Study» [16] анализируют 106 исследований по применению анализа настроений в программной инженерии с целью помощи практикам в выборе инструментов и выявления исследовательских пробелов. Авторы формулируют шесть исследовательских вопросов, охватывающих сферы применения, цели, данные, методы разработки, готовые инструменты и возникающие трудности. Результаты показывают, что большинство работ фокусируется на анализе коммуникаций в open-source проектах (GitHub, Jira, Slack), а доминирующими подходами являются нейросетевые архитектуры (BERT, LSTM) и метод опорных векторов (SVM). Среди готовых решений наилучшие метрики точности и F1-меры демонстрирует BERT. Однако авторы отмечают существенные ограничения: низкая устойчивость к иронии и сарказму, субъективность ручной разметки, деградация качества при кросс-платформенном применении. В контексте управления проектами подчёркивается критическая важност мониторинга эмоционального фона распределённых команд: раннее выявление негативных настроений позволяет менеджерам своевременно корректировать нагрузку, разрешать конфликты и поддерживать мотивацию. В качестве направлений будущих исследований предложены: создание специализированных датасетов из промышленного контекста, разработка гибридных моделей для обработки мультимодальных данных и интеграция инструментов анализа в CI/CD-пайплайны.

3. Постановка задачи и методология решения

3.1. Постановка задачи

Современные цифровые проектные команды функционируют в условиях распределённой коммуникации, где взаимодействие осуществляется через текстовые сообщения, видеоконференции, голосовые обсуждения и совместные онлайн-платформы. Эмоциональное состояние участников проекта проявляется не только в содержании текстов, но и в интонационных характеристиках речи, мимике и невербальных реакциях. Игнорирование данной информации может приводить к накоплению скрытых конфликтов, снижению вовлечённости и увеличению проектных рисков.

В рамках настоящего исследования ставится задача разработки модели мультимодального анализа эмоционально окрашенной информации, предназначенной для мониторинга эмоционального климата проектной команды и выявления потенциальных зон напряжения.

Требуется создать вычислительный механизм, обеспечивающий:

1. Автоматическую классификацию эмоционального состояния участников проекта на основе текстовых, аудио- и видеоданных.
2. Агрегацию результатов анализа по временным интервалам.
3. Формирование интегрального показателя эмоционального климата команды.
4. Выявление аномальных изменений эмоциональной динамики, потенциально связанных с ростом рисков проекта.

Пусть заданы следующие множества:

$T = t_1, t_2, \dots, t_n$ — множество текстовых сообщений.

$A = a_1, a_2, \dots, a_n$ — множество аудиофрагментов.

$V = v_1, v_2, \dots, v_n$ — множество видеосегментов.

$C = c_1, c_2, \ldots, c_k$ — множество классов эмоций (например: радость, гнев, тревога, нейтральность и др.).

Каждому участнику проекта соответствует тройка мультимодальных наблюдений:

$$x_i = (t_i, a_i, v_i)$$

Требуется найти отображение:

$$f: (T \times A \times V) \to C$$

где

$$f(x_i) = c_j$$

— эмоциональный класс, соответствующий совокупности мультимодальных признаков.

Дополнительно вводится агрегированная функция эмоционального индекса команды во временной момент τ:

$$E(\tau) = \frac{1}{N_\tau} \sum_{i=1}^{N_\tau} \omega_i \cdot f(x_i)$$

где:

N_τ — число наблюдений в интервале времени τ.

ω_i — весовой коэффициент (например, с учётом роли участника или достоверности модальности).

$E(\tau)$ — интегральный показатель эмоционального состояния команды.

Задача выявления рисков может быть сведена к задаче обнаружения аномалий во временном ряду:

$$R(\tau) = \begin{cases} 1, & \text{если } | E(\tau) - \mu | > \delta \\ 0, & \text{иначе} \end{cases}$$

где:

μ — средний эмоциональный уровень.

δ — порог чувствительности.

R(τ) — индикатор потенциального риска.

Таким образом, исследуемая проблема сводится к задаче мультимодальной классификации с последующей агрегацией и анализом временной динамики.

3.2. Методология решения

Этап 1. Сбор данных

Источники данных:

- текстовые сообщения (корпоративные чаты, системы управления задачами).
- аудиоданные (записи совещаний).
- видеопотоки (онлайн-встречи).

Данные анонимизируются и структурируются по участникам и временным интервалам.

Этап 2. Предобработка данных

Текстовая модальность:

- токенизация.
- лемматизация.
- удаление стоп-слов.
- нормализация эмодзи и специальных символов.

Аудиомодальность:

- извлечение акустических признаков (MFCC).
- спектральный анализ.
- нормализация уровня громкости.

Видеомодальность:

- детекция лиц.
- выравнивание кадров.
- выделение ключевых точек лица.

Этап 3. Извлечение признаков

Для каждой модальности используются специализированные модели:

- Текст: трансформерные модели (BERT, RuBERT);
- Аудио: wav2vec 2.0, HuBERT;
- Видео: сверточные нейронные сети (ResNet, 3D-CNN).

На выходе формируются эмбеддинги:

$$e_t, e_a, e_v$$

Этап 4. Мультимодальное слияние

Возможны стратегии:

- Раннее слияние — объединение эмбеддингов:

$$e = [e_t \oplus e_a \oplus e_v]$$

- Позднее слияние — объединение предсказаний моделей.
- Attention — механизм для адаптивного взвешивания модальностей.

Этап 5. Классификация эмоций

Используемые модели:

- Multimodal Transformer.
- LSTM для временной динамики.
- XGBoost (в качестве ансамбля).

Функция потерь — кросс-энтропия.

Этап 6. Анализ временной динамики

Формирование временного ряда эмоционального индекса команды.

Применение методов:

- скользящее среднее.
- экспоненциальное сглаживание.
- детекция аномалий.

Этап 7. Визуализация и поддержка принятия решений

Результаты представляются в виде:

- графиков эмоциональной динамики.
- тепловых карт по участникам.
- индикаторов риска.

3.3. Ожидаемые результаты

Предлагаемая методология позволит:

- повысить точность распознавания эмоций за счёт мультимодального подхода.
- выявлять скрытые конфликты на ранних этапах.
- формировать количественные показатели эмоционального климата.
- интегрировать результаты в системы управления проектами.

4. Заключение

В работе рассмотрены современные методы компьютерного анализа эмоционально окрашенной информации, применяемые в задачах управления проектами. Проанализированы подходы к обработке текстовой, аудио- и видеоинформации, а также методы мультимодального объединения признаков на основе глубоких нейронных сетей.

Предложена формальная постановка задачи мультимодального анализа эмоционального состояния проектной команды и разработана концептуальная архитектура решения, включающая этапы предобработки данных, извлечения признаков, их объединения и оценки эмоциональной динамики.

Разработанный подход позволяет интерпретировать эмоциональный фон коммуникаций как количественный показатель и использовать его для выявления потенциальных рисков и поддержки управленческих решений.

Перспективным направлением дальнейших исследований является экспериментальная реализация предложенной модели, подбор оптимальных архитектур нейронных сетей и апробация системы на реальных данных проектных команд.

Список используемых источников

1. Фазульянов Д.В., Гусева А.И. Разработка мультимодального метода сентимент-анализа для поддержки принятия решений в организациях // Современные наукоемкие технологии. – 2024. – № 5-2. – URL: https://top-technologies.ru/ru/article/view?id=40045.
2. Обзор подходов к анализу эмоциональной окраски текстов СМИ // Научный альманах. – 2024. – № 5. – URL: https://na-journal.ru/5-2024-informacionnye-tekhnologii/11975-obzor-podhodov-k-analizu-emocionalnoi-okraski-tekstov-smi.
3. Ram J. Sentiment analysis for project stakeholder management // IPMA World. – 2025. – URL: https://ipma.world/sentiment-analysis-for-project-stakeholder-management/.
4. Мажит А.Н., Кожамкулова Ж.Ж. Детектирование лиц и распознавание эмоций с использованием алгоритмов компьютерного зрения // Вестник науки и образования. – 2024. – URL: https://www.xn----8sbempclcwd3bmt.xn--p1ai/article/19904.
5. Шурдилов И.С., Нуриев М.Г., Лаптева М.Г., Толмачева А.В. Разработка системы распознавания эмоций по лицевым выражениям на основе машинного обучения // International Research Journal. – 2025. – № 6(156). – DOI: 10.60797/IRJ.2025.156.52.
6. Savchenko A.V. HSEmotion Team at ABAW-8 Competition: Audiovisual Ambivalence/Hesitancy, Emotional Mimicry Intensity and Facial Expression Recognition // URL: https://arxiv.org/pdf/2503.10399.
7. Исаева Е.В., Семенов С.В., Черных Д.Л., Гудовщиков А.В. Семиотический анализ текстов и интерпретация знаковых систем в цифровую эпоху: Sentiment-анализ с использованием платформы KNIME // Дискурс. – 2025. – № 4. – С. 121–138. – URL: https://discourse.etu.ru/assets/files/diskurs_4_2025-121-138.pdf.
8. Фадел В. Нейросетевые алгоритмы распознавания эмоций на основе сигналов ЭЭГ : магистерская диссертация / науч. рук. В.Г. Спицын. – Томск: ТПУ, 2023. – URL: http://earchive.tpu.ru/handle/11683/76400.
9. Rahul N.A. Strategic Alignment of Emotional Intelligence with Project Management: A Framework for Enhancing Project Success and Team Performance // Psychology Research and Behavior Management. – 2025. – Vol. 14. – P. 1–25. – URL: https://www.scirp.org/journal/paperinformation?pa perid=132717.
10. Лукашевич Н.В. Автоматический анализ тональности текстов: проблемы и методы //

Интеллектуальные системы. Теория и приложения. – 2022. – Т. 26, № 1. – С. 1–5. – URL: https://intsysmagazine.ru/pdfs/26-1/1-5-Lukashevich.pdf.

11. Vishnu V.M., Kumar P., Prajwal T.H., Prakruthi C.R., Rakshan S. Facial Emotion Recognition System // International Journal of Progressive Research in Engineering Management and Science (IJPREMS). – 2025. – Vol. 05, Issue 11. – P. 1727–1733. – DOI: 10.58257/IJPREMS45078.

12. Павлюченко К.А. Анализ методов выявления эмоционального контекста в текстах // Вестник БГУИР. – 2024. – URL: https://libeldoc.bsuir.by/bitstream/123456789/58097/1/Pavlyuchenko_Analiz.pdf.

13. Михненко П.А. Анализ мультимодальных данных в управлении проектами: перспективы использования машинного обучения // Управленческие науки. – 2023. – Т. 13, № 4. – С. 71–89. – DOI: 10.26794/2304-022X-2023-13-4-71-89. – URL: https://cyberleninka.ru/article/n/analiz-multimodalnyh-dannyh-v-upravlenii-proektami-perspektivy-ispolzovaniya-mashinnogo-obucheniya.

14. Cheng B., Arora C., Liu X., Hoang T., Wang Y., Grundy J. Multi-Modal Emotion Recognition for Enhanced Requirements Engineering: A Novel Approach // 2023 IEEE 31st International Requirements Engineering Conference Workshops (REW). – 2023. – URL: https://arxiv.org/pdf/2306.01492.

15. Nandwani P., Verma R. A review on sentiment analysis and emotion detection from text // Social Network Analysis and Mining. – 2021. – Vol. 11, № 1. – Art. 81. – DOI: 10.1007/s13278-021-00776-6. – URL: https://pmc.ncbi.nlm.nih.gov/articles/PMC8402961/.

16. Kratzwald B., Feuerriegel S., Prendinger H. et al. Sentiment Analysis in Software Engineering: A Systematic Mapping Study // arXiv preprint. – 2025. – arXiv:2502.07893. – URL: https://arxiv.org/pdf/2502.07893.

Разработка системы усовершенствованного управления процессом очистки углеводородов

Муравьева Е. А.
ИХТИ УГНТУ в г. Стерлитамаке
Стерлитамак, Россия
e-mail: muraveva_ea@mail.ru

Шарипов М. И.
ИХТИ УГНТУ в г. Стерлитамаке
Стерлитамак, Россия
e-mail: sharipovm@mail.ru

Аннотация[1]

Для повышения стабильности и эффективности технологического процесса в установке гидроочистки дизельного топлива разрабатывается и моделируется интеллектуальная система управления реакторной зоной на основе искусственной нейронной сети. Сложность нелинейного процесса и сильное взаимное влияние технологических параметров делают существующие системы управления на базе ПИД-регуляторов неэффективными, что требует создания более совершенных систем адаптивного управления.

Моделирование, проведенное в рамках исследования, включало в себя применение математических и имитационных методов. Для обучения нейронной сети применялась среда Matlab с Neural Network Toolbox. Программные продукты AIDA Pro и SMOC Pro использовались для создания модели объекта управления и анализа его работы в оптимизационном режиме. По результатам обучения, нейронная сеть достигла высокой точности среднеквадратичной ошибки, равной $3{,}3775 \cdot 10^{-7}$. Проведенное имитационное моделирование в SMOC Pro подтвердило способность системы поддерживать ключевые параметры в установленных технологических пределах при возникновении внешних воздействий, например, температуру в реакторе и концентрацию серы в продукте.

При использовании разработанного интеллектуального комплекса на базе нейронной сети для управления процессом гидроочистки обеспечивается стабильное качество готовой продукции, преодолевая недостатки традиционных ПИД-регуляторов. Эффективность данного комплекса подтверждена в заключении и позволяет достичь высокой стабильности в работе.

Ключевые слова: гидроочистка дизельного топлива, адаптивное управление, искусственная нейронная сеть, математическое моделирование, ПИД-регулятор, оптимизация технологического процесса.

1. Введение

Современные нефтеперерабатывающие производства предъявляют высокие требования к эффективности и стабильности технологических процессов, одним из ключевых среди которых является гидроочистка дизельного топлива. Основная задача этого процесса — снижение содержания серы в готовом продукте — осложняется его сложным нелинейным характером и сильной взаимосвязью множества технологических параметров, таких как температура и давление в реакторе, расход сырья и хладагента.

Традиционные системы автоматизации, построенные на классических ПИД-регуляторах, оказываются не в полной мере адекватны для управления подобными объектами. Их фундаментальное ограничение заключается в неспособности адаптироваться к нелинейностям и учитывать перекрестные влияния параметров, что приводит к колебаниям режима, снижению качества продукции и преждевременной деактивации дорогостоящего катализатора. В связи с этим актуальной задачей является разработка и внедрение интеллектуальных систем адаптивного управления, способных компенсировать указанные недостатки.

Перспективным направлением для создания таких систем является использование искусственных нейронных сетей, обладающих свойствами универсальных аппроксиматоров нелинейных функций и способных обучаться на основе экспериментальных данных. Однако успешная реализация нейросетевого подхода требует решения комплекса задач, включая обоснование самой целесообразности замены традиционных систем, разработку точной и адекватной нейросетевой модели, создание достоверной динамической модели объекта

[1] Труды XI Международной научной конференции "Информационные технологии интеллектуальной поддержки принятия решений", 11-13 ноября, Уфа-Чандигарх-Фандербейлпарк, 2025

управления, а также всестороннюю проверку эффективности всего интеллектуального комплекса в условиях, приближенных к реальным.

Таким образом, целью настоящего исследования является разработка и моделирование интеллектуальной системы адаптивного управления реакторной зоной установки гидроочистки на базе искусственной нейронной сети. Достижение этой цели предполагает последовательное решение задач по созданию архитектуры нейросети и ее обучению, идентификации динамической модели объекта в специализированном программном обеспечении и проведению имитационного моделирования для верификации работоспособности и устойчивости разработанного комплекса.

2. Основные закономерности процесса гидроочистки дизельного топлива и определение входных и выходных данных

При определенных температурах происходит реакция обессеривания. Если в реакторе Р-1 температура ТЕ 7-1 превышает норму, интенсивность реакций гидроочистки снижается, что приводит к образованию отложений кокса на катализаторе и сокращению срока его эксплуатации. Уменьшение температуры ТЕ 7-1 ниже нормы замедляет процесс и приводит к выпуску низкокачественной продукции. Повышение давления РЕ 8-1 способствует качественной гидроочистке топлива, сокращению образования кокса в реакторе и увеличению продолжительности его работы.

Изменения в объеме подачи охлаждающего воздуха, регулируемые клапаном 25-3, оказывают влияние на температуру и давление в реакторе, что отражается на общем ходе процесса. Также изменения в давлении подачи топлива в печь, регулируемые клапаном 6-3, влияют на температуру в печи и в дальнейшем на температуру в реакторе, что в свою очередь воздействует на процесс очистки дизельного топлива [1].

Эффективность всего технологического процесса напрямую зависит от работы реактора Р-1. Чтобы обеспечить его длительную и стабильную работу, необходим контроль и регулирование ключевых параметров. К ним относятся давление на подаче топлива, которое настраивается клапаном 6-3, температура продуктов на выходе из печи, также корректируемая этим же клапаном, температура реакции в самом реакторе ТЕ 7-1 и объем охлаждающего воздуха, подаваемого через клапан 25-3.

Применение традиционных ПИД-регуляторов для управления этой сложной системой имеет существенные ограничения. Главным недостатком является то, что они не способны адаптироваться к нелинейным характеристикам процесса и учитывать сильную взаимосвязь между различными параметрами установки гидроочистки дизельного топлива. Стоит подчеркнуть, что реакторная установка классифицируется как опасный производственный объект, а ее управление представляет собой сложную задачу. Достижение высокого качества технологического процесса возможно только при условии комплексного учета взаимного влияния всех его операционных переменных.

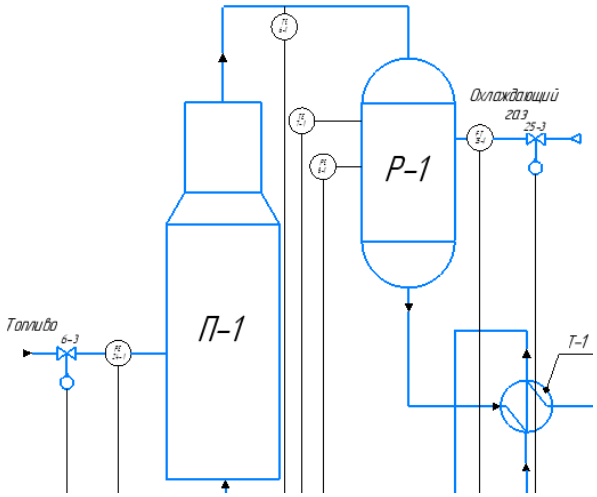


Рис. 1. Функциональная схема автоматизации работа реактора в процессе гидроочистки дизельного топлива

Для решения задачи с помощью нейронной сети необходимо собрать данные для обучения. Выходные параметры имеют прямую зависимость от соответствующих входных.

Далее производится подготовка данных для обучения нейронной сети.

Таблица 1. – Входные параметры процесса работы реактора гидроочистки дизельного топлива

Входные параметры			
Позиция	**Наименование**	**Пределы регулирования**	**Единицы измерения**
ТЕ 7-1	Температура в реакторе Р-1	380-420°С	°С
РЕ 24-1	Давление топлива, входящий в печь П-1	2,5-6	МПа
FT 25-1	Расход охлаждающего газа, подаваемого в реактор Р-1	100-200	$м^3/с$

ТЕ 6-1	Температура на выходе из печи	380-425	°С
АЕ 26-1	Концентрация серы в готовом продукте	0,1-0,3	% масс

Таблица 2. – Выходные параметры процесса работы реактора гидроочистки дизельного топлива

Выходные параметры			
Позиция	**Наименование**	**Пределы регулирования**	**Единицы измерения**
25-3	Регулирующий клапан подачи охлаждающего воздуха	0…100	%
6-3	Регулирующий клапан подачи топлива	0…100	%

1.2 Выбор типа и архитектуры нейросети

Наша сеть будет состоять из 5-ти слоев.

Первый шаг – это выбор структуры нейронной сети. Будет использована каскадная сеть с прямым распространением сигнала и обратным распространением ошибки (Cascade-forward backprop).

Далее необходимо выбрать количество нейронов в скрытом слое. В данном случае необходимое число нейронов равное 5.

Обучение сети будет производиться по модифицированному алгоритму оптимизации, направленному на решение задач о наименьших квадратах: Levenberg-Marquardt = 'trainlm'.

Затем выбираются параметры обучения сети, такие как, максимальное количество эпох, количество эпох между показами и параметр достижения цели. Данные параметры нужно выбирать экспериментально, руководствуясь при этом критерием завершения обучения [6,8,9].

Для завершения процесса обучения был установлен параметр net.trainParam.epochs равным 500, определяющий максимальное количество эпох обучения. Количество эпох между показами выбрано равным 25, устанавливая значение параметра net.trainParam.show.

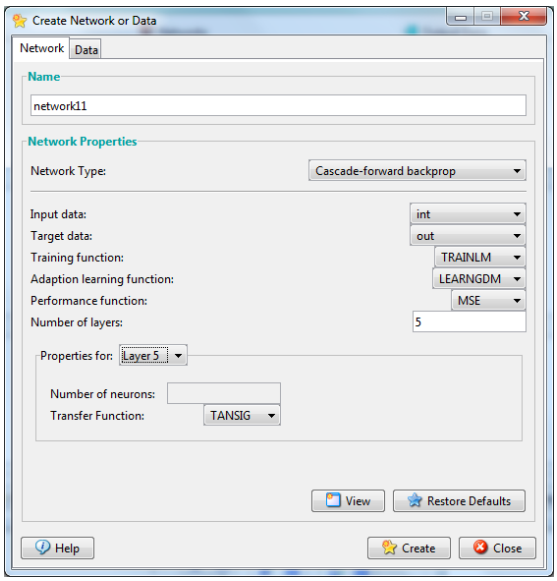


Рис. 2. Настройка параметров обучения нейронной сети

Чтобы завершить обучение, был задан целевой параметр отклонения net.trainParam.goal равным 0.

1.3 Построение и обучение сети в Matlab

Было реализовано обучение нейронной сети в Matlab.

На рисунке 3 показана выбранная структура искусственной нейронной сети. На вход нейронной сети поступает сигнал x, в текущем случае, это температура в реакторе, расход охлаждающего газа в реактор, давление топлива, подаваемого в печь, температура на выходе из печи, качество готового топлива. Сумматор «+» умножает каждый вход bi на вес wi и суммирует взвешенные входы. Затем значение проходит через функцию активации соответствующего слоя и рассчитывается выход: степень открытия клапана 25-3, степень открытия клапана степень открытия клапана 6-3.

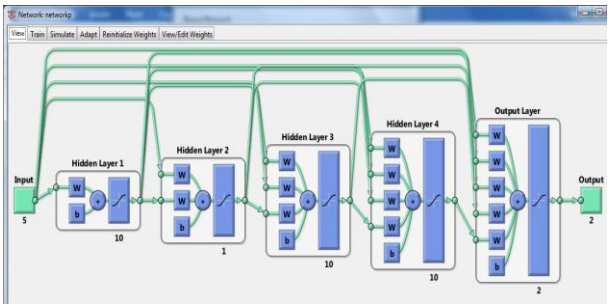


Рис. 3. Нейронная сеть

При нажатии на кнопку Performance в окне процесса обучения сети отображается график, показывающий изменение ошибки обучения (рисунок 4). На рисунке 2 видно, что как обучающее, так и тестовое множества сначала снижаются одинаково до 10 эпохи, а затем с небольшими отличиями приходят к минимальной ошибке. На 701-й эпохе обучения была зафиксирована конечная среднеквадратичная

погрешность в размере $3{,}3775 \cdot 10^{-7}$. Данная величина, являясь незначительной, свидетельствует об успешной адаптации нейронной сети к задаче управления процессом эфирации. Сравнительный анализ с предыдущими результатами показывает более позднее достижение минимальной ошибки, что говорит о следующем: применение упрощенной сетевой структуры для обработки значительных объемов информации снижает скорость выхода на точные показатели [3].

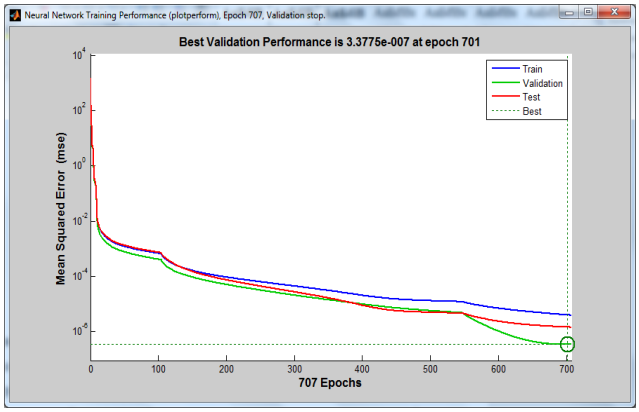


Рис. 4. График средней квадратичной ошибки

Еще одним инструментом оценки результата обучения нейронной сети может быть построение функций регрессии результатов (рисунок 5).

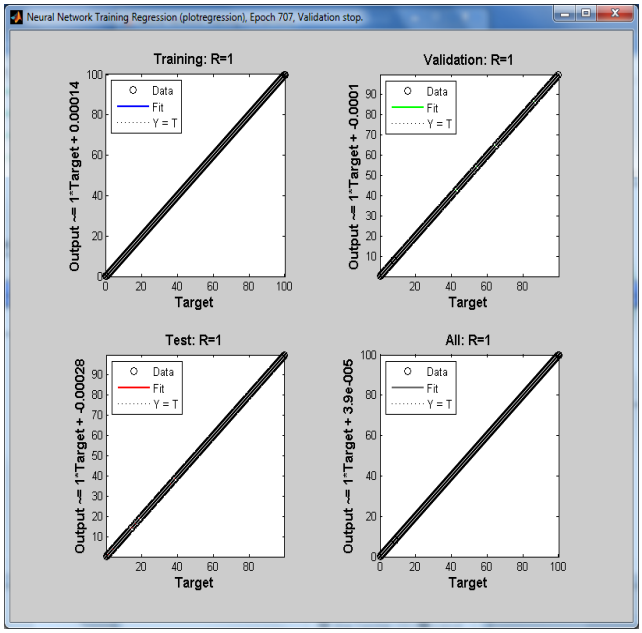


Рис. 5. Схема процесса обучения нейронной сети

Высокая точность нейронной сети подтверждается коэффициентом корреляции R, равным 1 и 1, что указывает на сильную связь между переменными. Результат обученной нейронной сети поражает своей точностью. Благодаря зависимости первой нейронной сети от первой достигается наилучшая точность, что и доказывают приведенные графики. Первый график служит явным показателем точности обучения и тестирования нейросети при коэффициенте градиента, стремящемся к нулю. В данном случае, Gradient = 0,0068642. Показатель выступает минимальным значением, указывая на повышенную точность обучения нейронной сети.

График "Validation Checks" отображает изменение ошибок на проверочном наборе данных. Визуальный анализ графика показывает, что были зафиксированы только три провала в процессе валидации нейронной сети, что не оказало существенного влияния на итоговые результаты.

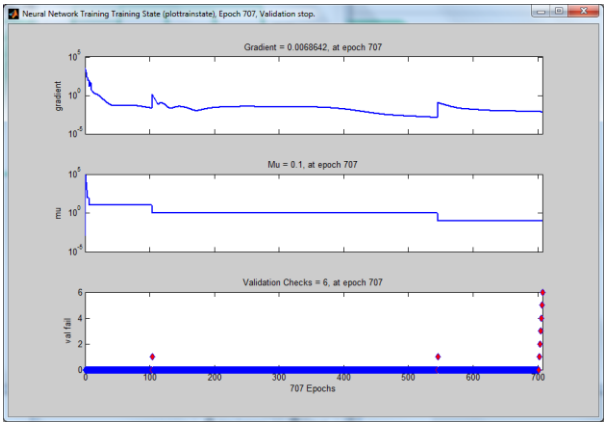


Рис. 6. Графики линейной регрессии результатов обучения

График «Mu» показывает изменение параметра обучения μ методом байесовской регуляризации, и чем ниже данное значение μ, тем точнее будет проходить обучение сети. Во втором случае обучение происходит точнее.

1.4 Проверка нейронных сетей

Для тестирования нейронных сетей было подано на вход 5 значений.

После выполнения команды для первой нейронной сети, было получено 5 значений (380,3732; 2,5326648; 100,9324; 0,10186648, 380,42). Близость полученных значений к заданному результату (0,93324; 0,93324), свидетельствует о применимости сети.

Таблица 3. – Проверочные данные

Вход				
TE 7-1	PE 24-1	FT 25-1	AE 26-1	TE 6-1
380,37324	2,5326648	100,9324	0,10186648	380,42
Выход				
253		63		
0,93324		0,93324		

Разработка интеллектуального комплекса для управления параметрами процесса гидроочистки дизельного топлива, основанная на нейронной сети, дает возможность эффективного контроля за работой реактора. Модель управления параметрами процесса учитывает сложные взаимосвязи между различными факторами производства и

управляющими сигналами. Использование автоматических систем контроля параметров технологического процесса гидроочистки дизельного топлива с учетом взаимодействия различных параметров объекта управления является обоснованным. В результате исследования подтверждена целесообразность использования такого подхода к оптимизации процесса очистки дизельного топлива.

2 Получение модели взаимосвязи входных и выходных параметров в пакете AIDA^{Pro} для управления работой реактора процесса гидроочистки дизельного топлива

Была создана рабочая область AIDAPro и введены текущая дата, время и набор данных. Набор данных состоит из 7500 примеров.

После загрузки данных с csv открылись теги трендов. Тренды тегов можно увидеть на вкладке trends рабочей области. Эти тренды можно выводить в любой последовательности, масштабировать во времени. На рисунке 7 мы видно, как изменяются значения параметров TE 7-1, PE 24-1, FT 25-1, TE 6-1, AE 26-1 в зависимости от увеличения или уменьшения степени открытия клапанов 25-3, 6-3.

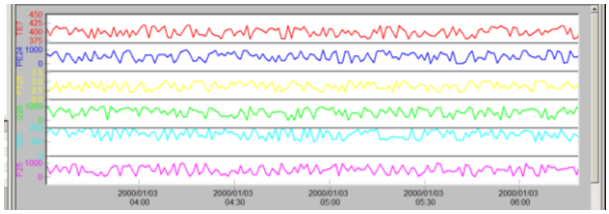


Рис. 7. Окно просмотра трендов переменных

Перед идентификацией модели необходимо создать новые группы аппроксимации параметров FT25, PE24, A26, TE6, TE7 «Fits» во вкладке «Fit Groups». Для достижения минимального отклонения от всех точек в экспериментально полученных данных с погрешностью в программе "Fits" проводится кривая, не обязательно проходящая через каждую точку. Важно определить входные и выходные переменные во вкладке "Fits" и настроить параметры и зависимости входных и выходных параметров работы реактора гидроочистки дизельного топлива во вкладке "FIR-FIT". После этого следует инициализировать параметризацию в модели для быстрой проверки параметров процесса и предотвращения основных ошибок.

Сначала необходимо применить непараметрический многомерный подход для подбора данных, который не требует соответствия определенным предположениям или параметрам в анализируемом наборе данных. Затем происходит уменьшение модели для каждого отношения "один вход - один выход" (SISO) до параметрической формы. На вход подается TE 7, и на выходе получается клапан 6, аналогично поступают с другими значениями (PE 24 приводит к F6; FT 25 приводит к F25; TE 6 приводит к F6; AE 26 приводит к F25). После этого будут представлены графики, показывающие прогнозируемые связи, которые позволят сделать вывод о правильности модели и, при необходимости, внести изменения.

В приведенном ниже рисунке 4.7 FT25, PE24, A26, TE6, TE7 - это входные данные, а F25, F26 - выходные. Установив флажки, как показано ниже, AIDAPro будет определять модели между переменными.

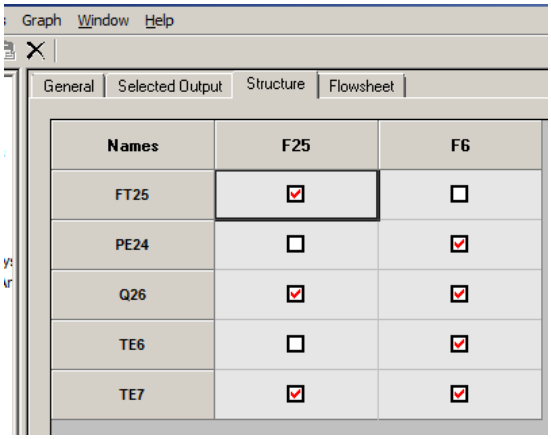


Рис. 8. Определение зависимостей

На вкладке "Fit Analysis" можно найти различные методы анализа и сравнения результатов аппроксимации параметров FT25, PE24, A26, TE6, TE7 для разных групп. Узел анализа соответствия применяется в случае наличия нескольких групп соответствия, в то время как анализ соответствия используется лишь для одной группы соответствия. FIR Fit Setup - это вложенная вкладка, где происходит настройка и запуск непараметрической FIR-аппроксимации параметров. Это позволяет получить первичную обратную связь относительно качества как данных, так и выбранной структуры модели. Первая модель, которая создается на этой вкладке, на всех последующих графиках именуется как "FIR".

Вкладка "Графики аппроксимации FIR" показывает результаты аппроксимации параметров FT25, PE24, A26, TE6, TE7 в виде ступенчатых графиков. На графиках TE7-F6, PE24-F6, FT25-F6, TE6-F6, AE26-F6, TE7-F25, AE26-F25 видно ступенчатое изменение. Графики TE6-F25, PE24-F25 отсутствуют, потому что они не зависят от клапана F25. Здесь можно менять порядок параметров и выбирать нужные графики для анализа.

Для выполнения параметрической аппроксимации данных FT25, PE24, A26, TE6, TE7 в окне Parametric Fit необходимо сначала провести аппроксимацию FIR на предыдущей вкладке, чтобы получить хорошую начальную оценку параметров. Окно Parametric Fit направлено на упрощение задачи анализа нелинейности системы путем рассмотрения

эквивалентной линейной системы с случайными параметрами. Его основная цель заключается в предоставлении краткого теоретического объяснения выявленным фактам, полученным в процессе моделирования.



Рис. 9. Графики ступенчатой характеристики

Окно Overlay представляет собой набор графических изображений переходных характеристик, снабженных таблицей легенды ниже графиков. Модель второго порядка типа бета является специфическим случаем вторичной модели. Эта модель полезна для анализа динамических реакций высокоуровневых моделей, когда возможен либо начальный обратный динамический отклик, либо реакция с перерегулированием без колебаний. Обратный динамический отклик наблюдается в случае, если значение β меньше нуля. Значения ступенчатой характеристики отображаются в окне Parametric Fit. Для TE6-F25 Gain=0,3495, Tau=49,88, Dt=2,013. Для FT25-F25 Gain=-221,2, Tau=131,9, Dt=2,0. Для TE6-F25 Gain=0,3495, Tau=49,88, Dt=2,013. Для FT 25-F6 Gain=0, Tau=15,17, Dt=0. Для TE 6-F6 Gain=0, Tau=30,33, Dt=0.

Окно Simulation отображает прогнозы и тренды выходных данных, которые связаны с группой параметров, подлежащих аппроксимации. Эффективность аппроксимации можно оценить по степени соответствия прогнозов текущим значениям выходных данных. Также здесь доступны изменения входных данных, подлежащих анализу. На первом графике показана симуляция зависимости между данными TE7, PE24, FT25, TE6, AE26 и открытием клапана F25, а на втором графике - между данными TE7, PE24, FT25, TE6, AE26 и открытием клапана F6. В этом окне можно делать прогнозы выходных данных и наблюдать за изменениями их тенденций.

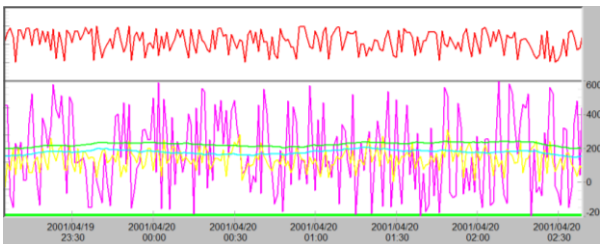


Рис. 10. График моделирования

Error – это разница между фактическим и прогнозируемым выходными данными (основанная только на детерминированной части модели) (рисунок 11). Модель можно считать хорошей, если ошибки распределены случайным образом.

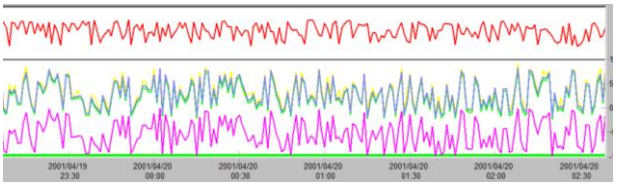


Рис. 11. График ошибки параметра AE26

На рисунке 11 показана разница между фактическими данными и данными прогноза. На графиках показана разница между фактическим значением AE 26 и данными прогноза, на который влияет степень открытия клапана F25.

Окно Cross Correlations (Взаимные корреляции) используется для выполнения анализа взаимной корреляции между аппроксимированными выходными параметрами и выходными данными (рисунок 12).

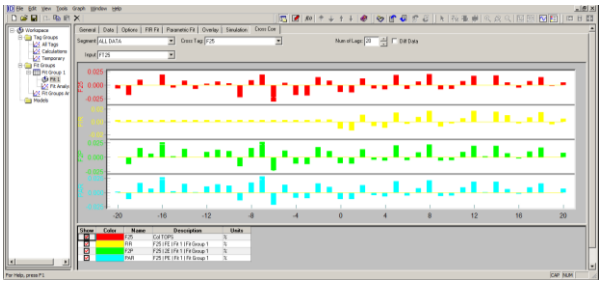


Рис. 12. Окно Cross Correlations

На рисунке 12 можно увидеть графики взаимных корреляций клапана F25, F2P, FIR и PAR. На рисунке 12 отражена взаимная корреляция между аппроксимированными выходными параметрами F25, F2P, FIR и PAR и выходными данными TE7, PE24, FT25, TE6, AE26. Взаимные корреляции являются мерой сходства двух рядов в зависимости от перемещения одной относительно другой.

Окно Fit Select / Export состоит из матрицы графиков переходной характеристики с таблицей легенды под графиками (рисунок 13). Для каждого отношения ввода-вывода обозначен результат соответствия как модель для этой пары ввода-вывода.

На рисунке 13 видны графики переходных характеристик всех параметрических

идентифицированных моделей. На данном окне видно графики связей между входными и выходными переменными. На рисунке 13 видно, что на графиках зависимости TE7 от F6, PE24 от F6, FT25 от F6, TE6 от F6, AE26 от F6 изменения параметров TE7, PE24, FT25, TE6, AE26 зависят от изменения открытия клапана 6-3, а на графиках зависимости FT25 от F25, TE6 от F25, AE26 от F25 изменения параметров FT25, TE6, AE26 зависят от изменения открытия клапана 25-3 и имеют ступенчатую характеристику.

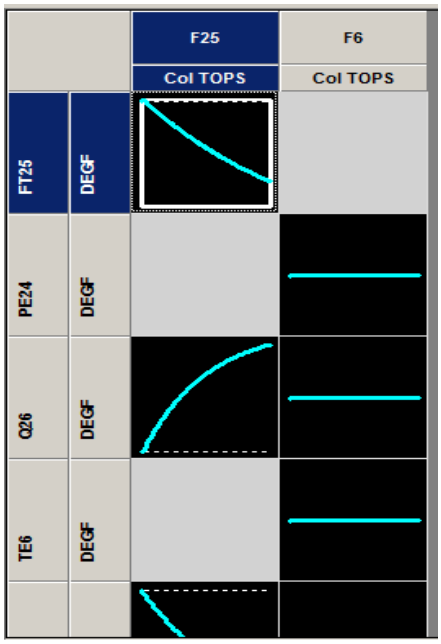


Рис. 13. Окно Fit Select/Export во вкладке Fit Analysis

Поля TE 6-1-25-3 PE 24-1-25-3 не имеют графиков, так как данные параметры не зависят от клапана 25-3. В данном окне можно изменять порядок параметрической модели и выбрать оптимальные графики.

В созданной модели, в окне Models, отображаются все отношения ввода-вывода и соответствующие им модели ступенчатой реакции. На рисунке 14 изображены тренды зависимостей TE7 от F6, PE24 от F6, FT25 от F6, TE6 от F6, AE26 от F6, FT25 от F25, TE6 от F25, AE26 от F25. На графиках отображаются те порядки ступенчатой функции, которые были выбраны. На графике зависимости FT25 от F25, TE7 от 25, AE26 от 25, AE26 от 6 построен график первого порядка. На графике зависимости TE7 от 6, PE24 от 6, FT25 от 6, TE6 от 6 построен график второго порядка. Так как между FT25 и F6, PE24 от 25, TE6 от 6 не существует взаимосвязи, она обозначается как нулевая модель (Zero Order).



Рис. 14. Окно Plots во вкладке Model

После всех проделанных шагов можно экспортировать получившуюся модель в SMOC^{Pro}. Для этого нужно выбрать Export во вкладке File и далее Export to SMOC.

2.2 Входные и выходные переменные для реактора гидроочистки дизельного топлива в контроллере SMOC^{Pro}

На рисунке 15 изображена модель процесса, которая экспортируется из программы AIDAPro. В этой модели показан реактор, куда поступают значения манипулируемых переменных, таких как расход охлаждающего газа (MV1) и температура на выходе из печи (MV2). Выходные переменные процесса включают в себя температуру в реакторе (POV1), температуру на выходе из печи (POV2) и концентрацию серы в готовом продукте (POV3). Фильтрация возмущений ПИД-регулятора применяется к температуре в реакторе TE7 и концентрации серы AE26 в готовом продукте. Поэтому необходимо внести изменения в сценарий задания по концентрации серы AE26 в готовом продукте, уменьшив значение с 30 до 20 на шаге 5 или с 10 до 20 на этом же шаге (см. рисунок 17,18).

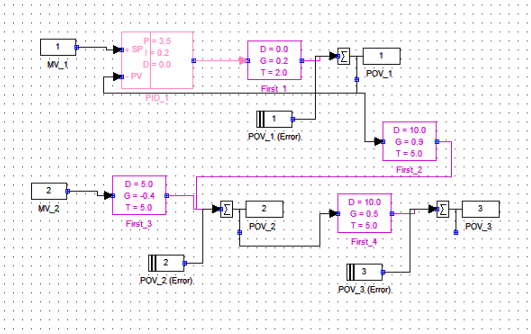


Рис. 15. Модель реактора с ПИД-блоком

В Модели реактора для обеспечения наиболее высокого быстродействия в системе использованы ПИД-регуляторы.

Анализ модели выявил потенциальный источник ошибок, такой как неизмеряемые возмущения на каждом выходе процесса, которые в текущей конфигурации представлены суммарным белым шумом. Отдельно генерируемое для ПИД-контура возмущение необходимо для моделирования алгоритма фильтрации помех, встроенного в ПИД-регулятор.

На рисунке 16 приводится отклик выходных переменных процесса (температура в реакторе (POV1), температура на выходе из печи (POV2), концентрация серы в готовом продукте (POV3)), на соответствующие им реакции на скачок параметра POV1(Error), POV2(Error), POV3(Error) (значения ошибок температуры в реакторе, температура на выходе из печи, концентрация серы в готовом продукте). Фильтрация возмущений ПИД-регулятора распространяется на температуру в реакторе (POV1) и концентрация серы в готовом продукте (POV3) при помощи динамических связей.

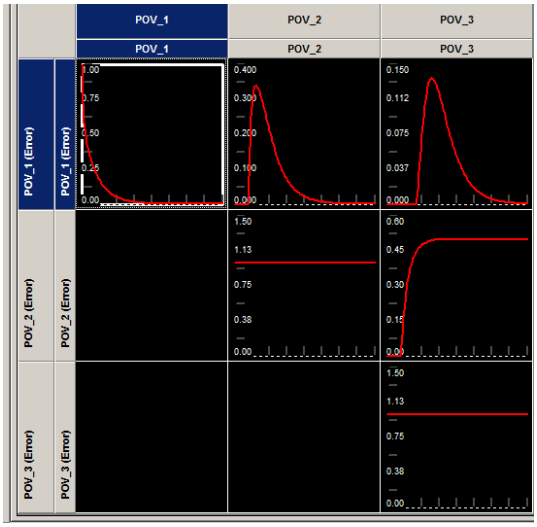


Рис. 16. Отклик параметра POV_1, соответствующий реакции на скачок параметра POV_1 (Error)

На вкладке Sub_Debutanizer было смоделировано поведение модели работы реактора в режиме оптимизации. Для этого был создан сценарий, где пропишем временные шаги, когда будут изменяться параметры.

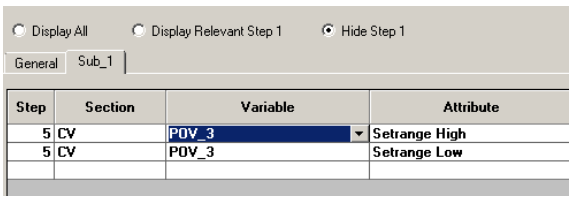


Рис. 17. Вкладка Subcontroller (Подконтроллер)

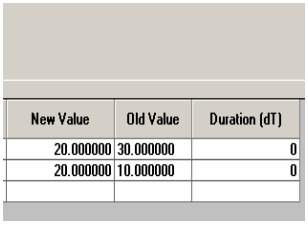


Рис. 18. Вкладка Subcontroller (Подконтроллер)

Фильтрация возмущений ПИД-регулятора распространяется на температуру в реакторе TE7 и концентрацию серы AE26 в готовом продукте, поэтому следует изменить в сценарии задания по концентрации серы AE26 в готовом продукте (с 30 до 20 на шаге 5 или с 10 до 20 на шаге 5) (рисунок 17,18).

На рисунке 8 представлены результаты моделирования при заданных характеристиках. В устойчивом режиме каждый график остается в пределах спецификации (таблица 1) благодаря использованию ПИД-регулятора. Эффективность фильтрации возмущений демонстрируется на данных графиках. После завершения моделирования на шаге 150, TE7 достигает 94,52°C, FT25 - 458,10 м3/с, TE6 - 614,99 °C, а AE26 - 82,49% массы. В таблице 1 представлены диапазоны значений для нормальной работы реактора процесса гидроочистки дизельного топлива.

3 Заключение

1. Разработана и апробирована архитектура интеллектуальной системы адаптивного управления на основе искусственной нейронной сети для сложного, нелинейного технологического процесса гидроочистки углеводородов. Данный подход преодолевает фундаментальные ограничения классических ПИД-регуляторов, обеспечивая учет взаимного влияния множества технологических параметров.

2. Экспериментально подтверждена высокая прогнозная точность разработанной нейросетевой модели прямого распространения, достигнувшая значения среднеквадратичной ошибки на уровне $3{,}3775 \cdot 10^{-7}$ при обучении в среде Matlab, что свидетельствует о ее адекватности и применимости для задач управления в реальном времени.

3. Создана и идентифицирована динамическая модель объекта управления в программном пакете AIDA Pro, которая адекватно отражает нелинейные взаимосвязи между ключевыми технологическими параметрами, такими как температуры в реакторе и печи, давление, расход хладагента и концентрация серы в продукте.

4. Проведенным имитационным моделированием в контроллере SMOC Pro доказана операционная эффективность разработанного интеллектуального комплекса, подтверждена его способность поддерживать выходные параметры процесса в установленных технологических пределах и

обеспечивать устойчивость работы при внешних возмущениях.

5. Теоретически и практически обосновано, что предложенный нейросетевой комплекс управления позволяет повысить стабильность технологического режима и гарантированное качество готовой продукции за счет комплексного учета нелинейности и взаимовлияния параметров, что недостижимо в рамках традиционных систем автоматизации.

Список используемых источников

1. Жандалдаев И. К., Калдарбек М. А., Наурызбаев К. К. Оптимизация системы управления процессом очистки и подготовки нефти // Международный журнал «Вестник науки» №6 (87) Том 7, 2025 (материалы журнала)
2. Таланов Н., Хлопонина В., Федоров М. С. Систематическое и аналитическое исследование методов анализа и определения технического состояния сложных технических систем // V Международная конференция по управлению в технических системах (CTS), 2023 (материалы конференции)
3. Муравьева Е. А. Адаптация нечеткого контроллера // Международная конференция «Инновации и перспективы развития горного машиностроения и электротехники» (IPDME), 2017 (материалы конференции)
4. Сахаров М., Коледина К., Губайдуллин И. Оптимальное управление процессом гидрирования углеводородов с помощью катализаторов // Математика, том 12, выпуск 22, 2024 (материалы журнала)
5. Основные технологии переработки нефтегазового сырья / М.А. Косарева, С.Г. Стахеев, Н.А. Третьякова. — Издательство Уральского Университета, 2022. — 20с. (книга)
6. Технология переработки нефти и газа / Е.Н. Ивашкина, Е.М. Юрьев, Н.И. Кривцова, Н.С. Белинская. — Издательство Томского политехнического университета, 2021. — 10с. (книга)

Анализ информации о геологическом разрезе на основе геоинформационных технологий и ИИ

Насырьянова Р. Б.
УУНиТ
Институт информатики,
математики и робототехники
Уфа, Россия
e-mail: nasyrianovar@mail.ru

Соколова А. В.
УУНиТ
Институт информатики,
математики и робототехники
Уфа, Россия
e-mail: beldyanova@yandex.ru

Аннотация[1]

В данной статье рассмотрены основные определения и методы моделирования геологических разрезов. Описывается какую информацию можно получить из графика профиля разреза, организация обработки данных об геологическом разрезе, описаны возможности искусственного интеллекта в геологии, а также представлен обзор разработанного веб-ресурса визуализации геологического разреза Мечетлино.

Ключевые слова: геологический разрез; график профиля разреза; визуализация; искусственный интеллект; веб-ресурс; разрез Мечетлино.

1. Введение

Геологический разрез является одним из ключевых инструментов в геологии, позволяющим визуализировать строение земных недр на заданном участке. Он отражает последовательность залегания горных пород, их литологический состав, тектонические нарушения и другие особенности геологического строения, геологический разрез хранит историю нашей планеты, так как каждый слой горных пород формировался в определенных условиях и содержащиеся в нем окаменелости позволяют ученым реконструировать древние ландшафты, климат и биологическое разнообразие прошлого [1]. Поэтому знания, полученные в результате их изучения, могут быть применены в различных сферах, в том числе в научных, просветительских и образовательных.

Существует несколько способов представления и описания (текстовый, табличный, графический и др.) геологического разреза, а современные методы и технологии позволяют автоматизировать процесс не только построения, но и анализа структуры геологического разреза [2]. Способ реализации зависит от исходных данных, масштаба моделируемой территории и задач, которые будет решать цифровая модель геологического разреза.

Так, например, сбор данных для изучения геологического разреза происходит разными методами – геофизические исследования, бурение скважин, анализ керна и геологическая съемка на местности и другие. Каждый из них дает свои характеристики данных – глубину залегания пород, их тип, толщину слоев, состав слоя и пространственные координаты на местности. Затем полученные данные обрабатываются, систематизируются и используются для визуализации геологического разреза, например, в виде графика профиля, который позволяет визуально оценить геологическое строение исследуемой территории. Из него можно узнать следующую информацию [3]:

- Какие породы залегают на данной территории и на какой глубине.
- Каковы толщины слоев и их протяженность вдоль профиля.
- Есть ли признаки тектонических нарушений или складчатости.
- Где располагаются водоносные горизонты или залежи полезных ископаемых.

В рамках данного исследования будет рассмотрено совместное применения геоинформационных, веб-технологий и искусственного интеллекта для построения и анализа графика профиля на примере геологического разреза Мечетлино.

2. Описание моделируемой территории

Разрез Мечетлино представляет собой важный объект для изучения, так как характеризуется сложной стратиграфией и наличием различных литологических комплексов. В данном геологическом разрезе сохранились отложения различных геологических эпох, содержащие богатый палеонтологический материал. Изучение этих слоев помогает не только понять геологическую историю региона, но и восстановить этапы эволюции жизни на Земле [4].

[1] Труды XI Международной научной конференции "Информационные технологии интеллектуальной поддержки принятия решений", 11-13 ноября, Уфа-Чандигарх-Фандербейлпарк, 2025

Разрез Мечетлино обладает уникальными характеристиками, в 2018 году разрезу был присвоен статус «Золотой гвоздь» под эгидой ЮНЕСКО и Международной комиссии по стратиграфии. Уникальность геологического разреза Мечетлино заключается в следующем:

- Это эталонный разрез нижней границы кунгурского яруса пермской системы. Здесь наиболее полно представлены толщи пород, сложенные на дне Мирового океана 280 миллионов лет назад, и много ископаемых организмов того периода.
- По всему разрезу встречаются конодонты, по которым устанавливают эталонные границы ярусов пермской системы.
- В 12 метрах разреза сконцентрировано более 10 миллионов лет геологической истории Земли, времени, когда происходило постепенное закрытие Уральского океана.
- Разрез содержит большое количество окаменелостей, которые хорошо сохранились в горных породах и доступны для изучения и демонстрации во время полевых экскурсий и в залах музея.

На рисунке 1 представлен геологический разрез Мечетлино.



Рис. 1. Геологический разрез Мечетлино

Для привлечения внимания туристов и повышения информационной осведомленности большого круга пользователей, заинтересованных в изучении разреза Мечетлино, необходимо разработать информационный ресурс (веб), который в интерактивном режиме позволит строить график профиля на основе геологических исследований, определять тип пород по фотографии и стратиграфическую шкалу. Далее рассмотрим данные и процесс их обработки для реализации подобного рода ресурса [5].

3. Обработка данных

Для эффективного анализа геологического разреза требуется системный подход к сбору и организации данных. Для комплексного изучения геологического разреза необходимо интегрировать разнородные данные в единую структурированную систему. Рассмотрим классификацию данных, необходимых для проведения полноценного анализа.

- Пространственные данные формируют географическую основу исследования и включают в себя цифровые модели рельефа (ЦМР), спутниковые снимки или векторные данные.

Необходимо обеспечить интеграцию всех видов пространственных данных, находящихся на территории моделировании разреза, которые могут иметь различное компьютерное представление.

Выделим основные виды пространственной информации, участвующие в моделировании графика геологического разреза.

Информация о скважинах может быть представлена в виде точечных объектах, которые описываются следующим образом

$$P^w = \left\{\left\{\left(x_j^w, y_j^w, z_j^w\right)\right\}, \left\{I_j^{atr}\right\}\right\}, j = \overline{1, n_w} \quad (1)$$

где P^w – точечные объекты, расположенные в пространстве; n_w– количество точечных объектов; I_j^{atr} – атрибутивная информация j-го точечного объекта.

Во время бурения скважины, может быть получена дискретная информация, которая содержит сведения о многослойной структуре (пластах) земной поверхности (рисунок 2), а именно глубине подошвы слоев и координатах их подошвы. Эта информация о многослойных структурах записывается в атрибуты I_j^{atr} каждой скважины.

Скважина N 1 Абсолютная отметка 280,00

Номер слоя	Описание пород	Мощность слоя, м	Глубина подошвы слоя, м	Абс. отметка подошвы слоя, м
1	Техногенные образования представленые насыпными грунтами состоящ из чернозема с включением щебня и дресвы карбонатных пород до 20	1,3	1,3	278,70
2	Глина коричневая, полутвердая, окристая, известковистая, с включением щебня и дресвы карбонатных пород до 20-25%, с гл. 2,2 м суглинок идентичного цвета и состояния	1,5	2,8	277,20
3	Песок гравелистый карбонатного состава, коричневый, с глинистым заполнителем в виде суглинка коричневого, полутвердого, известковистого, окристого до 15-30%, с включением дресвы до 20 %	2,7	5,5	274,50

Рис 2. Данные о скважинах бурения

Для решения задачи построения графика разреза в интерактивном режиме на всей поверхности моделируемой территории необходимо использовать непрерывные данные. Для этих целей могут быть использованы методы интерполяции, которые позволяют создавать поверхности в виде регулярной сетки (2)

$$R = \{R_{cell}\} = \{\{x; y\}_{cell}, h\}, \quad (2)$$

где R_{cell} — неделимые элементы, ячейки растра, которые содержат информацию о высоте подошвы залегания слоя.

Количество таких поверхностей для одной моделируемой территории будет равно количеству геологических пластов i (3).

$$L = \{l\}_r, r = \overline{1, i}\ . \quad (3)$$

- Геологические данные составляют содержательное ядро исследования и включают в себя стратиграфическую информацию или литологические данные.

Благодаря развитию технологий, источником стратиграфической информации теперь служит не только результат анализа информации о скважинах, но и методы искусственного интеллекта (ИИ). Они позволяют быстро и эффективно обрабатывать огромные массивы данных и интерпретировать геофизические данные.

Так, например, ИИ решает задачу компьютерного зрения, а именно — семантической сегментации. Его цель — присвоить каждому пикселю на фотографии геологического разреза метку, соответствующую типу породы или геологическому слою.

Для обучения ИИ понадобятся примерно 500 фотографий геологических разрезов. После обучения на новую, незнакомую фотографию разреза ИИ накладывает свою "маску" — раскрашивает его в условные цвета, показывая, где, по его мнению, находятся разные породы. Но в использование этого метода есть проблемы и ограничения, например:

- Освещение: один и тот же слой может выглядеть по-разному при солнце, в пасмурную погоду или при вспышке.
- Влажность: мокрая порода имеет совершенно другой цвет и контраст по сравнению с сухой.
- Нечеткие границы: Границы между слоями часто бывают переходными, размытыми, а не резкими.
- По одному 2D-изображению невозможно определить многие ключевые свойства: твердость, зернистость, состав минералов и так далее. Опытный геолог всегда использует тактильные и другие методы анализа. ИИ видит только цвет и текстуру.

Однако ИИ будет не заменой геолога, а его мощным инструментом-помощником. Его прогноз будет носить вероятностный и предположительный характер [6]. Идеальный сценарий работы выглядит следующим образом: геолог делает фотографию обнажения, ИИ предлагает предварительную интерпретацию в виде раскрашенной маски, а специалист затем проверяет, уточняет и подтверждает этот результат непосредственно в поле, используя традиционные методы — тактильный анализ, определение твердости и другие инструменты полевой диагностики.

Таким образом, симбиоз искусственного интеллекта и экспертизы геолога открывает новые возможности: ускоряет первичную документацию, стандартизирует описание разрезов и позволяет обрабатывать огромные архивы фотоданных, высвобождая время специалиста для решения более сложных интерпретационных и научных задач.

Для эффективного использования вся информация должна быть организована в единую базу данных с четкой структурой. Такая систематизация позволяет не только проводить комплексный анализ геологического разреза, но и использовать данные для построения трехмерных моделей и прогнозирования свойств пород в межскважинном пространстве.

4. Пример реализации информационного ресурса о геологическом разрезе Мечетлино

Предложенные методы и подходы реализованы при разработке информационного ресурса (веб-сайт) для визуализации геологических данных в цифровой среде на примере геологического разреза Мечетлино. Его интерфейс представлен на рисунке 3.



Рис. 3. Интерфейс веб-сайта

Веб-ресурс предлагает комплексное представление данных о геологическом разрезе Мечетлино – текстовая информация, карта территории, график профиля разреза со стратиграфической шкалой (рисунок 4). Кроме того, подробную информации о каждом слое геологического разреза можно получить непосредственно с карты (нажав на географический объект скважина), либо с графика

профиля (нажав на деление стратиграфической шкалы (рисунок 5)). В графике профиля реализовано включение временной шкалы, которая показывает сколько млн. лет назад образовался каждый слой.

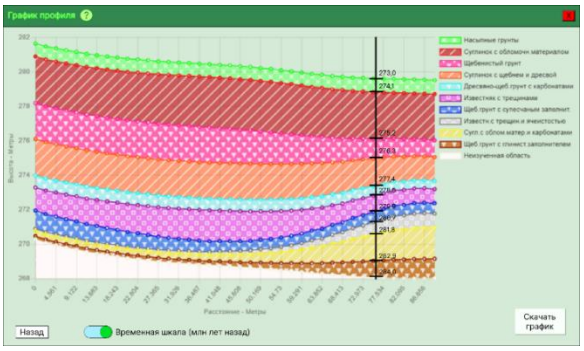


Рис. 4. График профиля геологического разреза Мечетлино

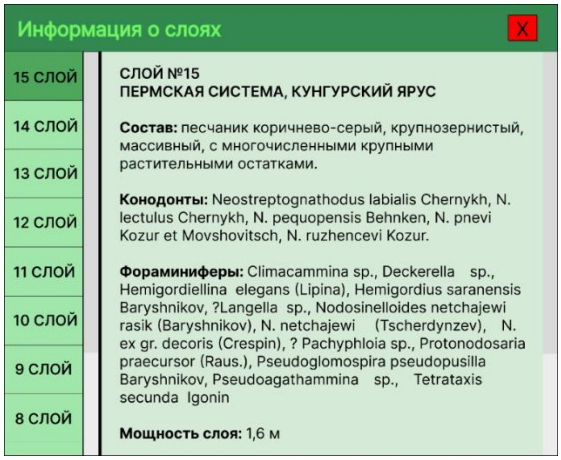


Рис. 5. Диалоговое окно «Информация о слоях»

Разработанный информационный ресурс представляет интерес для различных категорий пользователей (например, туристов или студентов геологических специальностей), поскольку обеспечивает интеграцию разнородных данных о территории и предоставляет их в наглядном (понятном) графическом виде. Подобного рода ресурсы могут выступать в качестве образовательных ресурсов для студентов, исследователей и заинтересованных туристов.

5. Заключение

В результате проведенного исследования были получены следующие результаты:

- Проведен анализ предметной области, который позволил выявить особенности моделируемого объекта.
- Проанализирован процесс организации обработки информации об геологическом разрезе, который показал, что на ряду с классическими методами представления и описания геологических разрезов эффективно использовать современные геоинформационные технологии и искусственный интеллект.
- Реализован информационный ресурс для визуализации геологического разреза на основе проведенных исследований.

Список используемых источников

1. Палеонтология. 2-е изд., перераб. и доп.: Учебник. / И.А. Михайлова, О.Б. Бондаренко. — М.: Изд-во МГУ, 2006. — 592 с. (книга).
2. Насырьянова Р.Б. Представление геологических разрезов в геоинформационных системах // Мавлютовские чтения : Материалы XVII Всероссийской молодёжной научной конференции. В 9-ти томах, Уфа, 21–23 ноября 2023 года. – Уфа: Уфимский университет науки и технологий, 2024. – С. 611-615. (статья из журнала).
3. Стратиграфическая схема верхнемеловых отложений Восточно-Европейской платформы. Объяснительная записка / А.Г. Олферьев, А.С. Алексеев. — М.: Палеонтологический институт РАН, 2005. — 204 с. (книга).
4. Черных В.В., Котляр Г.В., Кутыгин Р.В., Филимонова Т.В., Сунгатуллина Г.М., Мизенс Г.А., Сунгатуллин Р.Х., Исакова Т.Н., Бойко М.С., Иванов А.О., Мычко Э.В. Геологический разрез Мечетлино (Южный Урал). Палеонтологическая характеристика // Геолоический вестник. 2018. № 1. С. 119-137. (статья из журнала).
5. Насырьянова Р. Б. Разработка подсистемы визуализации геологического разреза в ГИС "Геопарки РБ" (на примере разреза Мечетлино): выпускная квалификационная работа — Уфа, 2024. – 80 с. (выпускная квалификационная работа)
6. Искусственный интеллект в геологии: «трудности перевода» и принцип опережения / [Электронный ресурс] // dprom.online. Территория недропользователей : [сайт]. — URL: https://dprom.online/explore/iskusstvennyj-intellekt-v-geologii-trudnosti-perevoda-i-printsip-operezheniya/ (дата обращения: 01.10.2025). (электронный ресурс).

Применение коллективного искусственного интеллекта для инвестиций на рынке золота

Желудков В. Д.
Уфимский университет науки и технологий
Институт информатики, математики и кибернетики
Уфа, Россия
vadim.winter2@mail.ru

Сазонова Е. Ю.
Уфимский университет науки и технологий
Институт информатики, математики и кибернетики
Уфа, Россия
ekaterina_rassadnikova@mail.ru

Сметанина О. Н.
Уфимский университет науки и технологий
Институт информатики, математики и кибернетики
Уфа, Россия
smoljushka@mail.ru

Аннотация[1]

В условиях высокой волатильности финансовых рынков и экономической неопределенности эффективное управление инвестициями на рынке золота приобретает особую актуальность. Работа посвящена разработке и экспериментальной проверка эффективности системы коллективного искусственного интеллекта для принятия инвестиционных решений на рынке золота, объединяющей преимущества различных инвестиционных философий. Разработана многоагентная система, включающая шесть специализированных AI-инвесторов на основе большой языковой модели DeepSeek: консервативного, агрессивного, технического аналитика, фундаментального аналитика, новостного трейдера и DeepSeek AI аналитика. Система продемонстрировала совокупную доходность 1005,3% при средней годовой доходности 14,4%. Коэффициент Шарпа составил 0,52 при максимальной просадке 15,9%. Стратегия коллективного ИИ превзошла пассивное инвестирование в золото в 2,6 раза, стратегию банковских вкладов - в 3,5 раза. Наибольшая эффективность системы наблюдалась в кризисные периоды. Разработанная система представляет собой эффективный инструмент для принятия инвестиционных решений на рынке золота, способный адаптироваться к различным рыночным условиям и обеспечивать устойчивую доходность при контролируемом уровне риска.

Ключевые слова: искусственный интеллект; коллективное принятие решений; инвестиционные стратегии; большие языковые модели; многоагентные системы.

1. Введение

В условиях растущей волатильности финансовых рынков и экономической неопределенности проблема эффективного управления инвестициями приобретает особую актуальность. Рынок золота традиционно считается одним из наиболее сложных для анализа из-за множества факторов, влияющих на ценообразование [1]. Исторически золото выполняло функции защитного актива, однако принятие решений о сроках инвестирования остается сложной задачей.

С развитием технологий искусственного интеллекта, в частности больших языковых моделей (LLM), появились новые возможности для анализа финансовых рынков [2, 3]. Однако существующие подходы часто ограничиваются использованием единого AI-агента, что не учитывает разнообразие инвестиционных стратегий. В отличие от индивидуальных алгоритмических подходов, система коллективного интеллекта способна объединять преимущества различных инвестиционных философий, что и определило цель настоящего исследования.

2. Задача принятия инвестиционных решений на рынке золота

В рамках исследования задача принятия инвестиционных решений на рынке золота формализуется как последовательный процесс принятия решений в условиях неопределенности.

[1] Труды XI Международной научной конференции "Информационные технологии интеллектуальной поддержки принятия решений", 11-13 ноября, Уфа-Чандигарх-Фандербейлпарк, 2025

Рассмотрим формальную постановку задачи. Инвестиционный процесс рассматривается на дискретном временном интервале, где каждый шаг соответствует одному году. Состояние портфеля описывается совокупностью параметров: денежный капитал, количество золота в портфеле, текущая цена золота, процентная ставка по вкладам и исторические данные с новостным фоном.

На каждом временном шаге система выбирает одно из двух возможных действий: инвестировать в золото или разместить средства в банковские вклады. Стратегия системы представляет собой отображение из пространства состояний в пространство действий, определяющее правила принятия решений на основе доступной информации.

Динамика состояния портфеля описывается детерминированными правилами. При выборе инвестирования в золото весь денежный капитал конвертируется в золото с учетом транзакционных издержек. При выборе банковских вкладов золото продается с учетом комиссии, а полученные средства размещаются под процент с начислением сложного процента.

Основной целью системы является максимизация конечной стоимости портфеля при ограничении на риск. Формально задача формулируется как максимизация ожидаемой полезности конечного состояния портфеля. Для учета риска используется модифицированный критерий, максимизирующий отношение ожидаемой доходности к волатильности, что соответствует максимизации коэффициента Шарпа.

В рамках коллективного подхода задача распределяется между несколькими агентами. Каждый агент формирует независимое решение на основе текущего состояния системы. Коллективное решение формируется через функцию агрегации, учитывающую индивидуальные решения агентов и их веса, которые могут определяться на основе уверенности агентов, их исторической эффективности или статических приоритетов.

Формальная модель предполагает ряд ограничений и допущений: отсутствие коротких позиций по активам, фиксированные транзакционные издержки, дискретность времени принятия решений, отсутствие налогов и дополнительных комиссий, а также предположение об идеальной ликвидности активов.

3. Об архитектуре системы коллективного искусственного интеллекта для инвестиций на рынке золота

Разработанная система коллективного искусственного интеллекта представляет собой многоуровневую архитектуру, состоящую из четырех взаимосвязанных модулей, обеспечивающих полный цикл принятия инвестиционных решений - от сбора данных до исполнения и анализа результатов.

Фундаментом системы является модуль данных, ответственный за сбор, хранение и предварительную обработку исторической информации. Ключевые функции модуля включают: использование исторических данных о ценах на золото в рублях за грамм и процентных ставках по вкладам в России за период 2000-2024 годов [6]; интеграцию с открытыми источниками (ЦБ РФ, Росстат) для обеспечения достоверности данных [7]; реализацию механизма валидации и очистки данных, исключающего аномалии и пропущенные значения; подсистему имитации новостного фона, генерирующую релевантные экономические новости для каждого анализируемого года.

Центральным элементом системы выступает модуль анализа, реализующий логику шести специализированных AI-инвесторов на базе большой языковой модели DeepSeek. Архитектура каждого агента включает: промпт-инженерную часть - формирование контекста запроса в соответствии с инвестиционным стилем; механизм взаимодействия с API DeepSeek - обеспечение стабильного соединения с языковой моделью; парсер ответов - извлечение структурированных данных из текстовых ответов модели.

Система объединяет разнонаправленные инвестиционные подходы через шесть типов агентов: консервативный инвестор - ориентация на минимизацию рисков, предпочтение банковских вкладов при ставках выше 7-8%; агрессивный инвестор - поиск возможностей для максимизации доходности, активные инвестиции в золото в периоды роста более 8%; технический аналитик - использование исторических ценовых данных и индикаторов для выявления трендов, и точек разворота; фундаментальный аналитик - сравнительный анализ ожидаемой доходности активов с учетом премии за риск; новостной трейдер - анализ рыночных настроений и реакции на экономические новости; DeepSeek AI аналитик - комплексный анализ, объединяющий все факторы и подходы.

Модуль голосования обеспечивает агрегацию индивидуальных решений в коллективное решение через четыре метода: простое большинство - учет количества голосов за каждый вариант; взвешенное по уверенности - умножение решений на уровень уверенности агента (в процентах); взвешенное по эффективности - дополнительный учет исторической успешности каждого агента; консенсус - требование не менее 75% голосов для принятия решения.

Каждый метод реализован в виде отдельного класса с единым интерфейсом, что обеспечивает простоту добавления новых алгоритмов агрегации. Метод взвешенного по эффективности дополнительно учитывает историческую успешность каждого агента, вычисляемую на основе предыдущих решений, что позволяет системе адаптивно перераспределять влияние агентов в зависимости от их фактической эффективности.

Модуль симуляции реализует среду для тестирования инвестиционных стратегий и включает: механизм выполнения решений, имитирующий реальные инвестиционные операции; учет транзакционных издержек в размере 0,5% при смене активов; отслеживание динамики капитала и сбор статистики для анализа; интерактивный интерфейс визуализации результатов.

Методология исследования предусматривала ретроспективное тестирование на исторических данных за период 2005-2024 годов с начальным капиталом 1 миллион рублей. Критически важным аспектом методологии было обеспечение условий, исключающих использование знаний о будущих событиях - система имела доступ только к историческим данным до текущего года и соответствующему новостному фону.

Алгоритм принятия решений реализовывал итерационный процесс: ежегодный анализ исторических данных каждым агентом; обработка новостной информации и формирование контекста; генерация индивидуальных рекомендаций с подробным обоснованием; коллективное голосование с использованием выбранного метода агрегации; исполнение решения и обновление портфеля

Для объективной оценки эффективности предложенного подхода была разработана система сравнительного анализа с тремя базовыми стратегиями: пассивное инвестирование в золото ("покупай и держи"); стратегия банковских вкладов с начислением сложного процента; смешанная стратегия с фиксированным распределением активов 50/50.

Каждая стратегия тестировалась на идентичных исторических данных с одинаковыми начальными условиями, что обеспечивало сопоставимость результатов и объективность оценки преимуществ коллективного подхода.

Все модули системы были интегрированы в единую платформу, разработанную на языке Python с использованием библиотек PyQt5 для графического интерфейса и matplotlib для визуализации. [5] Архитектура системы предусматривает возможность расширения через добавление новых типов агентов или методов голосования, что делает ее перспективной платформой для дальнейших исследований в области коллективного искусственного интеллекта для финансовых приложений.

4. Анализ эффективности разработанной системы коллективного искусственного интеллекта

Для всестороннего анализа эффективности разработанной системы коллективного искусственного интеллекта был применен комплексный подход, включающий как традиционные финансовые метрики, так и специализированные показатели, адаптированные для анализа AI-систем. Основу методологии оценки составили расчет абсолютной и относительной доходности, анализ рисков, сравнение с бенчмарками и статистическая проверка значимости результатов.

В качестве ключевых метрик доходности использовались абсолютная доходность за весь период, годовая доходность и средняя геометрическая доходность. Абсолютная доходность рассчитывалась как процентное изменение стоимости портфеля с начала до конца инвестиционного периода. Годовая доходность определялась через сложный процент, что позволяло нивелировать эффект различной продолжительности инвестиционных горизонтов. Особое внимание уделялось средней геометрической доходности, которая более точно отражает реальный рост капитала по сравнению со средним арифметическим показателем, особенно в условиях волатильности рынка.

Для оценки рисков применялась многоуровневая система показателей, центральное место в которой занял коэффициент Шарпа, измеряющий избыточную доходность на единицу риска. [4] Расчет коэффициента Шарпа проводился относительно безрисковой ставки, принятой на уровне 5% годовых, что соответствовало средним ставкам по государственным облигациям за анализируемый период. Дополнительно использовался коэффициент Сортино, учитывающий только понижательную волатильность - волатильность отрицательной доходности, что более точно отражает восприятие риска инвесторами. Максимальная просадка (Maximum Drawdown) рассчитывалась как наибольшее падение стоимости портфеля от пика до минимума, что позволяло оценить стрессоустойчивость стратегии в кризисные периоды.

Сравнительный анализ эффективности включал тестирование против трех базовых стратегий: пассивного инвестирования в золото ("покупай и держи"), стратегии банковских вкладов и смешанной стратегии с фиксированным распределением активов 50/50. Каждая стратегия оценивалась по идентичному набору метрик на тех же исторических данных, что обеспечивало

сопоставимость результатов. Для статистической валидации преимуществ коллективного подхода применялся t-тест для парных выборок, позволяющий определить значимость различий в доходности между предлагаемой системой и сравнительными стратегиями.

Анализ устойчивости системы проводился через тестирование на различных подпериодах, выделенных по критерию рыночной фазы: периоды роста (2005-2007, 2010-2012, 2016-2019), кризисные периоды (2008-2009, 2014-2015, 2020, 2022) и боковые тренды (2013, 2021, 2023). Такой подход позволял оценить способность системы адаптироваться к изменяющимся рыночным условиям и выявить потенциальные слабые стороны в специфических рыночных сценариях. Дополнительно анализировалась стабильность работы через расчет коэффициента вариации годовой доходности и показателя последовательности доходности.

Особое внимание уделялось оценке эффективности отдельных компонентов системы. Анализировалась успешность решений каждого типа инвестора через расчет частоту попаданий - доли правильных решений, а также сравнивалась эффективность различных методов голосования. Для оценки добавленной стоимости коллективного подхода рассчитывался показатель диверсификации - снижение волатильности портфеля по сравнению со средневзвешенной волатильностью индивидуальных стратегий агентов. Все расчеты проводились с учетом транзакционных издержек в размере 0,5% при смене активов, что обеспечивало реалистичность оценки практической применимости системы.

Экспериментальное исследование эффективности системы коллективного искусственного интеллекта для инвестиций на рынке золота проводилось на исторических данных за период с 2005 по 2024 год с начальным капиталом 1 000 000 рублей. За двадцатилетний инвестиционный горизонт система продемонстрировала выдающиеся результаты, существенно превзойдя все сравниваемые стратегии-бенчмарки (Рис. 1). Конечная стоимость портфеля составила 11 053 002 рубля, что соответствует совокупной доходности +1005,3% или средней годовой доходности +14,4%.

Анализ риск-профиля системы выявил сбалансированное соотношение доходности и риска (Рис. 5). Волатильность портфеля составила 17,9% при максимальной просадке 15,9%, что свидетельствует об эффективном управлении рисками в периоды рыночной турбулентности. Коэффициент Шарпа, рассчитанный относительно безрисковой ставки 5% годовых, достиг значения 0,52, подтверждая способность системы генерировать избыточную доходность на единицу принимаемого риска. Диапазон годовой доходности варьировался от -15,5% в худшем году до +58,6% в лучшем, при этом система демонстрировала устойчивую положительную динамику на большинстве временных отрезков.

Сравнительный анализ с базовыми инвестиционными стратегиями показал однозначное преимущество коллективного подхода (Рис. 2). Стратегия "покупай и держи" золото показала совокупную доходность +380,0%, стратегия банковских вкладов — +285,2%, а смешанная стратегия 50/50 — +332,6%. Таким образом, система коллективного ИИ превзошла ближайшего конкурента более чем в 2,6 раза по совокупной доходности за двадцатилетний период. Визуальный анализ графиков стоимости портфелей подтвердил, что стратегия команды AI-инвесторов не только обеспечила более высокую конечную стоимость, но и демонстрировала меньшую волатильность по сравнению с пассивным инвестированием в золото.

Распределение инвестиционных решений системы составило 60% в пользу золота и 40% в пользу банковских вкладов (Рис. 3), что отражает адаптивный характер стратегии, способной гибко перераспределять активы в зависимости от рыночных условий. Анализ эффективности системы в различные рыночные фазы выявил ее особую результативность в периоды кризисов (Рис. 4). В 2008-2009 годах, во время глобального финансового кризиса, система последовательно выбирала инвестиции в золото, получив доходность +30,7% и +39,7% соответственно. В 2014-2015 годах, на фоне геополитической напряженности и ослабления рубля, решение в пользу золота принесло +58,6% — максимальную годовую доходность за весь период тестирования. В 2020 году, в условиях пандемии COVID-19, система единогласно выбрала золото, получив +41,4% годовых.

Особый интерес представляет анализ индивидуальной и коллективной эффективности AI-инвесторов (Рис. 7). Каждый из шести типов инвесторов в отдельности показал относительно низкую успешность — 40,0% правильных решений при средней годовой доходности -2,8%. Однако при агрегации решений через механизм коллективного голосования была достигнута синергия, позволившая получить существенно положительный результат. Это подтверждает ключевую гипотезу исследования о том, что разнообразие инвестиционных подходов в рамках коллективной системы позволяет нивелировать индивидуальные ошибки и извлекать преимущество из комплементарности различных аналитических методов.

Эффективность различных методов голосования требует дополнительного исследования, однако предварительные данные указывают на потенциал взвешенных подходов (Рис. 6). В ходе эксперимента использовался метод простого большинства, при котором все агенты имели равные веса. При этом

исторический анализ решений показывает, что технический и фундаментальный аналитики демонстрировали более стабильную эффективность в определении рыночных трендов, в то время как новостной трейдер и DeepSeek AI аналитик оказывались более чувствительными к краткосрочным рыночным колебаниям.

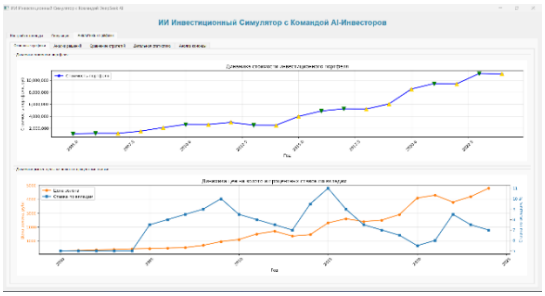


Рис. 1. Динамика стоимости инвестиционного портфеля за период 2005-2024 гг.

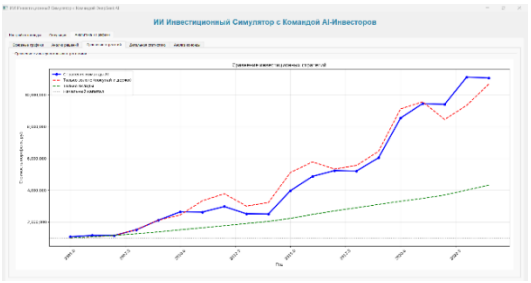


Рис. 2. Сравнение инвестиционных стратегий

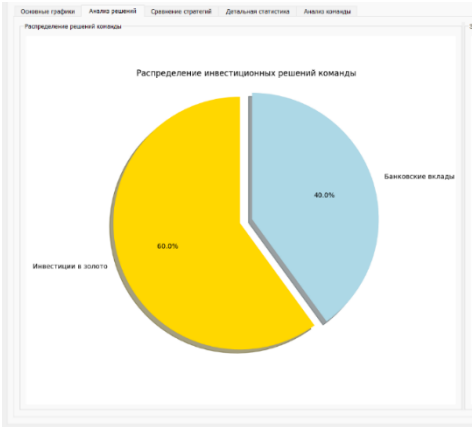


Рис. 3. Распределение инвестиционных решений команды

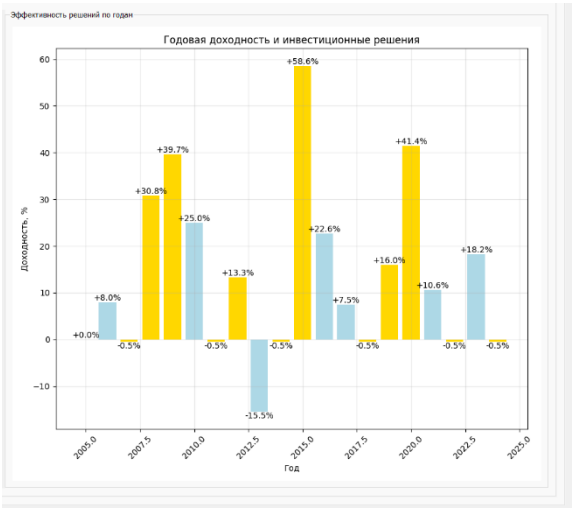


Рис. 4. Годовая доходность и инвестиционные решения

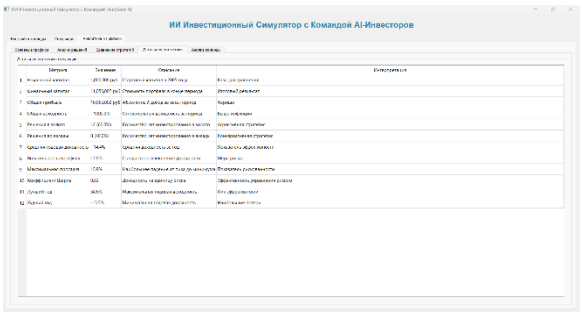


Рис. 5. Детальная статистика симуляции

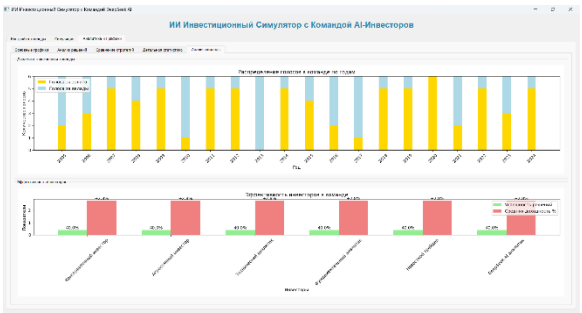


Рис. 6. Распределение голосов в команде по годам

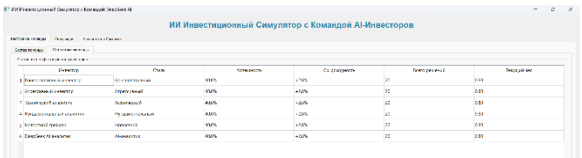


Рис. 7. Статистика эффективности инвесторов

Важным результатом является способность системы избегать значительных потерь в периоды коррекции рынка золота. В 2013 году, когда золото продемонстрировало падение на -20,6%, система единогласно выбрала банковские вклады, сохранив капитал и получив доходность 7,0%. Аналогично, в 2021 году, на фоне восстановления экономики после пандемии и сокращения закупок золота центральными банками, система верно определила временное снижение привлекательности

драгоценного металла, переключившись на банковские вклады.

Верификация результатов на различных подпериодах подтвердила устойчивость системы к изменениям рыночных условий. В периоды роста (2005-2007, 2010-2012, 2016-2019) система демонстрировала стабильную положительную доходность, в кризисные периоды (2008-2009, 2014-2015, 2020, 2022) — превышающую среднюю, а в периоды боковых трендов (2013, 2021, 2023) успешно минимизировала потери. Все расчеты учитывали транзакционные издержки в размере 0,5% при смене активов, что подтверждает практическую применимость системы в реальных рыночных условиях.

5. Заключение

Проведенное исследование подтвердило эффективность применения коллективного искусственного интеллекта для управления инвестициями на рынке золота. Разработанная многоагентная система продемонстрировала способность адаптивно принимать решения в изменяющихся рыночных условиях.

Основные результаты работы:

- Подтверждение гипотезы о преимуществе коллективного подхода - система обеспечила совокупную доходность 1005,3% за 20-летний период, тогда как индивидуальные стратегии AI-инвесторов показали отрицательную среднюю доходность.

- Создание эффективной модульной архитектуры - разработана и апробирована масштабируемая система, включающая модули данных, анализа, голосования и симуляции, допускающая легкое добавление новых агентов и методов агрегации решений.

- Высокая эффективность в кризисные периоды - в условиях финансовых кризисов, геополитической напряженности и пандемии система демонстрировала доходность от 30,7% до 58,6% годовых.

- Эффективное управление рисками - достигнуто ограничение максимальной просадки на уровне 15,9% при коэффициенте Шарпа 0,52, что подтвердило оптимальное соотношение риск-доходность.

- Преимущество над базовыми стратегиями - система превзошла стратегию "покупай и держи" золото в 2,6 раза, стратегию банковских вкладов в 3,5 раза и смешанную стратегию 50/50 в 3,0 раза.

Адаптивность к различным рыночным условиям - система успешно работала в периоды роста, кризисов и боковых трендов, демонстрируя стабильную положительную доходность.

Практическая значимость работы заключается в создании готового инструмента для поддержки инвестиционных решений, который может быть использован частными инвесторами и финансовыми учреждениями для управления портфелями в условиях нестабильности на рынке золота.

Список используемых источников

1. Baur D. G., McDermott T. K. J. Is Gold a Safe Haven? International Evidence // Journal of Banking & Finance. 2010. Vol. 34. № 8. P. 1886-1898.
2. DeepSeek: Official Documentation. URL: https://platform.deepseek.com/api-docs/ (дата обращения: 20.03.2025).
3. Goodfellow I., Bengio Y., Courville A. Deep Learning. Cambridge: MIT Press, 2016. 800 p.
4. Hull J. C. Risk Management and Financial Institutions. Hoboken: Wiley, 2018. 624 p.
5. Маккин У. Python и анализ данных. СПб.: Питер, 2022. 482 с.
6. Федеральная служба государственной статистики (Росстат). URL: https://rosstat.gov.ru (дата обращения: 20.03.2025).
7. Центральный банк Российской Федерации (Банк России). URL: https://cbr.ru (дата обращения: 20.03.2025).

Оценка применимости квантовых генераторов случайных чисел в системах защиты информации

Кучкарова Н. В.
Уфимский университет науки и технологий
Институт информатики, математики
и робототехники
Уфа, Россия
e-mail: nailya_kuchkarov@mail.ru

Ерошкин А. Д.
Уфимский университет науки и технологий
Институт информатики, математики
и робототехники
Уфа, Россия
e-mail: aderoshkin2005@gmail.com

Аннотация[1]

В данной работе представлена классификация квантовых генераторов случайных чисел (КГСЧ) с точки зрения использования в области информационной безопасности. Классификация проведена по трем критериям: физической природе явления, уровню доверия и архитектуре обработки данных. На ее основе выделены перспективные типы КГСЧ для систем защиты информации. Практическим результатом работы является разработка алгоритма прототипа генератора паролей, использующий суперпозицию и измерение кубитов для обеспечения криптографически стойкой случайности.

Ключевые слова: генератор псевдослучайных чисел (ГПСЧ); квантовая случайность; средства защиты информации; квантовый генератор случайных чисел (КГСЧ).

1. Введение

Массовая цифровизация повышает уязвимость критически важных объектов, включая государственные информационные системы, инфраструктуру промышленных предприятий и финансовый сектор. Противодействие кибератакам, нацеленным на кражу, искажение или перехват конфиденциальных данных, требует применения надежных механизмов защиты. Базовым подходом в обеспечении информационной безопасности в таких условиях остается использование криптографических методов [1].

На протяжении десятилетий, начиная с 1950-х годов, в криптографии доминируют генераторы псевдослучайных чисел (ГПСЧ). Однако детерминированная природа этих алгоритмов обуславливает ряд фундаментальных недостатков, ограничивающих их применение в современных системах безопасности [1-3].

К фундаментальным недостаткам традиционных генераторов псевдослучайных чисел (ГПСЧ) относятся:

- Предсказуемость последовательностей**.** Детерминированная природа ГПСЧ создаёт риск восстановления начального состояния (сида), что приводит к компрометации всей выходной последовательности.

- Цикличность. Из-за ограниченности внутренних ресурсов (например, размера состояния) последовательности ГПСЧ имеют конечный период, после которого повторяются.

- Наличие алгоритмической зависимости. Каждое последующее число в последовательности вычисляется на основе предыдущих, что создаёт уязвимые статистические взаимосвязи.

С появлением квантовых компьютеров эти системные уязвимости ГПСЧ становятся критическими. Злоумышленники, используя квантовые алгоритмы, могут эксплуатировать эти недостатки с беспрецедентной эффективностью, что ставит под угрозу безопасность криптографических систем, основанных на традиционной генерации случайных чисел [14]. В 2018 году Национальный институт стандартов и технологий США (NIST) представил генератор случайных чисел, использующий феномен квантовой запутанности. Согласно проведенным исследованиям], квантовые генераторы случайных чисел (КГСЧ) способны генерировать последовательности с истинной, или фундаментальной, непредсказуемостью. Это ключевое свойство позволяет устранить присущие традиционным ГПСЧ недостатки, обеспечивая высокий уровень энтропии и безопасности. Как следствие, КГСЧ открывают новые возможности для применения в задачах, требующих гарантированной случайности, – в криптографии, научном моделировании, криптовалютных протоколах и других областях [22].

В разделе 2 данной публикации будет представлена классификация КГСЧ в контексте применимости их

[1] Труды XI Международной научной конференции "Информационные технологии интеллектуальной поддержки принятия решений", 11–13 ноября, Уфа-Чандигарх-Фандербейлпарк, 2025

в средствах защиты информации. Раздел 3 содержит алгоритм предлагаемого прототипа генератора случайных паролей.

2. Классификация квантовых генераторов случайных чисел

Предлагаемая классификация КГСЧ включает три взаимодополняющих критерия, позволяющих оценить устройства с точки зрения их применимости в системах защиты информации.

2.1. По физической природе квантового явления

Оптические системы

- Деление фотонного пучка

Принцип работы такого генератора основан на квантовой суперпозиции одиночного фотона. Фотон направляется на полупрозрачное зеркало (светоделитель), которое с вероятностью 50% либо отражает, либо пропускает его. С квантово-механической точки зрения, фотон до взаимодействия с делителем находится в суперпозиции обоих путей, а акт измерения (факт срабатывания детектора в отраженном или прошедшем канале) коллапсирует это состояние в один из двух классических исходов: «отражение» (состояние $|0\rangle$) или «прохождение» (состояние $|1\rangle$). Эта принципиальная непредсказуемость отдельного события и служит источником истинной случайности. Данные системы отличаются высокой надежностью, но чувствительны к внешним условиям [16,17,18].

- Квантовая флуктуация вакуума

Данный метод основан на измерении фундаментальных квантовых флуктуаций электромагнитного поля в вакууме. Даже в идеальном вакууме, при отсутствии фотонов, напряжённости электрического и магнитного полей не равны нулю, а испытывают случайные колебания, предсказанные квантовой теорией поля. Преимущества такого подхода – высокая криптографическая стойкость и высокая скорость генерации. Недостатки – чувствительность к изменениям температуры и механическим вибрациям [19,24,25].

- Спонтанное параметрическое рассеяние (SPDC)

Мощный лазерный луч (луч накачки) с высокой энергией фотонов направляется на специальный нелинейный оптический кристалл (чаще всего используются кристаллы BBO, KTP или PPLN). С определённой (очень малой, порядка 10^{-10}) вероятностью один фотон накачки может спонтанно, под влиянием флуктуаций вакуума, распасться внутри кристалла на два фотона с меньшей энергией (и, следовательно, большей длиной волны).

При прохождении фотона лазерной накачки через нелинейный кристалл с малой вероятностью происходит его спонтанное расщепление на пару коррелированных фотонов (сигнальный и холостой). Случайность момента и направления распада служит источником энтропии. Особенности: высокая гибкость в выборе параметров для генерации битов, но требуется сложная оптическая стабилизация, а также имеется чувствительность к температуре кристалла и нестабильности накачки. Ввиду сложности и дороговизны данные системы являются менее пригодными для массовых коммерческих применений по сравнению с более простыми полупроводниковыми решениями [4,19,13].

Полупроводниковые системы.

- Лавовый (лавинный) шум

Этот тип генераторов относится к твердотельным (полупроводниковым) системам и использует квантовую случайность, возникающую в процессе лавинного пробоя обратносмещённого p-n-перехода, обычно в лавинном фотодиоде (APD), работающем в так называемом гегеровском режиме. Данные системы отличаются простотой реализации, высокой скоростью, но требуют фильтрации теплового шума [7,10].

- Туннелирование одиночных электронов

основано на квантовомеханическом прохождении электрона через тонкий потенциальный барьер (туннельный переход) в наноструктуре. Момент такого туннелирования принципиально непредсказуем и служит источником истинной случайности. Обеспечивает высокую стабильность процесса, обладает компактностью и технологической совместимостью с современной интегральной электроникой, но для эффективного выделения квантового сигнала на фоне теплового шума, как правило, требуется либо использование криогенных температур, либо создание сложных наноструктур.

Атомные системы представляют собой перспективный класс квантовых генераторов случайных чисел, использующих непредсказуемость квантовых переходов. Их работа основана на случайности спонтанного излучения, квантовых скачков между уровнями или измерений состояний отдельных атомов/ионов. Эту группу условно делят на две подкатегории:

- Спиновые системы

Частица с неопределённой ориентацией спина помещается в магнитное поле. При измерении в выбранном базисе её состояние коллапсирует в одно из двух: $|0\rangle$ («спин вверх») или $|1\rangle$ («спин вниз»). Вероятность каждого исхода определяется исходной суперпозицией, что обеспечивает фундаментальную случайность результата [8,9].

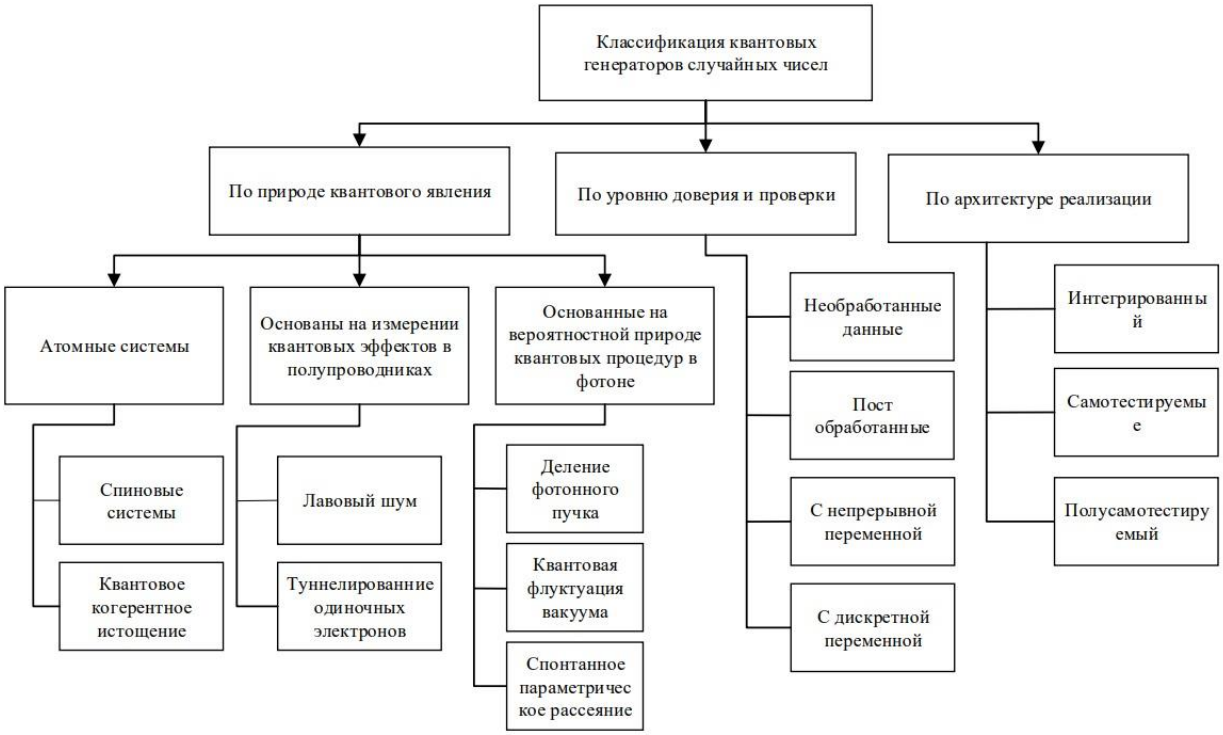


Рис. 1. Классификация квантовых генераторов случайных чисел

- Квантовое когерентное истощение

в когерентной системе слабое возбуждение может с низкой вероятностью вызвать случайный квантовый переход, нарушающий когерентность состояния. Непредсказуемость момента и канала этого перехода служит источником энтропии. Данная система отличается высоким потенциалом скорости и масштабируемости [13,19].

2.2. По уровню доверия и проверки.

Этот аспект критически важен для качества и скорости конечного случайного числа.

- Необработанные данные: «сырой» сигнал с детектора. отражает физический отклик на случайное квантовое событие. Этот сигнал содержит истинную квантовую случайность, но также может включать технические и классические шумы, что требует последующей обработки [6,11,22].

- Пост-обработанные данные представляют собой результат обработки «сырого» сигнала КГСЧ с помощью алгоритмов выравнивания, сжатия или экстракции случайности. Цель постобработки – удалить остаточные корреляции, снизить влияние технических шумов и гарантировать статистическую близость выходной последовательности к идеальному равномерному распределению [20,21].

По типу измеряемой величины:

- Непрерывные генераторы основаны на измерении квадратур электромагнитного поля или других аналоговых квантовых наблюдаемых, принимающих значения в непрерывном диапазоне. Источником случайности служат квантовые флуктуации вакуума или шум сжатого света. Данные генераторы позволяют достичь очень высоких скоростей генерации [11,22].

- Дискретные генераторы используют квантовые процессы с конечным числом исходов, такие как прохождение/отражение фотона, состояние спина электрона или факт детектирования пары фотонов при SPDC. Каждое измерение даёт дискретный, обычно бинарный результат ($|0\rangle$ или $|1\rangle$), который напрямую интерпретируется как случайный бит. Отличается простой интерпретацией в виде битов [17,18].

2.3. По архитектуре реализации

По архитектуре реализации КГСЧ делятся на три основные группы: интегрированные, самотестируемые и полусамотестируемые.

- Интегрированные генераторы реализуются в виде компактных микросхем – фотонных или полупроводниковых, где источник квантовой случайности, оптические/электронные элементы и детекторы размещены на одном кристалле – кремниевой платформе (Si photonics, CMOS) [5,12,15].

- Самотестируемые генераторы Основаны на нарушении неравенств Белла с использованием запутанных частиц. Обеспечивают максимальный уровень доверия без знания внутреннего устройства генератора [1,3,14].

- Полусамотестируемые генераторы предполагают частичные допущения о системе (например,

размерность гильбертова пространства). Данная группа КГСЧ предлагает оптимальный компромисс между криптографической надёжностью и практичностью реализации.

Проведенная классификация демонстрирует разнообразие современных КГСЧ, которые различаются физическими принципами и эксплуатационными параметрами. Общим для всех подходов является использование фундаментальных квантовых процессов (туннелирования, спонтанного излучения, вакуумных флуктуаций), что гарантирует высокий уровень истинной случайности.

Для массового внедрения наиболее перспективны полупроводниковые технологии (лавинный шум, туннелирование) и интегрированные CMOS/Si-фотонные платформы. Эти решения сочетают высокую скорость генерации, совместимость с существующей электроникой и приемлемую устойчивость. В то же время оптические и атомные системы, обеспечивая максимальную криптографическую надежность, требуют сложного оборудования и более чувствительны к внешним условиям.

Выбор архитектуры КГСЧ всегда представляет собой компромисс между скоростью, надежностью, стоимостью и требованиями к сертификации. Для задач информационной безопасности с повышенными требованиями к доказуемой случайности предпочтительны системы с дискретным бинарным измерением.

В третьем разделе представлен алгоритм прототипа квантового генератора.

3. Алгоритм работы прототипа квантового генератора

На основе анализа существующих методов и недостатков классических ГПСЧ в работе разработан прототип квантового генератора паролей. Его принципиальное отличие от псевдослучайных алгоритмов и аппаратных генераторов на классическом шуме заключается в использовании фундаментальной квантовой случайности, что обеспечивает криптографически стойкую основу. Ниже представлен алгоритм работы прототипа квантового генератора, состоящего из 3-х этапов (рис.2):

- Генерация отдельных квантовых битов;
- Преобразование кубитов в числа заданного диапазона;
- Построение символьного пароля на основе полученных чисел.

Разработанный алгоритм генератора паролей основан на фундаментальном принципе квантовой механики – непредсказуемости результата измерения кубита в состоянии суперпозиции. В отличие от детерминированных или псевдослучайных классических методов, этот подход обеспечивает истинную случайность для каждого бита. Ниже представлено пошаговое описание алгоритма: от инициализации квантового состояния и измерения до преобразования битовой последовательности в пароль и обработки параметров генерации.

- Генерация отдельных квантовых битов

Данный этап состоит из подэтапов 1–4:

Инициализация кубита (1). Создается один кубит – физический носитель квантовой информации. Он изначально устанавливается в базовое (детерминированное) состояние $|0\rangle$, что обеспечивает четкую отправную точку для последующих операций.

Применение оператора Адамара H (2). К кубиту применяется унитарный оператор Адамара, переводящий его из состояния $|0\rangle$ в суперпозицию. Это создаёт равновероятную суперпозицию двух базовых состояний, лежащую в основе квантовой неопределённости.

Измерение кубита (3). Производится измерение кубита в вычислительном базисе, в результате чего его суперпозиция коллапсирует в одно из классических состояний: 0 или 1. Каждый исход имеет вероятность 50%. На этом этапе квантовая неопределённость преобразуется в классическую случайность.

Сброс квантового состояния (4). После измерения кубит сбрасывается в исходное состояние $|0\rangle$. Это гарантирует возможность его повторного использования в следующих циклах генерации без корреляции с предыдущими результатами.

- Преобразование кубитов в числа заданного диапазона.

Данный этап состоит из подэтапов 5–7:

Определение минимального числа битов (5) для представления верхней границы диапазона (max). Определяется минимальное количество бит, достаточное для представления любого числа в заданном диапазоне [0, max]. Это обеспечивает полное покрытие всех возможных значений без потери данных.

Цикл измерений (6). На каждом шаге цикла генерируется независимый квантовый бит. Это эквивалентно выборке из идеального распределения Бернулли, где каждый бит статистически независим и имеет равную вероятность быть 0 или 1. Далее выполняется проверка условия (7): если сгенерированное число превышает максимальное значение max, оно отбрасывается, и функция запускается рекурсивно заново. В противном случае – число считается валидным и возвращается как результат. Если полученное число удовлетворяет условию и не превышает максимальную границу (число $\leq$ max), оно признаётся валидным результатом и немедленно возвращается в качестве выходного значения функции. Этот успешно сгенерированный

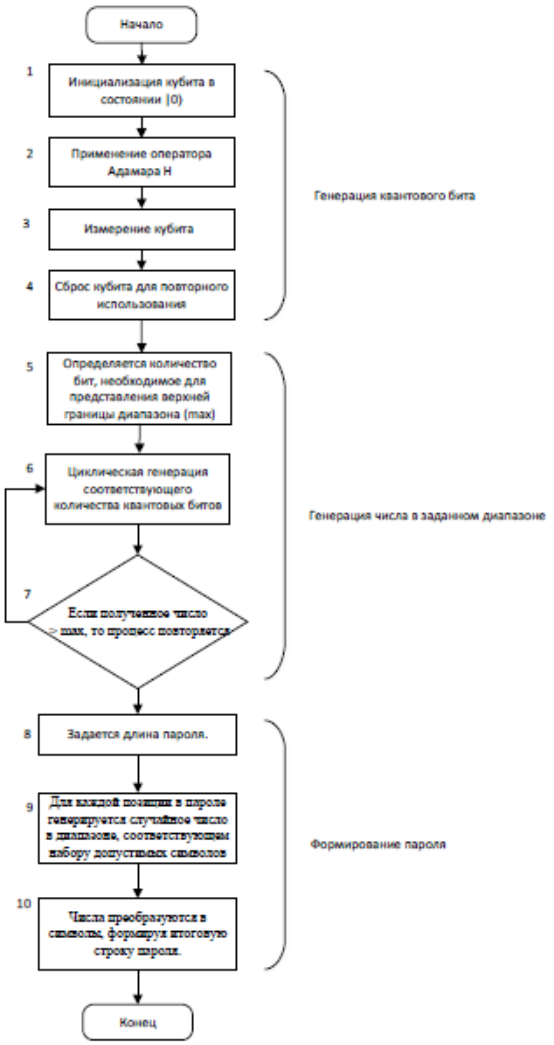


Рис. 2. Алгоритм работы прототипа квантового генератора

элемент передаётся далее в алгоритм для формирования символа пароля, становясь частью криптографически стойкой последовательности. В противоположном случае, когда сгенерированное число выходит за установленные рамки (число > max), алгоритм инициирует рекурсивный вызов самой себя. Этот процесс повторяется циклически до тех пор, пока не будет получено значение, укладывающееся в требуемый диапазон [0, max].

Данный шаг является ключевым для реализации метода отбора (rejection sampling). Его фундаментальная цель – обеспечить строго равномерное распределение вероятностей для всех допустимых чисел. Без этого механизма числа чьи двоичные представления попадают в нижнюю часть возможного битового пространства, имели бы более высокую вероятность выбора, что привело бы к статистическому смещению. Рекурсивная «перегенерация» некорректных результатов эффективно устраняет это смещение. Таким образом, каждый валидный вариант – от 0 до max – имеет абсолютно идентичную, математически выверенную вероятность 1/(max+1) быть сгенерированным. Это гарантирует криптографическую чистоту распределения, что критически важно для любого генератора, претендующего на истинную случайность.

- Формирование пароля

Работа третьего этапа состоит из подэтапов 8–10:

Задается длина пароля (8). Определяется длина итогового пароля. Этот параметр может быть либо фиксированным значением, либо задаваться пользователем. Длина пароля является ключевым фактором, напрямую определяющим его энтропию и, следовательно, криптографическую стойкость.

Далее для каждого символа пароля в отдельности вызывается квантовый генератор чисел (9). Это гарантирует статистическую независимость и истинную случайность каждого элемента итоговой последовательности. Генератор создает число, соответствующее позиции допустимого символа в выбранном алфавите (например, в таблице ASCII или заданном наборе знаков).

Затем каждое число в массиве преобразуется в соответствующий символ (10), как правило, с использованием таблицы ASCII или другого предопределённого набора знаков. Эта операция осуществляет отображение равномерного дискретного распределения чисел в пространство допустимых символов, формируя финальную строку пароля.

4. Заключение

- В представленной работе проведена классификация квантовых генераторов случайных чисел (КГСЧ) по следующим ключевым критериям: физической природе квантового явления, типу измеряемой переменной (дискретная/непрерывная), уровню доверия и методам верификации (сырые/постобработанные, самотестируемые/полусамотестируемые), а также архитектуре реализации (интегрированные, оптические, полупроводниковые и др.).

- На основе проведенного анализа современных методов генерации случайности и выявленных системных недостатков традиционных генераторов псевдослучайных чисел (ГПСЧ) в данной работе был разработан прототип квантового генератора паролей. Ключевое отличие предложенного решения от классических псевдослучайных алгоритмов и аппаратных генераторов, основанных на детерминированных процессах или измерениях классического шума, заключается в использовании фундаментального принципа квантовой механики – истинной, непредсказуемой случайности, возникающей при измерении квантовой системы. Это обеспечивает принципиально более высокий и доказуемый уровень криптографической стойкости, лежащей в основе генерации паролей.

Список используемых источников

1. Microsoft: генератор случайных чисел в Windows XP содержит недостатки // – URL:

https://www.securitylab.ru/news/308676.php (дата обращения: 05.09.2025).

2. Слеповичев И. И. Генераторы псевдослучайных чисел //Саратов: СГУ. – 2017. – Т. 117.
3. Случайные числа своими руками №1 // – URL: https://xakep.ru/2005/01/19/25259/ (дата обращения: 04.09.2025).
4. Неслучайная случайность, или Атака на ГПСЧ в .NET 1. // – URL: https://habr.com/ru/companies/skbkontur/articles/347758/ (дата обращения: 04.09.2025).
5. Kelsey J. et al. Cryptanalytic Attacks on Pseudorandom Number Generators // LNCS,Vol. № 1372. P. 168-188 (статья из журнала).
6. Marcin M. Jacak et al Quantum generators of random numbers. // – URL: https://www.nature.com/articles/s41598-021-95388-7 / (дата обращения: 04.09.2025).
7. Физики создали самый надежный генератор случайных чисел. // – URL: https://hi-tech.mail.ru/news/129031-kvantovye-fiziki-predstavili-samyj-nadezhnyj-generator-sluchajnyh-chisel/ / (дата обращения: 06.09.2025).
8. Жизан Н. Квантовая случайность: Нелокальность, телепортация и другие квантовые чудеса. – Альпина нон-фикшн, 2023.
9. Файер М Абсолютный минимум. – Издательский дом «Питер», 2016.
10. Методы экстракции квантовой случайности // – URL: http://qutes.org/wp-content/uploads/2023/03/QTS_2023.pdf (дата обращения: 07.09.2025).
11. Mannalath V., Mishra S., Pathak A. A comprehensive review of quantum random number generators: Concepts, classification and the origin of randomness //arXiv preprint arXiv:2203.00261. – 2022..
12. Ma X. et al. Quantum random number generation //npj Quantum Information. – 2016. – Т. 2. – №. 1. – С. 1-9.
13. Herrero-Collantes M., Garcia-Escartin J. C. Quantum random number generators //Reviews of Modern Physics. – 2017. – Т. 89. – №. 1. – С. 015004.
14. Sönmez Turan M. et al. Recommendation for the entropy sources used for random bit generation. – National Institute of Standards and Technology, 2016. – №. NIST Special Publication (SP) 800-90B (Draft)..
15. Jofre M. et al. True random numbers from amplified quantum vacuum //Optics express. – 2011. – Т. 19. – №. 21. – С. 20665-20672..
16. Kaiser, F., et al. (2020). Integrated photonic quantum random number generator. APL Photonics. – 2020. // – URL: https://doi.org/10.1063/5.0022362 (дата обращения: 05.09.2025)
17. Acín A., Masanes L. Certified randomness in quantum physics //Nature. – 2016. – Т. 540. – №. 7632. – С. 213-219..
18. Dynes J. F. et al. A high speed, postprocessing free, quantum random number generator //applied physics letters. – 2008. – Т. 93. – №. 3..
19. Wang, Y., et al.. Spin-based quantum random number generation using NV centers in diamond. Quantum Science and Technology. – 2022. // – URL: https://doi.org/10.1088/2058-9565/ac6d9e (дата обращения: 05.09.2025).
20. NIST Special Publication 800-90B. Recommendation for the Entropy Sources Used for Random Bit Generation. – U.S. Department of Commerce, 2018. – URL: https://csrc.nist.gov/publications/detail/sp/800-90b/final (дата обращения: 09.09.2025).
21. Dynes, J.F. et al. A high speed, postprocessing free, quantum random number generator // Applied Physics Letters. – 2008. – Vol. 93, № 3. – P. 031109. – URL: https://doi.org/10.1063/1.2961000 (дата обращения: 07.09.2025).
22. Acín, A., Masanes, L. Certified randomness in quantum physics // Nature. – 2016. – Vol. 540. – P. 213–219. – URL: https://doi.org/10.1038/nature20119 (дата обращения: 06.09.2025).

Модели и алгоритмы поддержки принятия решений при управлении геологоразведочными работами на основе методов тензорного анализа

Орлов Г. О.
Уфимский университет науки и технологий
Институт информатики, математики и робототехники
Уфа, Россия
e-mail: orlovgleb99@mail.ru

Аннотация[1]

Статья посвящена разработке моделей и алгоритмов поддержки принятия решений для управления геологоразведочными работами на основе методов тензорного анализа. Предложены концепция обработки данных, мультиосевая модель организации и метод визуализации геофизических полей. Разработанные решения обеспечивают сохранение структурной целостности многокомпонентных данных на всех этапах обработки. Экспериментально подтверждено преимущество предложенных методов над традиционными подходами по точности интерпретации и эффективности обработки данных.

Ключевые слова: геопространственные данные; геомагнитные данные; поддержка принятия решений; системы управления ГРР.

1. Введение

Современные геологоразведочные работы характеризуются нарастающим противоречием между сложной природой геофизических данных и применяемыми методами их обработки. Геофизические поля, включая геомагнитные, гравитационные и электромагнитные, играют ключевую роль в геологоразведке, обеспечивая дистанционное изучение геологического строения территорий и выявление месторождений полезных ископаемых [1, 2]. Однако традиционные подходы, основанные на скалярных и векторных представлениях, не учитывают тензорную организацию геофизических полей, что приводит к потере структурных взаимосвязей и снижению эффективности управленческих решений [3, 4].

Анализ существующих систем поддержки принятия решений показывает их ограниченную применимость для работы с многокомпонентными геофизическими данными [5, 6]. В научной литературе представлены исследования по многокомпонентной визуализации [7, 8], модернизации моделей хранения данных [9, 10], веб-ориентированным решениям обработки [11, 12] и поддержке принятия решений на основе разнородных данных в прикладных областях [13, 14]. Однако преобладание покомпонентного анализа и отсутствие целостного представления информации создают предпосылки для субъективных ошибок интерпретации и увеличивают временные затраты на подготовку данных [15, 16].

Целью исследования является разработка моделей и алгоритмов поддержки принятия решений на основе методов тензорного анализа. Практическая значимость работы заключается в создании комплекса решений, обеспечивающих сохранение структурной целостности геофизических данных на всех этапах обработки. Предлагаемый подход позволяет преодолеть системное противоречие между многомерной природой геофизических объектов и традиционными методами их анализа, что особенно актуально в условиях освоения месторождений со сложной геологической структурой.

Разрабатываемые в исследовании методы тензорного анализа направлены на создание целостной системы поддержки принятия решений, где геофизические данные формализуются как единый тензорный объект, сохраняющий пространственно-временные взаимосвязи между параметрами на всех этапах обработки - от

[1] Труды XI Международной научной конференции "Информационные технологии интеллектуальной поддержки принятия решений", 11-13 ноября, Уфа-Чандигарх-Фандербейлпарк, 2025

первичных измерений до формирования управленческих показателей.

2. Концепция управления геологоразведочными работами на основе тензорного анализа

Предлагаемая концепция представляет собой системный подход к управлению геологоразведочными работами, основанный на применении методов тензорного анализа для целостного представления многокомпонентных геофизических данных. Ключевой идеей концепции является формализация геофизических полей как единых тензорных объектов, сохраняющих структурные взаимосвязи между параметрами на всех этапах обработки информации.

Методологическая основа концепции включает системную интеграцию тензорных методов через формализацию многокомпонентных данных как единого тензорного объекта и разработку адаптивных механизмов интерпретации с сохранением пространственно-временных корреляций. Ключевым аспектом является сохранение структурной целостности данных при их преобразовании в управленческие показатели. В отличие от традиционных методов обработки, основанных на скалярных и векторных представлениях, предлагаемая тензорная модель учитывает комплексные пространственно-временные взаимосвязи и анизотропию физических свойств горных пород.

Концепция реализует трехуровневую архитектуру управления, включающую сбор и хранение многомерных данных, их тензорный анализ и интерпретацию, формирование управленческих воздействий (Рис. 1). Особенностью подхода является сквозная обработка информации без декомпозиции исходного тензорного представления, что минимизирует потери значимых корреляций между параметрами геофизических полей.

Математической основой концепции выступает адаптивный метрический тензор, вычисляемый по формуле

$$g_{ij} = \delta_{ij} + \alpha \cdot Cov(B_i, B_j),$$

где δ_{ij} – дельта Кронекера, α – коэффициент адаптации, Cov – ковариационная матрица компонент поля. Использование адаптивного метрического тензора позволяет учитывать пространственно-временные зависимости и анизотропию физических свойств геологической среды (Рис. 2).

Центральное место в концепции занимает операция метрического преобразования данных, включающая три ключевых элемента: исходный тензор поля B, метрический тензор g и его транспонированную форму g^T. Преобразование выполняется по формуле $\tilde{B} = g \cdot B \cdot g^T$, где адаптивный метрический тензор учитывает свойства геологической среды, а двойное умножение сохраняет структурные связи. Результатом преобразования являются скорректированные значения, в которых учтена анизотропия пород, подчеркнуты значимые аномалии и сохранены топологические свойства, что создает основу для последующего анализа границ и классификации геологических объектов.

Концепция включает механизмы спектрального анализа тензорных полей и вычисления инвариантов, позволяющих количественно оценивать неоднородность геофизических характеристик и выделять перспективные зоны для детального изучения. Особое внимание уделяется разработке специализированных методов визуализации, обеспечивающих наглядное представление сложных пространственных зависимостей для лиц, принимающих решения.

Реализация концепции позволяет преодолеть системное противоречие между многомерной природой геофизических объектов и традиционными методами их обработки, обеспечивая повышение обоснованности управленческих решений за счет сохранения системных взаимосвязей на всех этапах производственного цикла геологоразведочных работ.

3. Мультиосевая модель организации тензорных данных геофизических полей

В контексте управления геологоразведочными работами разработана мультиосевая модель организации данных, обеспечивающая сохранение структурной целостности информации о геофизических полях на всех этапах технологического процесса [2]. Модель реализует принцип тензорного метрического соответствия через систему взаимосвязанных компонентов, включающую статическую ось метаданных, динамическую ось оперативных измерений и архивную ось исторических данных [5]. Данный подход позволяет эффективно решать ключевые задачи геологоразведки, обеспечивая непротиворечивое и полное представление многокомпонентных данных для принятия управленческих решений.

Центральным элементом архитектуры выступает источник данных, объединяющий реальные

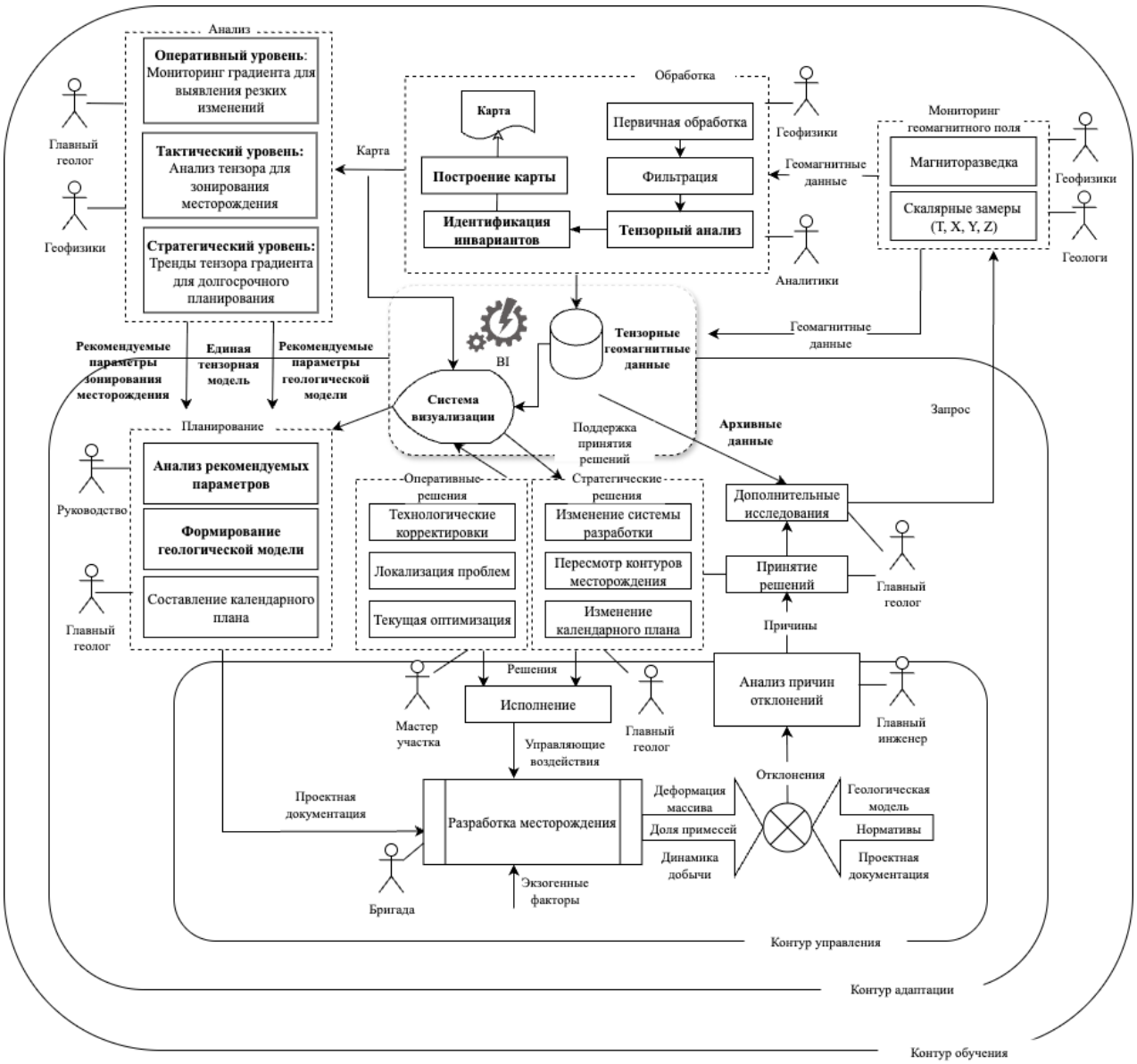


Рис. 1. Структурно-функциональная схема организационной системы управления геологоразведочными работами

измерительные станции и их цифровые аналоги [7]. Применительно к геологоразведочным работам, это позволяет осуществлять непрерывный сбор и структурирование данных о геофизических полях в процессе бурения, сейсмических исследований и других видов деятельности. Ключевой особенностью модели является автоматическое сохранение пространственно-временных взаимосвязей между параметрами при их цифровом представлении. Для геомагнитных измерений это обеспечивает согласованное хранение компонент *X*, *Y*, *Z* и их производных с привязкой к единой временной метке, что принципиально важно для достоверности геологической интерпретации [8]. Модель поддерживает работу с разнородными данными, обеспечивая целостность их тензорного представления при комплексной оценке перспективности месторождений.

Особое значение в модели имеет механизм взаимодействия между динамической оперативной осью и статичной архивной осью [1-4]. Специальная триггерная система реализует автоматизированный процесс миграции массивов данных по истечении заданного временного интервала ΔT, сохраняя при этом все структурные взаимосвязи между параметрами в их тензорном представлении. Механизм учитывает временные характеристики информации и уникальные идентификаторы узлов, исключая возможность дублирования или потери данных [3, 9]. Система обеспечивает целостность временных рядов информации о полях и позволяет одновременно работать с актуальными

и историческими данными, что особенно важно при сравнительном анализе.

Принципиальным отличием разработанной модели является ее способность сохранять метрические свойства тензорного представления данных [11]. Это достигается за счет специальных операторов агрегации, учитывающих не только значения отдельных параметров, но и их структурные взаимосвязи, выраженные через ковариационные матрицы и другие тензорные характеристики [12]. Такой подход позволяет избежать типичных проблем, связанных с потерей пространственно-временных корреляций в цифровом представлении полей [2].

Модель обеспечивает методологическую согласованность между этапами сбора, хранения и обработки геофизической информации [1]. Архитектурные решения включают механизмы автоматического контроля целостности массивов данных, верификации поступающей информации и алгоритмы согласования разнородных данных [2]. Это позволяет существенно сократить временные затраты на подготовку данных для анализа и повысить достоверность интерпретационных результатов.

Разработанная модель создает надежную основу для реализации адаптивных алгоритмов анализа, позволяя учитывать пространственно-временную динамику в цифровом представлении геофизических полей при принятии управленческих решений [6]. Сохраняя тензорные свойства данных на всех этапах их жизненного цикла, модель преодолевает традиционный разрыв между сложной природой геофизической информации и упрощенными методами ее хранения [7]. Внедрение модели в практику открывает новые возможности для повышения эффективности управления геологоразведочными процессами за счет обеспечения целостного цифрового представления многокомпонентных данных о геофизических полях.

4. Метод тензорной визуализации геофизических данных

Для практической реализации концепции тензорного управления геологоразведочными работами разработан специализированный метод визуализации, обеспечивающий целостное представление многокомпонентных геофизических полей [9]. Метод основан на использовании адаптивных суперэллипсоидных глифов (Рис. 2), которые преобразуют сложные тензорные данные в интуитивно понятные визуальные объекты, сохраняющие структурные взаимосвязи между параметрами поля.

В контексте геологоразведочных работ метод решает критически важную задачу – предоставление специалистам комплексного визуального представления геологической среды. Процесс визуализации начинается с преобразования исходного тензора поля и адаптивного метрического тензора, что позволяет учесть пространственные корреляции между компонентами [10]. По полученным значениям строятся оси глифа, где их длина и ориентация отражают ключевые свойства геофизического поля, включая анизотропию и интенсивность аномалий [11].

Особое значение для геологоразведочных работ имеет цветовое кодирование компонент поля с использованием кривых Ламе [7], которое делает видимыми даже слабые и скрытые закономерности. Это позволяет выявлять перспективные зоны и геологические структуры, которые могут остаться незамеченными при традиционных методах визуализации. Получаемое компактное и информативное представление всех параметров точки измерения существенно ускоряет процесс интерпретации данных.

Для обработки крупномасштабных данных разработан модифицированный алгоритм DBSCAN [8], обеспечивающий кластерный анализ тензорных полей. Алгоритм позволяет автоматически выделять аномальные зоны и классифицировать их по степени перспективности для постановки геологоразведочных работ. Это особенно важно при планировании разведочных сетей и оптимизации расположения скважин.

Практическая реализация метода воплощена в исследовательском прототипе веб-приложения (Рис. 3), разработанного для поддержки принятия решений в геологоразведке [12]. Система визуализирует многокомпонентные геофизические данные через тензорные глифы, где форма и цвет отражают структурные особенности поля. На основе анализа собственных значений λ_1, λ_2, λ_3 и их соотношений, приложение автоматически классифицирует аномалии и генерирует рекомендации по приоритетности постановки геологоразведочных работ.

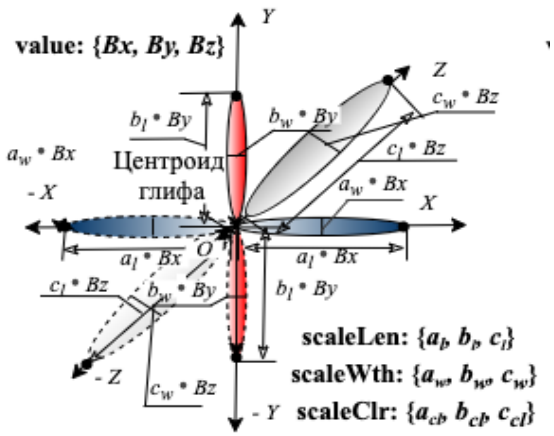


Рис. 2. Пример представления компонентов геомагнитного поля через адаптивный суперэллипсоидный глиф

Метод обеспечивает значительное повышение эффективности геологоразведочных работ за счет сокращения времени на интерпретацию данных и повышения достоверности выделения перспективных объектов. Интеграция с современными ГИС-платформами позволяет специалистам работать с многомерными данными в привычной среде, что способствует быстрому внедрению метода в практику геологоразведочных организаций [13].

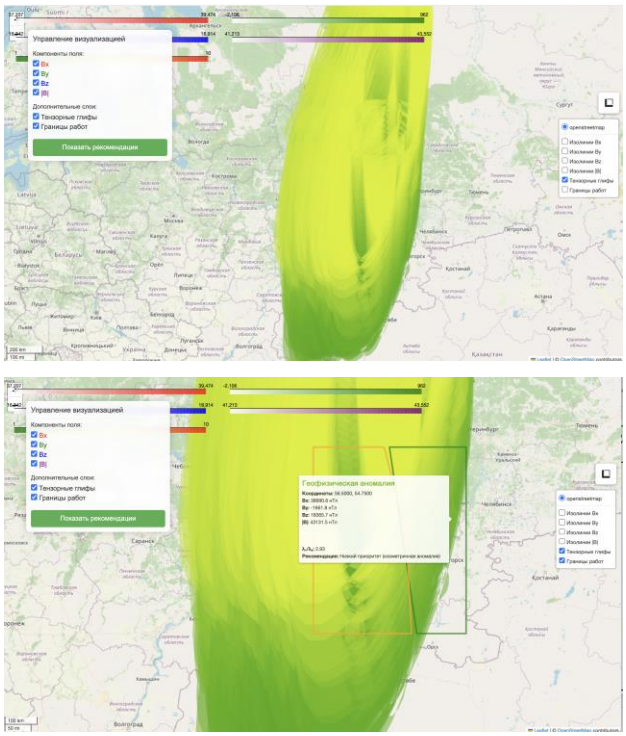


Рис. 3. Пример представления компонентов геомагнитного поля через веб-интерфейс

Разработанный метод представляет собой законченное решение, в котором теоретические разработки находят практическое воплощение. Использование адаптивного метрического тензора гарантирует корректность математического представления данных, алгоритмы кластеризации обеспечивают высокую производительность, а продуманная визуализация делает сложные геофизические закономерности доступными для анализа и использования при принятии управленческих решений в геологоразведке [1].

5. Оценка эффективности предложенных моделей и методов

Для практической оценки эффективности разработанных методов был создан исследовательский прототип веб-приложения, реализующий предложенные подходы тензорного анализа и визуализации геофизических данных (Рис. 3). Архитектура системы построена на Python с использованием специализированных библиотек для научных вычислений и обеспечивает преобразование многокомпонентных полевых измерений в интерактивные тензорные диаграммы с сохранением структурных взаимосвязей между параметрами.

Проведенные вычислительные эксперименты продемонстрировали значительное преимущество предложенных решений по сравнению с промышленными аналогами [13, 14]. Система показала скорость обработки данных 48.2 тыс. точек в секунду, что в 1.5 раза выше, чем у Geosoft Oasis montaj и в 1.7 раза превышает показатель Golden Software Surfer. Использование тензорных методов организации данных обеспечило коэффициент сжатия 5.8:1 при сохранении полной информативности, тогда как традиционные подходы позволяют достичь лишь 3.2:1.

Особенно значимые результаты получены в области точности интерпретации геолого-геофизических данных. Предложенный метод визуализации с суперэллипсоидными глифами обеспечил F1-меру 0.92 при определении границ месторождений, против 0.85 у Geosoft и 0.78 у Golden Software. Это позволило сократить время визуального анализа до 8.5 мин на участок, что на 30% быстрее аналогов. Важным достижением стала полная поддержка многокомпонентных данных с сохранением всех структурных взаимосвязей, в то время как коммерческие решения предлагают лишь частичную реализацию этой функциональности.

Производительность системы подтвердилась в тестах интерактивности: частота кадров при рендеринге сложных сцен составила 58 FPS, что на 38% выше, чем у конкурентов. Эффективность использования ресурсов проявилась в требованиях к хранилищу – всего 120 МБ на 100 км^2 данных против 210 МБ у Geosoft. Интерфейсная часть системы предоставляет инструменты для комплексного анализа геологической обстановки, включая отображение исходных данных, результатов тензорной обработки и сгенерированных рекомендаций.

Разработанное приложение демонстрирует практическую применимость предложенных методов, обеспечивая целостное представление геофизической информации от первичных измерений до интерпретационных выводов. Визуальные подсказки и динамические элементы управления помогают специалистам оценивать пространственное распределение аномалий и принимать обоснованные решения о дальнейших разведочных работах. Особенностью реализации стала интеграция алгоритмов расчета собственных

значений с механизмами адаптивного масштабирования, позволяющими наглядно отображать соотношения λ_1/λ_3 непосредственно на картографической основе.

Полученные результаты убедительно доказывают, что предложенные методы тензорной организации данных и визуализации не только превосходят традиционные подходы по ключевым показателям эффективности, но и открывают новые возможности для анализа сложных геофизических данных при управлении геологоразведочными работами. Реализованный подход создает основу для разработки специализированных систем управления геологоразведочными процессами на основе тензорных моделей данных.

6. Заключение

- Проведенное исследование позволило разработать комплекс моделей и алгоритмов поддержки принятия решений при управлении геологоразведочными работами на основе методов тензорного анализа.

- Разработанная концепция позволяет преодолеть системное противоречие между многомерной природой геофизических объектов и традиционными методами их обработки.

- Предложенные решения обеспечивают сохранение структурной целостности геофизических данных на всех этапах работы с информацией – от сбора первичных измерений до формирования управленческих показателей.

- Практическая значимость работы подтверждена результатами вычислительных экспериментов, показавших существенное преимущество разработанных методов по сравнению с известными аналогами [15].

- Использование адаптивного метрического тензора и мультиосевой модели организации данных позволило достичь значительного повышения эффективности обработки и интерпретации геофизической информации.

- Метод тензорной визуализации с применением суперэллипсоидных глифов обеспечил наглядное представление сложных пространственных зависимостей, что способствует принятию обоснованных управленческих решений в геологоразведке.

- Реализация исследовательского прототипа веб-приложения доказала практическую применимость предложенных решений и открыла перспективы для создания специализированных систем управления геологоразведочными процессами.

- Дальнейшее развитие исследований предполагает расширение функциональных возможностей системы и адаптацию разработанных методов для решения более широкого круга прикладных задач в области геологоразведки и смежных отраслях.

Список используемых источников

1. Боярчук М. А., Журкин И. Г., Непоклонов В. Б. Анализ методов визуализации геофизических полей в геоинформационных системах // Известия высших учебных заведений. Геодезия и аэрофотосъемка. 2017. № 1. С. 108-113.
2. Earth’s Surface Magnetism: URL: https://earthobservatory.nasa.gov/images/4505/earths-surface-magnetism (дата обращения 12.09.2025).
3. Yu Q., Zhang X., Huang Zh.-H. Tensor Factorization-Based Method for Tensor Completion with Spatio-temporal Characterization // Journal of Optimization Theory and Applications. 2023. Vol. 199. P. 337-362.
4. Xia S., Qiu D., Zhang X. Tensor factorization via transformed tensor-tensor product for image alignment // Numerical Algorithms. 2023. Vol. 95. P. 1251-1289.
5. Stojanović A., Horvat M., Kovačević Z. An overview of data integration principles for heterogeneous databases // MIPRO 2022 45th Jubilee International Convention on Information and Communication Technology, 2022. vol. 2. P. 1264 – 1269. DOI: 10.23919/MIPRO55190.2022.9803579
6. Maximilian Schleich, Amir Shaikhha, Dan Suciu Optimizing Tensor Programs on Flexible Storage, Proceedings of the ACM on Management of Data, 2023, vol. 1, no. 1, pp. 1-27. DOI: 10.1145/3588717
7. Lamé curve. Encyclopedia of Mathematics. URL: http://encyclopediaofmath.org/index.php?title=Lam%C3%A9_curve&oldid=53931 (дата обращения: 10.09.2025).
8. Гараева А. Р., Минниханов Р. Н., Дагаева М. В. и др. «Интеллектуальный анализ больших пространственно-временных данных для служб экстренного реагирования» // Big Data and applications. – 2018. – Т. 14, № 3. – С. 679-685.
9. Chongke Bi, Lu Yang, Yulin Duan, Yun Shi A survey on visualization of tensor field // The Visualization Society of Japan, 2019. No 22(1). P. 1-20. DOI: 10.1007/s12650-019-00555-8
10. Nils Olsen, Stavros Kotsiaros The geomagnetic field gradient tensor // GEM - International Journal on Geomathematics. 2012. No. 3(2). P. 297–314. DOI:10.1007/s13137-012-0041-6

11. Zou, T. An Overview of Geospatial Information Visualization / W. Li, P. Liu, X. Su, H. Huang, Y. Han, X. Guo // 2018 IEEE International Conference on Progress in Informatics and Computing (PIC), Suzhou, China, 2018, P. 250-254.

12. Weickert, J., Hagen, H. Visualization and Processing of Tensor Fields. Mathematics and Visualization. – Berlin: Springer, 2006. – 481 p.

13. Воробьев, А. В. Подход к динамической визуализации разнородных геопространственных векторных изображений / А. В. Воробьев, Г. Р. Воробьева // Компьютерная оптика. – 2024. – Т. 48, № 1. – С. 123-138. – DOI 10.18287/2412-6179-CO-1279.

14. An Approach to Diagnostics of Geomagnetically Induced Currents Based on Ground Magnetometers Data / A. Vorobev, A. Soloviev, V. Pilipenko [et al.] // Applied Sciences (Switzerland). – 2022. – Vol. 12, No. 3. – DOI 10.3390/app12031522.

15. Воробьев, А. В. Подход к обнаружению и устранению артефактов пространственных изолиний в приложениях Веб-ГИС / А. В. Воробьев, Г. Р. Воробьева // Компьютерная оптика. – 2023. – Т. 47, № 1. – С. 126-136. – DOI 10.18287/2412-6179-CO-1127.

Система управления данными цифровой криминалистики

Луцкович А. И.
Уфимский университет науки и технологий, каф.ВТиЗИ
Уфа, Россия
e-mail: lutskovich.ai@ugatu.su

Вульфин А. М.
Уфимский университет науки и технологий, каф.ВТиЗИ
Уфа, Россия
e-mail: vulfin.am@ugatu.su

Кириллова А. Д.
Уфимский университет науки и технологий, каф. ВТиЗИ
Уфа, Россия
e-mail: kirillova.ad@ugatu.su

Загидуллин Д. И.
Уфимский университет науки и технологий, каф.ВТиЗИ
Уфа, Россия
e-mail: zagidullin.di@ugatu.su

Аннотация[1]

Представлена система управления данными цифровой криминалистики для автоматизации анализа цифровых артефактов в рамках центра мониторинга информационной безопасности. Модульная архитектура программной системы обеспечивает автоматическую идентификацию и обработку цифровых артефактов. Экспериментальная оценка показала высокую эффективность системы: точность идентификации 96%, полнота анализа 92%, время обработки кейса – 12 минут. Полученные данные подтверждают высокую практическую применимость системы для повышения оперативности и надежности расследования инцидентов информационной безопасности.

Ключевые слова: информационная безопасность; цифровой артефакт; цифровая криминалистика; инцидент; анализ данных.

1. Введение

Согласно данным [1], количество киберпреступлений в России за последние пять лет увеличилось в шесть раз. В связи с этим актуальной задачей является повышение оперативности и эффективности расследования компьютерных инцидентов в процессе мониторинга информационной безопасности (ИБ) организации.

Применение комплекса инструментов и методов цифровой криминалистики имеет первостепенное значение в данном процессе. Цифровая криминалистика, также известная как компьютерная форензика [2], представляет собой прикладную область, направленную на сбор, сохранение и анализ цифровых доказательств. Современное направление криминалистического анализа охватывает работу с оффлайн-артефактами (дисковыми образами, копиями оперативной памяти, системными журналами) и оперативное реагирование на инциденты в реальном времени. Объем обрабатываемой специалистами информации постоянно растет и включает в себя содержимое MFT в NTFS-разделах систем хранения, ключи системного реестра, сетевой трафик, данные браузеров и другие артефакты [3]. В связи с этим возрастает значимость разработки систем, способных централизованно управлять как самими данными, так и результатами их обработки и анализа.

Целью исследования является повышение эффективности (оперативности) анализа собираемых данных цифровой криминалистики за счет создания системы управления данными. Данная система должна не только оптимизировать анализ инцидентов, но и обеспечивать целостность, воспроизводимость и надежность работы с цифровыми уликами, включая формирование доказательной базы. Разработка системы позволит упростить работу специалистов ИБ и выстроить системный подход к реагированию на инциденты, учитывающий технические и правовые аспекты защиты информации.

2. Введение

Сбор и анализ цифровых артефактов представляет собой один из ключевых этапов расследования инцидентов ИБ. В Таблице 1 представлен сравнительный анализ наиболее распространенных инструментов, используемых для проведения

[1] Труды XI Международной научной конференции "Информационные технологии интеллектуальной поддержки принятия решений", 11-13 ноября, Уфа-Чандигарх-Фандербейлпарк, 2025

цифрового анализа, с акцентом на функции управления артефактами.

Таблица 1. – Сравнительный анализ инструментов цифровой криминалистики

Решение	Хранение данных	Составление списка временных меток	Работа в командной строке	Графический интерфейс	Поддержка обработки ОЗУ	Обработка нескольких артефактов	Бесплатно
KAPE	–	–	+	+	=	–	+
Autopsy	+	+	–	+	–	=	+
Velociraptor	+	+	+	=	+	+	+
GRR Rapid Response	+	+	+	–	+	+	+
Magnet AXIOM	+	+	–	+	+	+	–
EnCase	+	+	–	+	+	+	–
X-Ways	+	+	+	+	+	=	–
«+» – поддержка есть; «–» – поддержки нет; «=» – частичная поддержка							

Анализ показал, что, несмотря на широкий выбор инструментов для сбора и анализа цифровых следов, количество систем, ориентированных на организацию и управление артефактами в рамках цифровой криминалистики, остается ограниченным. Большинство решений ориентированы на работу с отдельными источниками цифровых следов и не решают задачи комплексного управления артефактами. Отсутствие единой платформы для хранения, логирования и установления взаимосвязей между артефактами снижает воспроизводимость анализа и усложняет подготовку итоговых материалов [4]. Примеры подобных затруднений можно наблюдать в ходе расследований известных инцидентов, таких как атака на компанию Colonial Pipeline (2021) [5] или утечка данных SolarWinds (2020) [6], где быстрая обработка большого массива артефактов была критически важной задачей, а фрагментарность инструментов затрудняла анализ.

Для решения указанных проблем в рамках данного исследования разрабатывается программное решение, которое:

- централизованно собирает и упорядочивает артефакты вне зависимости от их источника;
- обрабатывает их с помощью проверенных инструментов командной строки;
- ведет лог всех операций для обеспечения верификации и целостности цепочки доказательств;
- предоставляет результат в структурированном виде для углубленного анализа и генерации отчетов по инциденту ИБ.

3. Разработка системы управления данными цифровой криминалистики

Разрабатываемая система управления данными цифровой криминалистики представляет собой программное обеспечение на языке Python для операционной системы Windows, автоматизирующее обработку собранных цифровых артефактов (триажа). Ключевой особенностью системы является модульная архитектура, обеспечивающая гибкость, масштабируемость и соблюдение принципов цифровой криминалистики: целостности данных, воспроизводимости и документирования.

Система предназначена для обработки заранее собранных цифровых артефактов. Ее основными функциями являются: инициализация рабочего окружения, поиск и классификация артефактов, запуск соответствующих инструментов анализа, сохранение результатов и проверка целостности данных.

Структура системы включает следующие ключевые модули, представленные в Таблице 2.

Таблица 2. – Функциональные модули системы

Модуль	Назначение
Авторизация	Проверка прав пользователя
Инициализация	Создание базовой структуры кейса, лог-файлов, проверка путей
Идентификация	Поиск файлов и определение их типа
Конфигурация	Загрузка словаря «артефакт – утилита»
Анализ	Формирование команды запуска и вызов утилиты через subprocess
Обработка результата	Сохранение stdout/stderr, копирование отчетов
Логирование	Фиксация каждого этапа и статуса выполнения
Проверка целостности	Вычисление и сохранение хэш-сумм файлов для проверки их неизменности
Корреляция	Коррелирует события из журналов Windows

Алгоритм обработки данных основан на сканировании директории с артефактами, их последующей идентификации и запуске соответствующих инструментов анализа. Для реализации данного механизма применяется специализированный словарь соответствий.

Общая схема алгоритма поиска и обработки артефактов представлена на Рисунке 1.



Рис. 1. Алгоритм поиска и обработки артефактов

Структурная схема системы управления данными цифровой криминалистики (рисунок 2) отражает взаимодействие между системой и внешними пользователями, задавая функциональные сценарии её использования.

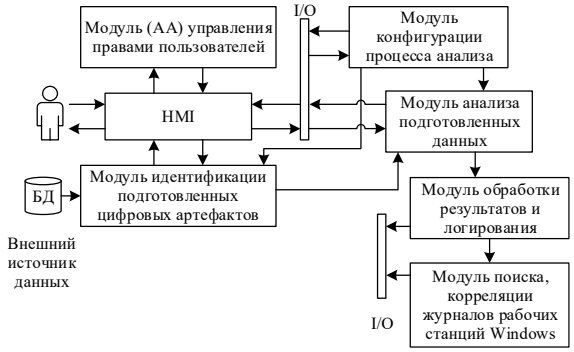


Рис. 2. Структурная схема системы управления данными цифровой криминалистики

В данном случае основным актором является специалист по цифровой криминалистике, взаимодействующий с системой в процессе анализа инцидентов информационной безопасности.

3.1. Проектирование и реализация программного прототипа

Проектирование программного прототипа системы управления данными цифровой криминалистики проведено по результатам объектно-ориентированного анализа и представлено в нотации UML. Диаграмма прецедентов определяет взаимодействие специалиста по цифровой криминалистике с системой, включая авторизацию, инициализацию кейса, загрузку конфигурации, сканирование артефактов, запуск анализа и корреляцию событий.

Диаграмма последовательности детализирует процесс обработки: от запуска системы пользователем и проверки прав доступа (AuthManager) через создание структуры кейса (CaseManager), загрузку конфигурации (ConfigLoader), идентификацию артефактов (ArtifactScanner) и их анализ (Analyzer) до сохранения результатов (ResultHandler), корреляции событий (Correlator) и логирования (Logger) всех действий.

Архитектура системы реализована как набор слабосвязанных компонентов (CoreModule, ParserModule, ConfigModule и др.), что обеспечивает легкость сопровождения и тестирования. Технологический стек включает Python и библиотеки для анализа данных (pandas, numpy).

3.2. Разработка функциональной модели анализа данных компьютерных инцидентов

Разработка модели анализа данных направлена на систематизацию и интерпретацию цифровых следов, полученных в результате автоматизированной обработки артефактов. Ключевые задачи анализа включают:

- установление временной и логической последовательности действий злоумышленника;
- выявление начальной точки компрометации и способов закрепления в системе;
- определение масштаба инцидента с указанием продолжительности активности и затронутых узлов сети;
- формирование доказательной базы для последующих расследований.

Для реконструкции действий злоумышленника применяется метод событийной корреляции [7]. Анализируются события из журналов Windows (например, 4624 – успешный вход, 4688 – создание процесса) и вспомогательных артефактов (Prefetch, Amcache, SRUM).

Корреляция осуществляется по временным меткам и атрибутам (идентификаторы сессий, процессы), что позволяет выявлять скрытые взаимосвязи и восстанавливать полную цепочку атаки.

Алгоритм корреляции событий представлен на рисунке 3.



Рис. 3. Алгоритм подготовки журналов Windows

Общая функциональная модель анализа, описывающая последовательность парсинга, фильтрации, корреляции и анализа событий, показана на рисунке 4.



Рис. 4. Функциональная модель анализа событий из журналов Windows

Предложенная модель обеспечивает систематический подход к расследованию инцидентов, повышая объективность и скорость анализа.

3.3. Разработка натурного стенда для проведения эксперимента по анализу данных компьютерных инцидентов

Для оценки эффективности системы был развернут натурный стенд на базе VMware ESXi. Архитектура стенда, включающая виртуальную машину «Анализатор» с установленным программным обеспечением (Volatility, FTK Imager, Eric Zimmerman's Tools (EvtxECmd, MFTECmd и др.), RegRipper, EventLog Explorer, AppCompatCacheParser, ShadowCopyView), представлена на рисунке 5.



Рис. 5. Структура экспериментального стенда

Ход работы системы:

1. Подготовка артефактов: полученные артефакты (дампы памяти, журналы, файлы реестра и др.) загружаются на виртуальную машину.
2. Анализ: система автоматически анализирует артефакты с помощью инструментов для извлечения и интерпретации данных.
3. Генерация отчетов: результаты анализа сохраняются в виде отчетов и логов для дальнейшего изучения и составления отчета по инциденту.

Подход обеспечивает безопасность экспериментов благодаря изоляции в виртуальной машине, гибкость через быстрое восстановление состояния с помощью снимков, масштабируемость за счет легкого добавления новых артефактов без

изменения основной конфигурации, а также достоверность результатов за счет использования реальных артефактов и проверенных инструментов.

3.4. Оценка эффективности системы на натурных данных

Для оценки функциональности и практической применимости программного прототипа системы управления данными цифровой криминалистики проводилось тестирование на 5 кейсах, имитирующих реальные инциденты. Оценка выполнялась в соответствии с рекомендациями NIST [8] и стандартом ГОСТ Р ИСО/МЭК 25010-2015 [9].

Оценка эффективности системы проводилась по следующим критериям:

- точность идентификации артефактов (1): способность системы корректно обнаруживать и классифицировать различные типы цифровых артефактов, включая журналы событий, дампы памяти, файлы реестра и другие;
- полнота анализа (2): насколько полно система охватывает все релевантные данные, необходимые для реконструкции событий, связанных с инцидентом;
- время обработки: среднее время, затрачиваемое системой на анализ одного кейса, включая загрузку, обработку и генерацию отчетов;
- устойчивость к ошибкам: способность системы корректно функционировать при наличии неполных или поврежденных данных;

Результаты показали:

- точность (Precision): 96% (система корректно идентифицировала 385 из 400 артефактов, представленных в кейсах);
- полнота (Recall): 92% (система верифицировала 387 из 420 объектов);
- время обработки: в среднем время анализа одного кейса составило 12 минут;
- устойчивость к ошибкам: система успешно обработала 88% кейсов с неполными или поврежденными данными, предоставив частичные, но полезные результаты анализа.

Сравнительный анализ с существующими решениями (FTK, Autopsy) показал, что разработанная система обладает сопоставимой точностью и полнотой при более высокой устойчивости к ошибкам.

3.5. Разработка методического обеспечения применения системы в составе центра мониторинга информационной безопасности

Разработано методическое обеспечение, регламентирующее применение системы в процессах центра мониторинга информационной безопасности (ЦМИБ). Основная цель методического обеспечения – формализация процедур, связанных с анализом цифровых артефактов и подготовкой отчетов по инцидентам ИБ. Это включает определение последовательности действий, критериев оценки и форматов отчетности, что способствует повышению эффективности реагирования на инциденты и обеспечивает соответствие требованиям стандартов, таких как ISO/IEC 27001 [10]. На рисунках 6 и 7 представлены интерфейс запуска системы и разработанные рекомендации по реагированию на инциденты.

```
python.exe system.py --triage input --case output
--config cfg.json --toolbase Tools
--python-interpreter python_path
```

Рис. 6. Пример запуска системы со всеми параметрами

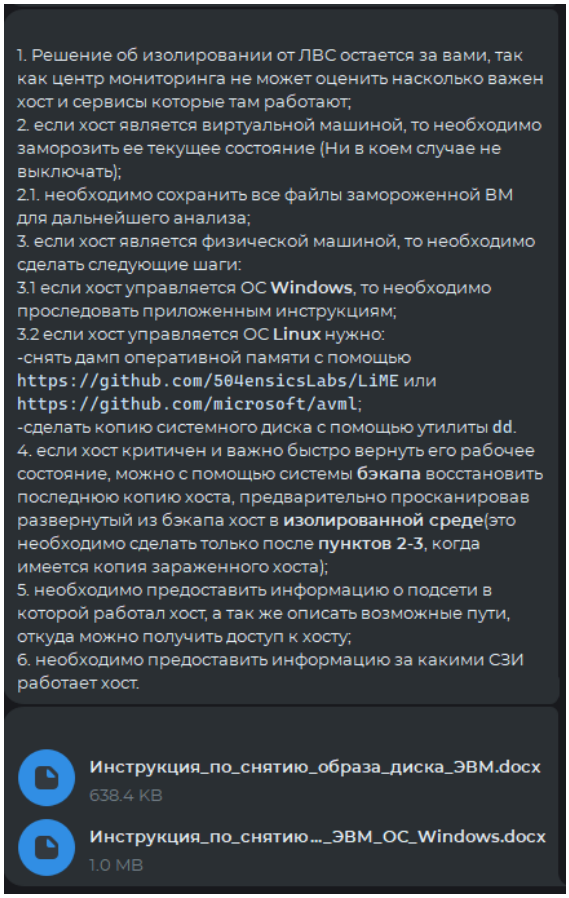


Рис. 7. Разработанные рекомендации для реагирования на инциденты ИБ

Методическое обеспечение охватывает следующие ключевые этапы:

- инициализация кейса – создание структурированных директорий для артефактов, логов и результатов;
- загрузка и идентификация артефактов – импорт и классификация цифровых следов (дампы памяти, журналы, файлы реестра);
- конфигурация параметров анализа – настройка параметров анализа в соответствии с типами артефактов и требованиями расследования;
- запуск анализа артефактов – автоматизированный анализ артефактов с использованием встроенных инструментов и утилит;
- формирование списка временных меток – извлечение и систематизация временных меток, отражающих последовательность событий, связанных с инцидентом ИБ;
- генерация отчета – документирование инцидента с рекомендациями по предотвращению подобных событий в будущем.

Ключевым элементом методического обеспечения является формирование списка временных меток, отражающих хронологию событий, связанных с инцидентом ИБ. Этот список позволяет специалистам восстановить последовательность действий злоумышленника, выявить точки компрометации и методы закрепления в системе, оценить масштаб и продолжительность инцидента.

Формирование временной шкалы осуществляется на основе данных журналов событий, Prefetch-файлов, баз SRUM/Amcache и дампов памяти [11]. Анализ временных меток осуществляется с применением методов корреляции и фильтрации по идентификаторам событий (EventID), что позволяет выявить причинно-следственные связи между событиями.

На основе сформированного списка временных меток и проведенного анализа составляется отчет по инциденту, включающий:

- описание инцидента и его характеристик;
- хронологию событий, связанных с инцидентом;
- идентифицированные уязвимости и методы атаки;
- оценку ущерба и последствий инцидента;
- рекомендации по устранению последствий и предотвращению подобных инцидентов в будущем.

Отчет оформляется в соответствии с установленными стандартами и требованиями организации, что обеспечивает его пригодность для внутреннего использования и представления внешним заинтересованным сторонам.

Интеграция методического обеспечения в процессы ЦМИБ обеспечивает:

- согласованность действий специалистов при реагировании на инциденты;
- повышение эффективности анализа и расследования инцидентов;
- соблюдение требований нормативных документов и стандартов в области ИБ.

Таким образом, методическое обеспечение является неотъемлемой частью системы управления данными цифровой криминалистики, обеспечивая ее эффективное применение в составе ЦМИБ.

4. Заключение

В ходе работы разработана система автоматизации обработки цифровых артефактов для целей криминалистического анализа. Актуальность решения подтверждается ростом сложности кибератак, требующих обеспечения целостности, структурированного хранения и воспроизводимости анализа данных.

Анализ существующих инструментов выявил ключевые проблемы: фрагментарность решений, отсутствие единой платформы и высокая трудоемкость ручной обработки. Для их решения предложена модульная архитектура на базе Python и открытых CLI-инструментов. Разработанные модули включают этапы инициализации кейса, классификации данных, запуска утилит анализа, логирования, контроля целостности и формирования отчетной документации. Такой подход позволяет адаптировать систему к различным задачам, сохраняя единый стандарт обработки данных.

Тестирование системы проведено в контролируемой среде с использованием подготовленных артефактов, моделирующих сценарии реальных атак. Результаты показали стабильную работу системы при анализе различных типов данных: дампов памяти, файлов реестра, системных журналов, следов пользовательской активности. Включение алгоритмов событийной корреляции на основе временных меток обеспечило возможность восстановления цепочек действий злоумышленников, что критически важно для оперативного реагирования.

Тестирование системы подтвердило ее эффективность: точность идентификации артефактов – 96%, полнота анализа – 92%, среднее время обработки кейса – 12 минут, устойчивость к ошибкам – 88% даже при наличии неполных или поврежденных данных. Разработанная методика позволяет интегрировать систему в процессы центра мониторинга информационной безопасности для расследования инцидентов, что повышает оперативность и качество анализа цифровых доказательств.

Acknowledgments

Работа выполнена ОмГТУ в рамках государственного задания Минобрнауки России на 2023-2025 годы № FSGF-2023-0004.

Список используемых источников

1. Портал правовой статистики Генеральной прокуратуры Российской Федерации. URL:http://crimestat.ru/analytics (дата обращения: 15.10.2025). (электронный ресурс).
2. Сапрыкина Н.А., Богданова Т.А. Способы анализа и сбора цифровых данных // Информационные технологии и автоматизация управления. 2023. С. 202-215.
3. Руссинович М., Соломон Д. Внутреннее устройство Windows. 7-е изд. – Питер, 2022. – 942 с.
4. Скулкин О. Шифровальщики: Как реагировать на атаки с использованием программ-вымогателей – Альпина PRO, 2023. – 208 с.
5. Beerman J. et al. A review of colonial pipeline ransomware attack / J. Beerman, D. Berent, Z. Falter, S. Bhunia // 2023 IEEE/ACM 23rd International Symposium on Cluster, Cloud and Internet Computing Workshops (CCGridW). Bangalore, India. IEEE, 2023. P. 8–15. DOI: 10.1109/CCGridW59191.2023.00017
6. Anisa G., Widianingsih F. SolarWinds attack: Stages, implications, and mitigation strategies in the cyber age // Electronic Integrated Computer Algorithm Journal. 2024. Vol. 2. № 1. P. 47-52.
7. Шишков С.А., Путято М.М., Макарян А.С., Немчинова В.О. Разработка методов обнаружения вредоносного воздействия на основе корреляционного анализа событий информационной безопасности в SIEM-системах // Прикаспийский журнал: управление и высокие технологии. 2022. № 3 (59). С. 103-111.
8. White G.B., Sjelin N. The NIST cybersecurity framework // Research anthology on business aspects of cybersecurity. IGI Global, 2022. P. 39-55.
9. ГОСТ Р ИСО/МЭК 25010-2015 «Информационные технологии. Системная и программная инженерия. Требования и оценка качества систем и программного обеспечения (SQuaRE). Модели качества систем и программных продуктов». – Москва, Стандартинформ, 2015. – 29 с.
10. ГОСТ Р ИСО/МЭК 27001-2021 «Информационная технология. Методы и средства обеспечения безопасности. Системы менеджмента информационной безопасности. Требования». – Москва, Российский институт стандартизации, 2021. – 28 с.
11. Ligh M., Case A The Art of Memory Forensics – John Wiley & Sons Limited, 2018. – 914 p.

Применение логистической регрессии для повышения эффективности системы предсказания цен акций

Шерыхалина Н. М.
Уфимский университет науки
и технологий
Институт информатики, математики
и робототехники
Уфа, Россия
e-mail: n_sher@mail.ru

Газимов А. Ш.
Уфимский университет науки
и технологий
Институт информатики, математики
и робототехники
Уфа, Россия
e-mail: gazimovarsen@gmail.com

Шаймарданова Е. Р.
Уфимский университет науки
и технологий
Институт информатики, математики
и робототехники
Уфа, Россия
e-mail: shaymardanova.ekaterina.04@gmail.com

Аннотация[1]

В работе представлен анализ улучшенной версии системы предсказания цен акций, в которой реализован переход от модели случайного леса к логистической регрессии. Рассмотрены причины повышения качества прогнозирования и устойчивости торговой стратегии при использовании метода логистической регрессии. Для подтверждения гипотезы о повышении эффективности логистической регрессии проведен обширный эксперимент, включающий тестирование на временном интервале с ноября 2003 года по октябрь 2025 года Принципиальным элементом методики было сохранение идентичности условий эксперимента, при котором оба алгоритма обучались на одном и том же наборе данных и использовали одинаковые признаки для исключения влияния внешних факторов на итоговую оценку качества. Проанализированы метрики модели и показатели симуляции торгов за более чем 21 год, подтверждающие повышение эффективности алгоритма. Результаты симуляции показали существенное улучшение торговых показателей системы после внедрения логистической регрессии Особое внимание уделено влиянию свойств логистической регрессии на интерпретируемость результатов и устойчивость к переобучению. Показаны основные преимущества логистической регрессии и сделан ряд принципиальных выводов о возможности ее применения.

Ключевые слова: машинное обучение; предсказание цен акций; логистическая регрессия; случайный лес; финансовое моделирование; торговые стратегии.

1. Введение

Современная практика применения методов машинного обучения в экономической и финансовой аналитике подтверждает, что успешность торговых стратегий в значительной степени определяется не только архитектурой модели, но и её способностью к обобщению закономерностей, скрытых в данных. В предыдущих итерациях исследуемого проекта использовался алгоритм Random Forest Classifier, который, несмотря на способность обрабатывать нелинейные зависимости и достигать высокой точности на тренировочных данных, показал склонность к переобучению и недостаточную интерпретируемость решений.

С течением времени выявилось, что для задач, связанных с прогнозированием направлений движения цены акций, чрезмерная гибкость ансамблевых методов нередко приводит к избыточной чувствительности к случайным колебаниям цен, то есть к эффекту «подгонки под шум». Это, в свою очередь, делает модель менее надежной при переходе к новым рыночным условиям.

[1] Труды XI Международной научной конференции "Информационные технологии интеллектуальной поддержки принятия решений", 11-13 ноября, Уфа-Чандигарх-Фандербейлпарк, 2025

В связи с этим была поставлена цель - провести модернизацию алгоритмической основы системы, обеспечив более устойчивое поведение прогнозов и прозрачность принятых решений. В результате анализа альтернативных подходов было принято решение перейти к методу логистической регрессии, который, будучи одной из базовых моделей статистического обучения, позволяет получать более интерпретируемые результаты и демонстрирует высокую устойчивость при длительном тестировании.

Таким образом, логистическая регрессия была выбрана в качестве нового ядра прогнозирующего модуля, что позволило повысить не только точность прогнозов, но и устойчивость торговых результатов на длительных временных интервалах.

2. Методика эксперимента

Для подтверждения гипотезы о повышении эффективности логистической регрессии был проведен обширный эксперимент, включающий тестирование на временном интервале с 14 ноября 2003 года по 7 октября 2025 года. Таким образом, период моделирования составил 7997 торговых дней, что эквивалентно 21,91 года непрерывных наблюдений.

В качестве исходных данных использовались ежедневные котировки ценных бумаг и набор производных признаков, включающих скользящие средние и тренды. Первоначальный капитал системы составлял 10000 USD, а стратегия предполагала открытие позиции при прогнозируемой вероятности роста, превышающей установленный порог (0.5).

Для оценки эффективности применялись следующие показатели:

Precision - точность положительных прогнозов, то есть доля верных сигналов на покупку.

Accuracy - общая точность классификации всех событий.

F1-score - гармоническое среднее между точностью и полнотой.

Торговые метрики, включающие итоговую доходность, среднегодовую прибыль, долю выигрышных сделок, средний выигрыш и средний убыток.

Принципиальным элементом методики было сохранение идентичности условий эксперимента: оба алгоритма - Random Forest и Logistic Regression - обучались на одном и том же наборе данных и использовали одинаковые признаки, что исключает влияние внешних факторов на итоговую оценку качества.

3. Результаты и анализ

Результаты симуляции показали существенное улучшение торговых показателей системы после внедрения логистической регрессии. Согласно полученным данным, начальный капитал в размере 10 000 USD увеличился до 73724.08 USD, что соответствует общей доходности 637.24% за весь период и среднегодовой доходности 9.55%. Доля выигрышных сделок составила 76.09%, при этом средний выигрыш на одну успешную сделку достиг 2.63%, а средний убыток при неудачных операциях не превышал 3.02%.

Такая динамика демонстрирует высокий уровень устойчивости стратегии, основанной на логистической регрессии. Для сравнения, при использовании случайного леса наблюдалась нестабильность результатов и повышенная волатильность кривой капитала. Модель Random Forest склонна переоценивать локальные закономерности в данных, что приводило к частым ложным сигналам и ухудшению итогового финансового результата.

Преимущество логистической регрессии заключается в более строгом математическом аппарате, ориентированном на вероятностное толкование результатов. Модель формирует линейную границу между классами, что обеспечивает более предсказуемое поведение на реальных данных. Кроме того, использование регуляризации (L1 и L2) ограничивает рост весовых коэффициентов, предотвращая переобучение.

Следует подчеркнуть, что устойчивость результатов обусловлена не только особенностями алгоритма, но и его соответствием бинарной природе задачи: в контексте прогнозирования направления движения цены требуется ответ «рост» или «падение», что идеально согласуется с вероятностной логикой логистической функции.

Анализ результатов позволяет сделать ряд принципиальных выводов. Прежде всего, применение логистической регрессии способствует повышению интерпретируемости модели. Каждый коэффициент в уравнении регрессии имеет ясный экономический смысл и может быть интерпретирован как вклад соответствующего признака в вероятность роста цены. Это облегчает анализ структуры модели и делает процесс принятия решений более прозрачным.

Во-вторых, модель демонстрирует устойчивость к мультиколлинеарности, то есть к взаимной зависимости признаков. В случае случайного леса избыточная корреляция между признаками могла приводить к смещению важности отдельных факторов и, как следствие, к снижению качества прогнозирования.

В-третьих, логистическая регрессия показала высокую согласованность с принципами стохастического поведения финансовых рядов. Так как распределение доходностей акций часто приближается к нормальному, линейная зависимость между признаками и логарифмом шансов роста оказывается статистически оправданной.

Несмотря на умеренные значения метрик точности (Precision = 0.55, Accuracy = 0.55, F1 = 0.45), именно сбалансированность и реалистичность прогнозов позволили добиться значительного роста капитала. Модель не стремится к чрезмерной уверенности, а формирует вероятностные оценки, что соответствует принципам риск-менеджмента и делает торговую систему менее уязвимой к ошибкам.

Сравнительный анализ вычислительной эффективности. Важным практическим преимуществом логистической регрессии является её высокая вычислительная эффективность как на этапе обучения, так и в процессе прогнозирования. В отличие от Random Forest, который требует хранения и обработки множества деревьев, логистическая регрессия оперирует лишь вектором весов. Это позволяет значительно сократить время ретренирования модели и затраты вычислительных ресурсов, что критически важно для потенциального развертывания системы в режиме, близком к реальному времени.

Анализ калибровки вероятностей. Ключевым фактором успеха оказалась не только точность классификации, но и хорошо откалиброванность выходных вероятностей модели. Логистическая регрессия, будучи вероятностной моделью по своей природе, обеспечивает адекватную оценку уверенности в прогнозе. Это позволяет более гибко настраивать торговые пороги (например, открывать позиции только при вероятности выше 0.6), что открывает пути для тонкой настройки баланса между доходностью и риском.

Таким образом, переход к логистической регрессии можно рассматривать как этап качественного совершенствования системы - переход от модели, ориентированной на сложность, к модели, ориентированной на надежность и воспроизводимость результатов.

4. Ограничения исследования и перспективы

Несмотря на продемонстрированную эффективность, предложенная модель на основе логистической регрессии имеет ряд ограничений, обусловленных как её архитектурой, так и природой финансовых данных.

Основное допущение логистической регрессии — линейная зависимость между признаками и логарифмом шансов. Хотя этого оказалось достаточно для данной задачи, сложные нелинейные паттерны, присутствующие на рынке, могут оставаться невыявленными. В будущем представляет интерес эксперимент с обобщенными аддитивными моделями (GAM) или неглубокими нейронными сетями, которые способны улавливать нелинейности, сохраняя при этом относительную интерпретируемость.

Зависимость от качества признаков. Эффективность модели критически зависит от тщательного отбора и инженерии признаков. В данной работе использовался классический технический анализ. Перспективным направлением является расширение признакового пространства за счёт включения макроэкономических индикаторов, данных новостного фона (с применением методов обработки естественного языка) или показателей волатильности (например, индекса VIX).

Стационарность рынка. Модель демонстрирует устойчивость в течение длительного периода, однако это не гарантирует её эффективность в условиях фундаментального изменения рыночных режимов (структурных сдвигов). Для повышения робастности необходима реализация механизма адаптивного переобучения (online learning) или детектора концептуального дрейфа, который будет сигнализировать о необходимости обновления модели.

Управление капиталом (Money Management). В работе использовалась фиксированная сумма на сделку. Значительного улучшения риск-менеджмента и доходности можно достичь, интегрировав в стратегию методы динамического управления капиталом, такие как Критерий Келли, который оптимизирует размер позиции в зависимости от прогнозируемой вероятности успеха.

5. Заключение

Проведённое исследование подтвердило, что переход от ансамблевого метода Random Forest Classifier к классической логистической регрессии обеспечивает значительное повышение эффективности и устойчивости системы прогнозирования цен акций. Логистическая регрессия продемонстрировала способность поддерживать баланс между точностью предсказаний и устойчивостью к переобучению, что особенно важно в задачах финансового моделирования, где избыточная адаптация к историческим данным может приводить к катастрофическим потерям на реальном рынке.

К числу основных преимуществ логистической регрессии относятся:

интерпретируемость параметров и возможность экономической трактовки коэффициентов;

устойчивость к переобучению благодаря регуляризации;

линейная прозрачность зависимости между признаками и откликом;

соответствие бинарной структуре задачи прогнозирования направления движения цен;

Результаты исследования имеют непосредственную практическую ценность для разработчиков алгоритмических торговых систем. Они демонстрируют, что приоритет при выборе модели должен смещаться от сложности и максимальной точности на исторических данных в сторону устойчивости, интерпретируемости и способности к обобщению. Предложенная модель на основе логистической регрессии представляет собой надежный и прозрачный инструмент, который может служить либо самостоятельным ядром торговой стратегии, либо robust-базисом, относительно которого имеет смысл тестировать более сложные и современные алгоритмы.

Таким образом, работа подтверждает старую истину в контексте машинного обучения для финансов: «Лучшее — враг хорошего». В условиях высокой неопределенности и шума надежная и понятная простая модель часто оказывается выигрышней сложного, но хрупкого «черного ящика».

В конечном итоге целью системы является заработок наибольшего количества денег, а не поиск наилучших значений метрик логистической регрессии, так как даже при большей точности чем нынешняя может сложится рыночная ситуация, при которой итоговый капитал будет ниже. Зная это, можем сравнить среднегодовой заработок модели с первым доступным финансовым вкладом в Сбербанке - оказывается, что вклад приносит на 0.05% меньше.

Также, изучив графики частоты сделок, можем увидеть, что система “держала” бумагу на протяжении 90% и более от общего количества торговых дней, верно предугадывая кризисы 2008, 2014, 2020 и 2024 годов.

Таким образом, даже относительно простые статистические модели при корректной настройке и надлежащем отборе признаков способны обеспечивать высокий уровень практической эффективности. Результаты исследования подтверждают, что не всегда сложность модели означает ее превосходство — напротив, в финансовых задачах более устойчивые и интерпретируемые методы нередко оказываются предпочтительными.

Список используемых источников

1. Hosmer, David W., and Stanley Lemeshow. Applied Logistic Regression. Third Edition. Hoboken, New Jersey: John Wiley & Sons, Inc., 2013. - 528 p.
2. Breiman, Leo. Random Forests. In Machine Learning, Volume 45, Issue 1, Springer, 2001. - pp. 5-32.
3. Zhang, Jie, and Wang, Yong. Financial Forecasting Using Logistic Regression and Ensemble Models. In IEEE Access, Vol. 8, 2020. - pp. 23010-23021.
4. Murphy, Kevin P. Machine Learning: A Probabilistic Perspective. Second Edition. Cambridge, Massachusetts: The MIT Press, 2022. - 1263 p.
5. Bishop, Christopher M. Pattern Recognition and Machine Learning. New York: Springer Science+Business Media, LLC, 2011. - 738 p.
6. James, Gareth; Witten, Daniela; Hastie, Trevor; Tibshirani, Robert. An Introduction to Statistical Learning with Applications in Python. New York: Springer, 2023. - 614 p.
7. Friedman, Jerome H.; Hastie, Trevor; Tibshirani, Robert. The Elements of Statistical Learning: Data Mining, Inference, and Prediction. Second Edition. New York: Springer, 2009. - 745 p.
8. Shalev-Shwartz, Shai; Ben-David, Shai. Understanding Machine Learning: From Theory to Algorithms. Cambridge: Cambridge University Press, 2014. - 410 p.
9. Kuhn, Max; Johnson, Kjell. Applied Predictive Modeling. New York: Springer, 2013. - 600 p.
10. Géron, Aurélien. Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems. Third Edition. Sebastopol, California: O’Reilly Media, 2023. - 1120 p.
11. Russell, Stuart J.; Norvig, Peter. Artificial Intelligence: A Modern Approach. Fourth Edition. Hoboken, New Jersey: Pearson Education, 2021. - 1168 p.
12. Goodfellow, Ian; Bengio, Yoshua; Courville, Aaron. Deep Learning. Cambridge, Massachusetts: The MIT Press, 2016. - 800 p.
13. Aggarwal, Charu C. Machine Learning for Text. New York: Springer, 2018. - 471 p.
14. Hastie, Trevor; Tibshirani, Robert; Friedman, Jerome. Statistical Learning with Sparsity: The Lasso and Generalizations. Boca Raton, Florida: CRC Press, 2015. - 367 p.
15. Ripley, Brian D. Pattern Recognition and Neural Networks. Cambridge: Cambridge University Press, 1996. - 416 p.
16. Efron, Bradley; Hastie, Trevor. Computer Age Statistical Inference: Algorithms, Evidence, and

Data Science. Cambridge: Cambridge University Press, 2016. - 475 p.

17. Kuhn, Max; Silge, Julia. Tidy Modeling with R: A Framework for Modeling in the Tidyverse. Sebastopol, California: O'Reilly Media, 2022. - 528 p.
18. Alpaydin, Ethem. Introduction to Machine Learning. Fourth Edition. Cambridge, Massachusetts: The MIT Press, 2020. - 640 p.
19. Zhang, Harry; Ma, Zhihua. Ensemble Machine Learning: Methods and Applications. New York: Springer, 2012. - 360 p.
20. Mitchell, Tom M. Machine Learning. International Edition. New York: McGraw-Hill Education, 1997. - 432 p.

Разработка стратегии бренд-маркетинга продукта

Кузнецова Е. В.
Уфимский университет науки и технологий
Институт экономики управления и бизнеса
Уфа, Россия
e-mail: 5kuznetsova@mail.ru

Аннотация[1]

Актуализируется концепция бренд-маркетинга продуктов компаний. Дается определение бренд-маркетинга продукта как долгосрочного маркетингового плана компании по формированию, поддержанию и развитию уникального по отношению к конкурентам бренда продукта, поддерживаемого приверженными клиентами. Рассматривается классификация стратегий бренд-маркетинга на примере продуктов известных компаний: по уровню бренда, стратегическим целям бренда, целевой аудитории бренда, позиционированию бренда, архитектуре бренда, взаимодействию бренда с рынком. Особый акцент в статье делается на алгоритм разработки стратегии бренд-маркетинга-продукта. Первый этап включает в себя отраслевой анализ рынка, конкурентов и потребителей. Второй этап описывает достоинства и конкурентные преимущества продукта, формирует уникальное торговое предложение продукта, выявляет нишу марки. Третий этап позволяет определить целевую аудиторию, провести сегментацию, составить потрет потребителя. Четвертый этап идентифицирует характер и ценности бренда с позиции архетипов, уточняет модель позиционирования бренда продукта. Пятый этап формирует айдентику бренда, его фирменный стиль и визуальный неповторимый образ. Шестой этап отмечает важность маркетинговых коммуникаций бренда. На седьмом этапе проводится пилотное тестирование, оптимизация, реализация и мониторинг стратегии бренд-маркетинга продукта. Итоговый восьмой этап оценивает эффективность стратегии бренда продукта.

Ключевые слова: стратегия; бренд-маркетинг продукта; целевая аудитория; позиционирование; ценности бренда; айдентика бренда; ребрендинг.

1. Введение

В современных реалиях острой конкуренции выигрывает бренд, управлением которым опирается на системный подход и стратегическое маркетинговое мышление [1]. Согласно исследованиям, компании, имеющие структурированные бренд-стратегии на 20 % рентабельнее конкурентов. Если бренд-стратегия не определена, компании сталкиваются с проблемами, которые могут привести к оттоку клиентов и сворачиванию бизнеса.

Бренд является важнейшей составляющей маркетинга [2, 3, 4, 5]. На наш взгляд, целесообразно использовать объединенную дефиницию бренд-маркетинг. Бренд-маркетинг – стратегическое маркетинговое управление брендом продукта, компании, территории или личности с целью создания и поддержания уникального образа бренда в сознании потребителей посредством имиджевого и репутационного маркетинга.

Под стратегией бренд-маркетинга продукта автором понимается долгосрочный маркетинговый план компании по формированию, поддержанию и развитию уникального по отношению к конкурентам бренда продукта, поддерживаемого лояльными потребителями.

2. Цель

Цель исследования: разработать алгоритм стратегии бренд-маркетинга продукта.

Задачи исследования:

- уточнить понятие «стратегии бренд-маркетинга продукта»;
- рассмотреть классификацию стратегий бренд-маркетинга продукта;
- сформулировать и описать этапы разработки стратегии бренд-маркетинга продукта.

Научная новизна исследования:

- дано определение стратегии бренд-маркетинга продукта. Уточнена классификация стратегий

[1] Труды XI Международной научной конференции "Информационные технологии интеллектуальной поддержки принятия решений", 11-13 ноября, Уфа-Чандигарх-Фандербейлпарк, 2025

бренд-маркетинга продукта на примере компаний;

- сформулирован алгоритм разработки стратегии бренд-маркетинга продукта.

3. Методы

В качестве методологии исследования используется системный подход; в качестве методов – метод дедукции, метод анализа и синтеза, методы классического и стратегического маркетинга, инструменты брендинга.

4. Результаты

Стратегия бренд-маркетинга выстраивается в зависимости от таких критериев как: уровни бренда, цели стратегии, целевая аудитория, позиционирование, архитектура бренда, взаимодействие с рынком (рис. 1).



Рис. 1. Виды стратегий бренд-маркетинга продукта

Рассмотрим классификацию стратегий бренд-маркетинга продукта на примере компаний.

Уровень бренда.

1. Корпоративная стратегия. Компания Samsung создает единый имидж бренда качественной техники.

2. Стратегия продуктового бренда. Компания Tesla производит электромобили высокого качества с инновационными технологиями.

3. Стратегия линейки продуктов. Компания Mars обновляет рецепты сладостей, следствием чего появляются новые вкусы конфет и батончиков.

4. Стратегия суббренда. Суббренд «Яндекс.Еда» фокусируется исключительно на доставке еды, не пересекаясь с другими суббрендами компании «Яндекс».

Цели стратегии бренда.

1. Стратегия создание бренда. Появление каждого бренда уходит в удивительную историю его создания. Например, корни бренда в сфере развлечений и мультипликации Walt Disney отражаются в творчестве Уолтера Диснея, который монетизировал свой талант и брал заказы на рисунки.

2. Стратегия сохранения и поддержания бренда продукта. Компания Nike в период пандемии коронавируса делает акцент на продажу спортивных костюмов вместо кроссовок.

3. Стратегия расширения бренда. Компания Amazon начинает с книжного магазина и расширяется до маркетплейса гигантских размеров.

4. Ребрендинг. «Тинькофф» в 2024 г. переименовывается в Т-Банк для привлечения молодежной аудитории в отрасли банковских услуг. Т-Банк реализует концепцию молодежного неформального тона общения с целевыми аудиториями, делая упор на гибкость, надежность и цифровые технологии.

Целевая аудитория бренда.

1. Стратегия массового рынка. Компания Procter&Gamble функционирует на массовом рынке, предлагая потребителям продукты своих многочисленных брендов.

2. Стратегия среднего рынка. Стратегия middle-сегмента выступает мостиком между массовым и премиум сегментами. 12Storrez – российский бренд одежды мидл-маркета, 12 раз в год обновляющий мини-коллекцию, использующий долговечные натуральные материалы и «принцип капсульного гардероба, когда все оттенки и фактуры гармонично дополняют друг друга».

3. Стратегия премиум рынка. Rolls-Royce производит роскошные автомобили люкс-сегмента.

4. Стратегия «нишевого» рынка. Бренд очков Warby Parker выбирает особенную нишу, не занятую «большими брендами»: предлагает гибкую модель продаж, когда потребители могут примерять очки дома перед покупкой.

Позиционирование бренда.

1. Позиционирование по стоимости. Стоимость может рассматриваться как низкая, так и высокая. Если цены на продукты высокие, то такие бренды, как правило, говорят о лучшем сочетании цены и качества. Если цены на продукты низкие, то бренды отмечают выгодное ценообразование для потребителей, предлагая тарифы эконом-сегмента, которые оправдываются отсутствием сервиса. Например, такое позиционирование поддерживается сетью магазинов FixPrice.

2. Позиционирование по качеству. Французский модный дом Hermes, позиционирующий себя

в люксовом сегменте, отличается изысканной ручной работой, высочайшим сервисом и высококачественными материалами при производстве уникальных изделий.

3. Позиционирование по инновациям. Бренд SpaceX «приближает» к потребителям космос, специализируется на инновационных спутниковых технологиях и космических путешествиях.

4. Позиционирование по устойчивому развитию в рамках повестки ESG. Бренд Natura Siberica декларирует экологический вектор: компания применяет в производстве своей косметики исключительно натуральные ингредиенты из сибирских трав, имеет сертификаты международных экостандартов.

5. Позиционирование по традициям. Бренд Your Yool – бренд одежды и украшений, выстраивающий коммуникации и создающий модели в татарских традициях [6].

Архитектура бренда.

1. Стратегия монобрендинга. Компания Apple, производящая разные продукты (MacBook, iPhone, iPad, AirPods, Apple Watch) и продвигающая их под своим единственным брендом.

2. Стратегия мультибрендинга. Примером мультибрендинговой стратегией можно назвать корпорацию LVMN, реализующую разнообразные бренды Dior, Louis Vuitton, Moët&Chandon на различных сегментах и рынках, от моды до алкогольных напитков.

3. Стратегия суббрендинга. Корпорация Сбер запускает суббренды «СберЗдоровье», «Купер», «Сберфуд», «Ситидрайв» для выхода на новые рынки, не связанные с финансово-кредитными операциями.

Взаимодействие с рынком.

1. Агрессивная стратегия бренда. Агрессивная стратегия часто сочетается с провокационным маркетингом. Бренд Apple выпускает скандальный ролик новой модели iPad, транслирующий пренебрежительное отношение к культуре и искусству, когда показывается, как под прессом уничтожаются картины, музыкальные инструменты, книги.

2. Оборонительная стратегия бренда. Компания Blockbuster занимает оборонительную стратегию по отношению к основному конкуренту Netflix, запуская платформу по просмотру фильмов онлайн.

3. Адаптивная стратегия бренда продукта. Адаптивную стратегию бренда можно назвать также стратегией ситуационного маркетинга. В период COVID-19 российская сеть фитнес-клубов World Class вместо традиционных тренировок в зале предлагает потребителям онлайн-фитнес, что позволяет не потерять приверженных клиентов и привлечь новичков за счет появления новых виртуальных услуг.

4. Стратегия ниши. Компания выбирает узкую нишу и фокусируется на ней. Например, на рынке технического оборудования для массажа лидирует бренд массажных кресел Yamaguchi.

Этапы разработки стратегии бренд-маркетинга продукта, на наш взгляд, следует свести в следующий алгоритм (рис. 2).



Рис. 2. Алгоритм разработки стратегии бренд-маркетинга продукта

1. Анализ рынка, конкурентов и потребителей.

Стратегический анализ рынка необходим, чтобы понять, какую отправную точку следует установить для выстраивания стратегии бренд-маркетинга продукта. Методами анализа рынка, конкурентов и потребителей являются: STEEP-анализ, матрица угроз, матрица возможностей, SWOT-анализ, SNW-анализ, отраслевой анализ рынка [7], модель пяти сил конкуренции М. Портера [8], карта стратегических групп, лепестковая диаграмма конкурентоспособности, прогнозирование рыночных тенденций и спроса, качественные маркетинговые исследования (наблюдение, фокус-группы, глубинные интервью), количественные маркетинговые исследования (опросы, анализ вторичных данных).

Итогом анализа является разработка портрета покупателя – собирательного образа клиента с учетом его демографических, психографических, поведенческих, географических характеристик и особенностей.

2. Описание продукта и ниши, формирование УТП.

Продукт целесообразно описать с точки зрения его уровней, согласно концепции Ф. Котлера: товар по

замыслу – решение проблем и болей потребителей; товар в реальном исполнении – физические, дизайнерские, функциональные характеристики продукта (бренд, свойства, внешнее оформление, качество, упаковка); товар с подкреплением – дополнительный сервис, получаемый клиентами (доставка, монтаж, производство под заказ, гарантия, консультации, ремонт, дробление или укрупнение партий продукта) [9].

Под «нишей» понимается часть рынка, на которой функционирует компания со своими продуктами. Поэтому изучается деятельность конкурентов, их ценности и целевые аудитории.

Под уникальным торговым предложением в бизнесе и маркетинге рассматривается описание преимуществ продукта (товара или услуги компании), которое привлекает потенциальных клиентов и отстраивает бренд от конкурентов. Также благодаря УТП, формируется лояльность потребителей и повышается коэффициент конверсии. Раскроем формулы успешных УТП:

«Продукт + уникальность», например, кофейня Capybara Coffee & Tea во Вьетнаме с капибарами, когда потребители наслаждаются не только изысканным кофе, но и получают удовольствие от фотосессий и контактной игры с милыми животными.

«Продукт + решение проблемы», например, шоколад из кэроба предлагается для тех, кому противопоказан сахар.

«Продукт + экспертность», например, косметика из натуральных ингредиентов, подтвержденная врачом-дерматологом высшей категории.

«Продукт + выгода», например, курс SMM-специалиста с последующим устройством на работу в digital-маркетинговое агентство.

«Продукт + ограничение по времени», например, Domino's Pizza: «Доставка пиццы в Москве за 30 минут».

3. Определение целевой аудитории.

Целевую аудиторию (ЦА) исследуют для того, чтобы понимать, как она будет реагировать на поведение бренда. Инструментами для анализа выступают CRM-системы (преимущественно в b2b-сегменте) и характеристика социальных сетей (в b2c-сегменте). ЦА можно классифицировать с точки зрения нескольких признаков. По степени вовлеченности в продукт выделяют «горячую», «теплую» и «холодную» ЦА. По отношению к продукту бренда – основную, вторичную и потенциальную ЦА.

Для работы с ЦА следует провести сегментацию несколькими методами, например, классическим методом, методами 5W М. Шеррингтона [10], VALS, Khramatrix, LTV, RFM, построить «лестницу Бена Ханта».

Также нами рекомендуется составить карту пути клиента [11]. Карта путешествия клиента (Customer Journey Map, CJM) – современный маркетинговый инструмент анализа опыта клиента, пользователя, сотрудника при взаимодействии с продуктом бренда [12]. Карта пути клиента позволяет оценить, где имеются страхи и боли у клиента, как он коммуницирует с брендом, какими ценностями он живет, что движет им при покупке и эксплуатации продукта. CJM решает несколько управленческих задач (рис. 3) [13].



Рис. 3. Задачи CJM

4. Обозначение характера и ценностей бренда, разработка позиционирования.

Каждый бренд формирует уникальные ценности – ключевые принципы бизнеса, транслируемые потребителям, партнерам, стейкхолдерам. Ценности связаны с самобытным характером бренда, демонстрирующим определенное поведение. Основоположник аналитической психологии, гениальный философ, психиатр, педагог и культуролог Карл Юнг в своей замечательной книге предлагает понятие «архетип» [14]. Через архетип выстраиваются ценности, эмоции, характер бренда.

Эксперты в области маркетинга и психологии М. Марк и К. Пирсон в труде «Герой и Бунтарь. Создание бренда с помощью архетипов» описывают алгоритм брендинга посредством архетипирования брендов [15]. Это позволяет создать и оценить неповторимый имидж, понять идеи, желания, ценности целевых аудиторий, привлечь и удержать новых клиентов, повысить узнаваемость бренда, связанную с оригинальным образом компании и ее продуктов, ассоциациями и персонажами [16]. В брендинге так же, как и в психологии, выводятся 12 архетипов. Ключевая идея заключается в том, что бренды подобно живым существам должны чувствовать потребителей и передавать эмоции через свое поведение, фразы, эмоциональные триггеры, tone of voice, сторителлинг, геймификацию [17].

5. Создание визуального образа и фирменного стиля бренда.

Визуальный образ и фирменный стиль бренда представляют собой айдентику – неповторимые визуальные элементы, узнаваемые потребителями. К подобным элементам относят логотип продукта или компании, символ, эмблема, буквенный знак,

цветовую палитру, шрифты, линии, иконки, персонажей, стиль иллюстраций, рекламные слоганы.

Логотипы, буквенные знаки, символы, эмблемы бренда являются самым узнаваем атрибутом бренда. Компании уделяют отдельное внимание визуалу марки. С течением времени логотипы претерпевают изменения, связанные преимущественно с изменением позиционирования, а также с учетом модных тенденций. В современных реалиях превалируют такие тренды как минимализм, асимметрия, 3D-эффекты, динамизм, адаптивность, персонализация, смешанные шрифты, винтажная и экспериментальная типографика, монохром и упрощенные палитры одновременно с яркими градиентами и сложными цветовыми переходами, линии, разделенные на сегменты, многослойность, «управляемый беспорядок», триггерные надписи, небрежные, «нарисованные от руки» логотипы.

Компании изменяют логотипы в рамках ребрендинга – маркетингового инструмента, создающего новый бренд продукта, ориентированного на новые сегменты рынка (рис. 4).

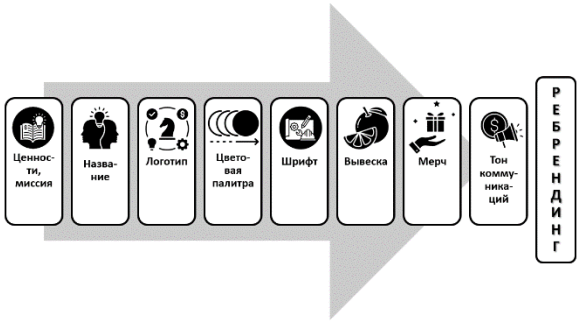


Рис. 4. Изменение компонентов при полном ребрендинге

6. Разработка коммуникационной стратегии бренда.

Коммуникационная стратегия бренда – долгосрочный план по продвижению компании и ее продуктов и коммуникациям бренда с целевыми аудиториями. Алгоритм коммуникационной стратегии бренда состоит из нескольких этапов, начиная с проведения ситуационного анализа, постановки целей коммуникаций, определения стратегического плана, выбора коммуникационных каналов и инструментов контента, разработки медиаплана и оценки эффективности маркетинговых коммуникаций бренда [18].

7. Тестирование, оптимизация, реализация и мониторинг стратегии бренда.

В процессе тестирования происходит апробация пилотной версии коммуникационной стратегии с целью оценки реакции потенциальных клиентов, выявления «слабых мест» и внесения корректив. Проводятся эксперименты в полевых и лабораторных условиях, приглашаются фокус группы для сбора мнений относительно привлекательности айдентики бренда. По итогам пилотного тестирования вносят стратегические изменения: выбирают другие сегменты рынка, модифицируют концепцию и позиционирование бренда, проводят A/B-тесты стратегий. В процессе реализации коммуникационной стратегии бренда компания осуществляет мониторинг взаимодействия с целевыми группами с учетом их клиентского опыта [19], отслеживает выполнение обещаний перед потребителями и взаимных ценностей. Оцениваются ключевые показатели эффективности: приверженность к бренду, узнаваемость бренда, восприятие бренда, финансовые показатели маркетинга и продаж.

8. Оценка эффективности стратегии бренда.

Анализ подходов к оценке эффективности брендинга основательно проводят в своей работе отечественные исследователи С.А. Старов, О.Н. Алканова. Ядром их методологии выступает «колесо контактов», с которым последовательно связаны четыре группы показателей [20].

Первая группа – «метрики восприятия»: распознавание бренда и осведомленность о нем, дифференциация бренда, релевантность бренда, доверие бренду, выделение бренда продукта среди конкурентов, понимание бренда, влияние бренда на принятие решений о покупке.

Вторая группа – «поведенческие метрики»: превосходство и объем продаж бренда, предпочтение и приверженность бренда, привлечение новых покупателей за счет известности бренда, удержание постоянных покупателей, выполнение обещания бренда, индекс лояльности клиентов (NPS), индекс удовлетворенности клиентов (CSI).

Третья группа – «рыночные метрики»: доля рынка бренда, география дистрибьюции бренда, индекс развития бренда (BDI), расширение бренда.

Четвертая группа – «финансовые метрики»: оценка текущей стоимости бренда, показатель отдачи инвестиций в бренд (ROBI).

5. Заключение

Можно резюмировать, что стратегия бренд-маркетинга продукта является важнейшим звеном маркетинговой стратегии компании. Бренд продукта – образ товаров и услуг в сознании потребителей. Бренд – это коммуникационная связь между компаний и целевыми аудиториями. Известный бренд продукта приводит к высоким продажам, лояльным потребителям, дифференцированной позиции на рынке, расширенным рыночным возможностям по использованию марки, повышенной ценности продукта. Поэтому маркетологам и специалистам по

брендингу целесообразно грамотно разрабатывать и реализовывать стратегии бренд-маркетинга продуктового портфеля компании, региона, государства [21].

Список используемых источников

1. Родионова Л. Н., Давлетшина С. М. Конкурентная разведка. – Уфа, 2024. – 310 с. (учебное пособие).
2. Аакер Д. Создание сильных брендов. – М.: Гребенников Бизнес Букс, 2003. – 440 с. (монография).
3. Гэд Т., Розенкрейц А. Создай свой бренд. – СПб.: Издательский дом Нева, 2004. – 191 с. (монография).
4. Домнин В. Н. Брендинг. – М.: Юрайт, 2025. – 555 с. (учебник и практикум для вузов).
5. Панкрухин А. П. Бренды и брендинг // Практический маркетинг. – 2011. – № 4 (170). – С. 4-15. (статья из журнала).
6. Бикметов Е. Ю., Кузнецова Е. В., Рувенный И. Я. Основы стратегического маркетингового управления. – Уфа: РИЦ УУНиТ, 2023. – 150 с. (учебное пособие).
7. Родионова Л. Н., Ибрагимова Л. Д. Формирование системы отраслевого управления в регионе // Актуальные вопросы экономической теории: развитие и применение в практике российских преобразований. Сборник научных трудов одной из площадок Уфимского гуманитарного научного форума. – Уфа, 2022. – С. 223-226. (материалы конференции).
8. Портер М. Конкурентная стратегия: Методика анализа отраслей и конкурентов; пер. с англ. – М.: Альпина Диджитал, 2016. – 456 с. (монография).
9. Котлер Ф. Основы маркетинга: краткий курс. – М Вильямс, 2019. – С. 233. (учебник).
10. Sherrington M. Added Value: The Alchemy of Brand-Led Growth, 2003. – 208 p. (монография).
11. Калбах Дж. Путь клиента: создаем ценность продуктов и услуг через карты путей, блупринты и другие инструменты визуализации. – М.: Манн, Иванов и Фербер, 2022. – С. 230. (монография).
12. Балахнин И. Найден более быстрый маршрут. Применение карт путешествия потребителя для повышения продаж и лояльности. Теперь и в B2B. – М.: Альпина PRO, 2022. – С. 36. (монография).
13. Карта пути клиента: от диагностики до целевого сценария. Гайд для продуктовых команд. – М.: Лаборатория Wonderfull, 2021. – С. 5. (монография).
14. Юнг К. Г. Архетипы и коллективное бессознательное. – М.: Издательство АСТ, 2023. –С. 8. (монография).
15. Марк М, Пирсон К. Герой и бунтарь. Создание бренда с помощью архетипов. – М.: Бомбора, 2023. – С. 18. (монография).
16. Касимова Э. Р., Кузнецова Е. В., Рувенный И. Я. Методы оценки имиджевого маркетинга // Бизнес. Образование. Право. – 2025. – № 2 (71). – С. 39-45. (статья из журнала).
17. Касимова Э. Р., Кузнецова Е. В., Рувенный И. Я. Геймификация как управленческая технология // Бизнес. Образование. Право. – 2020. – № 2 (51). – С. 171-175. (статья из журнала).
18. Кузнецова Е. В. Стратегия контент-маркетинга // Управление экономикой: методы, модели, технологии: Материалы XXI Международной научной конференции. – Уфа: УГАТУ, 2021. – С. 129. (материалы конференции).
19. Бикметов Е. Ю., Галиханова Л. И., Кузнецова Е. В., Рувенный И. Я. Использование клиентского опыта для принятия маркетинговых решений в здравоохранении // KANT. – 2023. – №3 (48). – С. 4-9. (статья из журнала).
20. Старов С. А., Алканова О. Н. Методология оценки эффективности брендинга // Вестник Санкт-Петербургского университета. – 2009. – Сер. 8. – Вып. 4. – С. 142. (статья из журнала).
21. Социальная политика государства в контексте долгосрочных национальных целей развития развития / Карелин И. Ю., Юлдыбаев Б. Р., Мусина Г. А. и др. – Уфа, 2024. – 410 с. (монография).

Перспективы реставрации фотографий с применением нейронных сетей

Богомолов В. П.
Уфимский университет науки и технологий
каф. АСУ
Уфа, Россия
e-mail: valentbog@yandex.ru

Аннотация[1]

В данной статье рассматриваются современные методы реставрации старых фотографий с использованием генеративных самообучающихся нейросетевых моделей. Проведен анализ различных подходов и сравнительная оценка их эффективности и практичности. Особенностью исследования стало тестирование моделей на снимках с серьёзными повреждениями. Результаты показали, что модели, работающие на локальном оборудовании, превосходят облачные сервисы по качеству восстановления и детальности изображений. Кроме того, модели, размещенные локально не требуют предварительной оплаты и не имеют ограничений на объем данных. Была выбрана нейросеть с максимальной эффективностью, которая могла восстанавливать детали и сохранять аутентичность изображений. В статье подробно обсуждаются как преимущества выбранного метода, так и его ограничения, рассматриваются перспективы дальнейшего развития технологий реставрации с использованием нейронных сетей.

Ключевые слова: нейронные сети, реставрация фотографий, stable diffusion, генеративные модели, фотореставрация, VAE, остаточные блоки, нейронные архитектуры.

1. Введение

В последнее время появилась возможность восстанавливать старые фотографии с помощью различных методов, основанных на нейронных сетях. Актуальность восстановления поврежденных исторических фотоснимков обусловлена:

1. Эти фотографии значительно утратили качество и целостность из-за воздействия времени, механических повреждений и внешних факторов.
2. Ценность подобных фотографий заключается в сохранении культурного наследия.

Современные методы реставрации, основанные на нейросетевых моделях, позволяют автоматизировать этот процесс с высокой степенью точности.

2. Перспективы реставрации фотографий

Цель исследования заключается в сравнительном анализе и оценке эффективности различных нейросетевых методов реставрации старых фотографий. Основная задача — выявить наиболее подходящие модели, параметры и подходы для восстановления изображений с разного рода повреждениями, используя как облачные, так и локальные вычислительные ресурсы.

Для исследования использовали монохроматический отпечаток фотографического изображения, отсканированный в растровый графический формат jpeg, который имел множественные механические повреждения и смещение фрагментов относительно друг друга.



Рис. 1. Исходный образец до манипуляций

Моделям, основанным исключительно на традиционных алгоритмах и использующим низкоуровневые признаки, сложно обнаруживать и эффективно исправлять дефекты таких фотографий. Более того, ни один из этих методов не рассматривает восстановление некоторых не структурированных дефектов, таких как выцветание

[1] Труды XI Международной научной конференции "Информационные технологии интеллектуальной поддержки принятия решений", 11-13 ноября, Уфа-Чандигарх-Фандербейлпарк, 2025

или низкое разрешение фотографии. С широким распространением нейронных сетей появились специализированные модели, предназначенные для реставрации изображений со смешанной деградацией.

Общие принципы работы генеративных самообучающихся моделей, построенных на принципе нейронных сетей, использующихся для реставрации изображений конечных отпечатков фотографического изображения со смешанной деградацией таковы: дефекты, такие как царапины и кляксы, сначала идентифицируются по низкоуровневым признакам, а затем закрашиваются путем заимствования текстур из окрестностей.

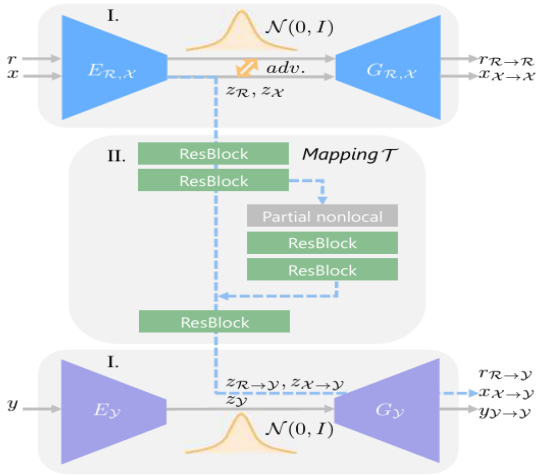


Рис. 2. Архитектура реставрирующей нейросетевой модели

Как видно на Рис. 2, в основе лежит архитектура из двух вариационных автокодировщиков (VAE): один обучается на реальных и синтетических искажённых фото, другой - на “чистых” (неповрежденных) изображениях (I на Рис. 2). Это позволяет перевести все искажённые изображения (реальные и синтетические) в латентное пространство (II на Рис. 2) (абстрактное, многомерное представление данных, в котором похожие точки расположены близко друг к другу, предоставляющее сжатое понимание данных). [1,2]

Вариационный автокодировщик (Variational Autoencoder — VAE) — генеративная модель, используемая в машинном обучении для генерации новых данных в виде вариаций входных данных, на которых она обучалась. Модель состоит из шифратора, который учится выделять важные скрытые переменные из обучающих данных, и дешифратора, который затем использует эти скрытые переменные для реконструкции входных данных. Вариационные автокодировщики не просто учатся кодировать исходные данные в более компактное внутреннее представление (сжимать) им восстанавливать из этого представления данные обратно (разжимать), но и моделировать их вероятностное распределение. Это позволяет им генерировать новые данные, похожие на исходные, а также восстанавливать данные с большей точностью.

В латентном пространстве происходит обучение преобразования, способного превращать искажённые изображения в чистые, опираясь на пары синтетических данных (искажённое — чистое). Residual blocks включают соединения, которые добавляют выход предыдущего слоя к выходу текущего.

Глобальная и локальная ветви — это ключевые компоненты архитектуры восстановления старых фотографий, каждая из которых предназначена для устранения определённого типа дефектов.

Глобальная ветвь (global branch):

Основная задача - восстановление структурированных дефектов: царапин, пятен, дыр, крупных загрязнений. В основе лежит partial nonlocal block - модуль, который учитывает глобальный контекст изображения. Это позволяет эффективно восстанавливать большие дефекты, используя информацию из удалённых, но неповреждённых областей снимка. В отличие от обычных nonlocal-блоков, partial nonlocal block использует маску повреждений, чтобы не привлекать для восстановления пиксели из уже испорченных областей. Таким образом, при заполнении дефектов используются только «здоровые» участки изображения, что обеспечивает структурную согласованность и минимизирует артефакты.

Локальная ветвь (Local branch):

Ориентирована на устранение неструктурированных дефектов: шум, размытость, выцветание, зернистость плёнки. Реализована с помощью нескольких residual blocks (остаточных блоков), которые анализируют и обрабатывают небольшие, ограниченные участки изображения и восстанавливают детали, используя соседние пиксели. Локальная ветвь хорошо справляется с дефектами, равномерно распределёнными по всему изображению, где не требуется учитывать глобальную структуру. Эти две ветви объединяются в латентном пространстве, что позволяет эффективно восстанавливать фотографии с множественными смешанными дефектами. Одновременно можно устранять как структурированные, так и неструктурированные дефекты.

В ходе исследования облачных реставрационных нейросетей было выявлено, что ни одна из них не смогла выполнить качественную реставрацию данного изображения. Результаты оказались неудовлетворительными для восстановления этого изображения. Также большинство из них требовало предварительной оплаты.

Методология исследования основана на применении генеративных самообучающихся моделей, построенных на принципах нейронных сетей, в частности архитектуры вариационных автокодировщиков (VAE) и модели Stable Diffusion.

Исследование включало следующие этапы:

• Выбор и подготовка исходных данных — монохроматический JPEG-файл с множественными механическими повреждениями.

• Использование модели SD Webui Old Photo Restoration, основанной на Stable Diffusion, запущенной на персональном компьютере с конфигурацией: Windows 11, процессор Intel i7-13700KF, видеокарта MSI GeForce RTX 4080 16GB, 32 ГБ оперативной памяти.

• Подбор параметров обработки: выбор включения реставрации царапин (Process Scratch), лиц (Face Restore), разрешения (High Resolution) и аппаратного ресурса (Use CPU). Эти параметры варьировались для оценки влияния на качество восстановления.

• Детальный разбор архитектуры нейросети с описанием двух ключевых ветвей: глобальной для устранения структурированных дефектов (царапины, пятна, дыры) с использованием partial nonlocal блока, и локальной — для неструктурированных дефектов (шум, размытость, выцветание), реализованной через residual blocks.

• Обоснование выбора параметров, их влияние на качество итогового изображения и объяснение отклонений от стандартных настроек, например, необходимость использовать CPU при ошибках CUDA Out of memory.

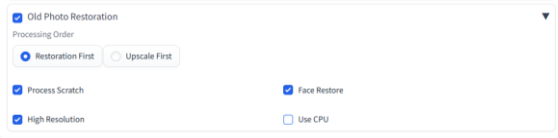


Рис. 3 Интерфейс локальной нейросетевой модели SD Webui Old Photo Restoration

Элементы графического пользовательского интерфейса, представленные на Рис. 3, принадлежат к типу “включить-выключить” и имеют следующий функционал:

Process Scratch предоставляет выбор либо реставрировать царапины и иные схожие с ними механические повреждения на растровом изображении конечного отпечатка фотографического изображения, либо не реставрировать.

Face Restore предоставляет выбор либо реставрировать лица на данном изображении, либо не реставрировать.

High Resolution предоставляет более точную настройку процесса обработки изображения исходя из разрешающей способности исходного растрового изображения.

Use CPU предоставляет выбор использовать процессор или графический ускоритель для произведения вычислений. Данный параметр следует включить при возникновении программной ошибки с идентификатором “CUDA Out of memory”, возникающей при обнаружении программой несоответствия фактически доступной памяти графического ускорителя, минимально требуемой для корректной работы. Также данный параметр следует включить при том условии, что графический процессор был разработан производителем, отличным от Nvidia Corporation.

Увеличение проводилось в 4 раза.



Рис. 4. Отреставрированное изображение с параметрами согласно Рис. 3

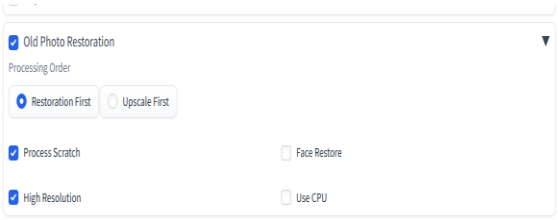


Рис. 5 Альтернативный вариант параметров



Рис. 6. Отреставрированное изображение с параметрами согласно Рис. 5

Эксперимент проведён на реальном растровом изображении с выраженными механическими повреждениями и смещением фрагментов.

• Техника проведения тестирования: изображение обрабатывалось с использованием разных конфигураций параметров интерфейса SD Webui Old Photo Restoration, фиксировались результаты реставрации, которые затем визуально

анализировались на наличие артефактов, степень восстановления лиц и устранение дефектов.

• Источник данных: оригинальный отсканированный монохроматический фотоснимок, сохранённый в формате JPEG, взятый из сети “Интернет”.

• Критерии оценки результатов включали полноту восстановления дефектов, качество деталей лиц, отсутствие визуальных искажений и артефактов, а также сравнение с результатами облачных сервисов.

Исходя из Рис. 4 и Рис. 6, можно сделать вывод, что параметры, изображенные на Рис. 3, являются предпочтительными в сравнении с параметрами на Рис. 5. Результат реставрации изображения, отображённый на Рис. 4, не представляет полной реставрации изображения оцифрованного растрового монохроматического отпечатка фотографического изображения с множественными механическими повреждениями, однако большинство повреждений было реставрировано, лица также отреставрированы, однако подверглись заметным искажениям в процессе реставрации ввиду несовершенства использованной в процессе реставрации нейросетевой модели.

3. Заключение

Локальные нейросетевые модели, основанные на архитектуре Stable Diffusion, превосходят оба варианта — облачные сервисы и другие онлайн-решения — по качеству восстановления и доступности. Такие модели позволяют эффективно устранять как структурированные, так и неструктурированные дефекты, обеспечивая высокую детализацию и точность, особенно в восстановлении лиц и мелких деталей. Они также позволяют более гибко настраивать параметры под конкретные изображения и не нуждаются в постоянной оплате облачных услуг, что делает их более экономически выгодными в долгосрочной перспективе.

Ключевые плюсы использования локальных моделей:

• Высокое качество восстановления изображений

• Возможность адаптации под конкретные задачи

• Отсутствие необходимости платить за облачные ресурсы

Недостатки:

• Для работы на локальном оборудовании требуются значительные вычислительные ресурсы, что предъявляет высокие технические требования к производительности устройств.

• Могут возникать искажения при реконструкции лиц, особенно если обучающие данные имеют недостатки или неполноту, что снижает точность восстановления.

• Имеются ограничения при восстановлении участков с серьёзными повреждениями.

Сравнение с другими методами показало, что локальные модели обеспечивают минимальные артефакты и наилучшее качество, особенно при работе с сложными повреждениями. В то же время, облачные сервисы часто уступают по скорости и уровню детализации, что особенно заметно при обработке архивных или высоко поврежденных материалов.

Исследование подтверждает, что применение нейросетей, реализованных на базе Stable Diffusion, значительно упрощает и ускоряет процесс реставрации старых фотографий, что делает такие технологии привлекательными для музеев, архивов и частных коллекционеров. Для массового применения предлагается создать пакетный режим работы, который позволит массово восстанавливать коллекции фотографий с минимальным участием человека.

Список используемых источников

1. Ziyu Wan, Bo Zhang, Dongdong Chen, Pan Zhang, Dong Chen, Jing Liao, Fang Wen (2020) Bringing Old Photos Back to Life. https://arxiv.org/abs/2004.09484
2. Zhangmozhe (2021) Old photo restoration (Официальная реализация на PyTorch). https://github.com/microsoft/Bringing-Old-Photos-Back-to-Life

Программное обеспечение для обнаружения и классификации электрических отказов

Тен Э. Р.
ООО «ЭТК»
Уфа, Россия
e-mail: eriksonten@mail.ru

Абдрахманова Р. П.
Уфимский университет науки
и технологий
Уфа, Россия
e-mail: vmk_rimma@mail.ru

Аннотация[1]

В статье рассматривается актуальная проблема своевременного выявления электрических отказов электродвигателей. Обосновывается значимость оперативного диагностирования неисправностей (обрыв фазы, короткое замыкание на корпус, межвитковое короткое замыкание), позволяющих предотвратить экономические потери, экологические риски и угрозы безопасности.

Цель исследования — повышение эффективности диагностики отказов электродвигателей и разработка специализированного программного обеспечения для их выявления и классификации. В работе анализируются существующие методы обнаружения неисправностей, отличающиеся либо недостаточной информативностью, либо высокой ресурсозатратностью.

Предложен комплексный подход к диагностике на основе анализа трёхфазного тока/напряжения с применением:

– преобразования Кларка (переход к системе координат αβγ);

– преобразования Парка (переход к вращающейся системе dq);

– дискретного преобразования Фурье и быстрого преобразования Фурье (анализ гармонических составляющих);

– расчёта амплитудного и фазового спектров.

Разработан алгоритм обнаружения конкретных типов отказов по характерным изменениям спектральных характеристик. Создан программный модуль с интерфейсом для взаимодействия с осциллографом, сбора, обработки и визуализации данных в реальном времени.

Результаты исследования позволяют существенно сократить время и затраты на диагностику электродвигателей, повысив надёжность их эксплуатации.

Ключевые слова: отказы электродвигателей; преобразование Кларка; преобразование Парка; быстрое преобразование Фурье; фазовый спектр; амплитудный спектр.

1. Введение

Электродвигатели используются в различных сферах, начиная с бытовых приборов, заканчивая промышленными станками. Поломка двигателя приводит к значительным экономическим потерям, кроме того, может вызвать негативные последствия для окружающей среды и безопасности людей. Поэтому своевременное выявление неисправности является важной задачей. В статье рассматриваются наиболее частые и значимые электрические отказы: обрыв фазы, короткое замыкание на корпус, межвитковое короткое замыкание.

Многие существующие программы либо оповещают о наличии проблемы без классификации, либо требуют изъятия двигателя из устройства и проводят полный анализ двигателя, используя большое количество датчиков и вспомогательных устройств. Второй способ сопровождается значительными финансовыми и временными затратами. Поэтому необходимо наличие портативного программного обеспечения, которое позволит в кратчайшие сроки провести проверку.

Объектом исследования являются признаки и классификация отказов электродвигателя. Предметом исследования - модели и алгоритмы классификации отказов электродвигателя на основе анализа 3-фазного тока/напряжения с помощью координатного преобразования и дискретного

[1] Труды XI Международной научной конференции "Информационные технологии интеллектуальной поддержки принятия решений", 11-13 ноября, Уфа-Чандигарх-Фандербейлпарк, 2025

преобразования Фурье, амплитудного и фазового спектров. [1,2,4]

Целью работы является повышение эффективности выявления отказов электродвигателя и разработка программного обеспечения для их выявления и классификации.

Для достижения цели поставлены следующие задачи:

1. Анализ методов и алгоритмов выявления отказов электродвигателя.

2. Разработка алгоритма обработки 3-фазного тока/напряжения электродвигателя с помощью координатного преобразования, дискретного преобразования Фурье, амплитудного и фазового спектров.

3. Программная реализация модуля анализа состояния электродвигателя и оценка его эффективности.

2. Теоретическая часть

Функция обнаружения обрыва обмотки фазы вызывается 3 раза на каждую фазу. В программе реализован счетчик, который увеличивается на единицу, если считанное значение тока находится в определенном промежутке. При достижении значения 150 счетчик сигнализирует о наличии отказа. Подобный способ позволяет избежать ложных срабатываний, однако уменьшает достоверность.

Для обнаружения коротких замыканий необходимо воспользоваться преобразованиями Кларка и Парка. В зависимости от поведения γ-составляющей определяется вид поломки. У короткого замыкания обмотки фазы на корпус амплитуда повышается, у межвиткового короткого замыкания значение γ уменьшается.

2.1. Электрические отказы

Обрыв обмотки фазы случается чаще чем другие отказы. При этом электродвигатель начинает быстро терять мощность и сильно нагревается.

Короткое замыкание происходит, когда обмотка касается корпуса или при соединении разных витков катушек. В таких случаях электродвигатель потребляет слишком большой ток, а сопротивление обмоток уменьшается. [6]

2.2. Преобразование Кларка

Считанные значения трех фаз тока/напряжения образуют входной вектор. Его нужно перемножить на матрицу (1) для получения выходных значений α, β, γ.

$$\sqrt{\frac{2}{3}} \cdot \begin{bmatrix} 1 & -\frac{1}{2} & -\frac{1}{2} \\ 0 & \frac{\sqrt{3}}{2} & -\frac{\sqrt{3}}{2} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \quad (1)$$

На рисунке 1 изображены компоненты 3-фазной системы координат.

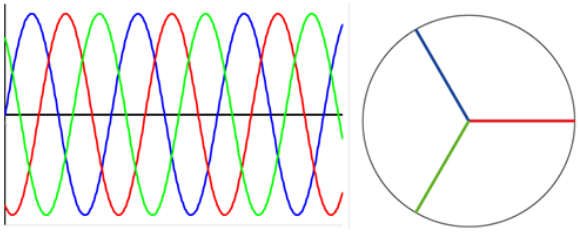


Рис. 1. Система координат abc

На рисунке 2 изображены компоненты 2-фазной системы координат, полученные после преобразования Кларка.

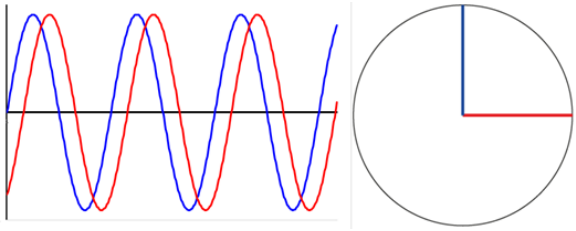


Рис. 2. Система координат αβγ

2.3. Преобразование Парка

Векторы α, β, γ нужно перемножить на матрицу (2). Таким образом будут получены компоненты dq вращающейся системы координат.

$$\begin{bmatrix} \cos(\theta) & \sin(\theta) & 0 \\ -\sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad (2)$$

где θ - мгновенный угол произвольной частоты.

На рисунке 3 изображен результат преобразования Парка.

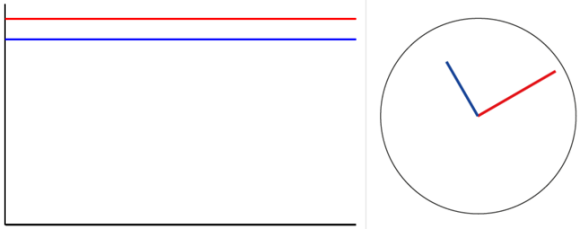


Рис. 3. Система координат dq

Наличие прямых линий говорит об отсутствии каких-либо отказов.

2.4. Преобразование Фурье

Дискретное преобразование Фурье позволяет получить амплитудный и фазовый спектры. Для выполнения ДПФ используется формула (3).

$$X_k = \sum_{n=0}^{N-1} x_n(\cos(\frac{2\pi kn}{N}) - i \cdot \sin(\frac{2\pi kn}{N})), \quad (3)$$

где:

– N – количество дискретных значений на входе;

– xn, n=0,…,N-1 – измеренные значения сигнала;

– Xk, k=0,…,N-1 – комплексные числа на выходе;

– k, k=0,…,N-1 – индекс частоты.

Однако в программе используется большое количество данных, и, поэтому, для подобного вычисления требуется много времени. Решением проблемы является быстрое преобразование Фурье (Fast Fourier transform). Этот метод сократит количество вычислений, но для него необходимо определенное количество входных значений (оно должно равняться степени 2). [3]

Алгоритм БПФ:

1. Входные данные делятся на две группы. В первую идут элементы с четным индексом, во вторую с нечетным;

2. Этап 1 повторяется до тех пор, пока в группе не останется два элемента;

3. Расчет новых значений с использованием ДПФ;

4. Группы объединятся путем переплетения элементов (этап 1 в обратном порядке).

Пример БПФ изображен на рисунке 4.

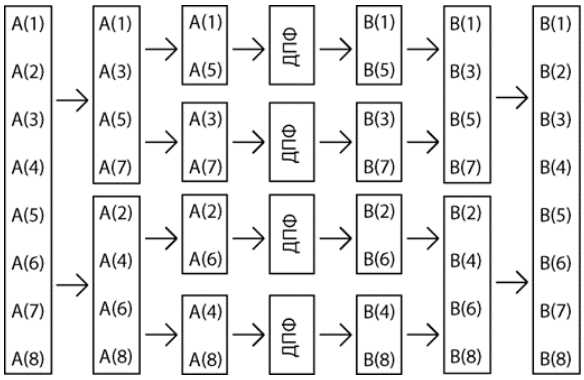


Рис. 4. Принцип быстрого преобразования Фурье

2.5. Амплитудный спектр

Вычислить аплитудный спектр можно по формуле (4).

$$A_k = \sqrt{Re(X_k)^2 + Im(X_k)^2}, \quad (4)$$

где:

– k – индекс значения,

– Re(Xk) – действительная часть значения, полученного после преобразования Фурье,

– Im(Xk) – мнимая часть значения, полученного после преобразования Фурье.

Стоить заметить, что для анализа не требуется большое количество значений, поэтому можно использовать около 5% начальных значений амплитудного спектра. [5]

2.6. Фазовый спектр

Вычислить фазовый спектр можно по формуле (5).

$$\varphi_k = arctg\frac{Im(X_k)}{Re(X_k)} \quad (5)$$

2.7. Интерфейс программы

На рисунке 5 изображен интерфейс разрабатываемой программы.

1) Поле выбора осциллографа;

2) Поле вывода сообщения;

3) Поле вывода графиков;

4) Выбор способа подключения и кнопка поиска устройств;

5) Кнопки, необходимые для получения данных:

6) Поле выбора фазы и кнопки для работы с сохраненными файлами;

7) Индикаторы обнаруженных отказов.

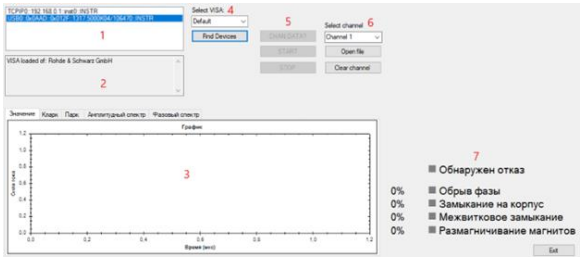


Рис. 5. Интерфейс

При нажатии кнопки «CHAN DATA?» запускается функция, создающая сообщения для осциллографа.

При нажатии кнопки START запускается таймер. Таймер отправляет запрос на получение данных с осциллографа. После задержки он будет вызывать саму функцию, тем самым запустится циклический опрос и обработка сигнала в реальном времени. Для остановки таймера используется кнопка STOP и прекращается передача данных.

Для отправки запроса на осциллограф используются функции готовой библиотеки visa32. viOpen занимает канал связи осциллографа, тем самым начинает передачу информации; возвращает значение статуса работы осциллографа. Если оно превышает заданный порог, то будет выведен код ошибки.

Для функции Query нужно создать SCPI команду, которая будет отправлена осциллографу. Команда записывается в переменную chan_comm и состоит из постоянной части «CHAN» (запрос к каналу данных) и переменной части, полученной из

функции btnCHANDATA_Click, которая определяет номер канала данных и запрос данных. Она возвращает дискретные значения, считанные осциллографом, которые разделены между собой точкой с запятой. Возвращаемые значения представляют собой строку, а значит впоследствии она должна быть преобразована в массив числовых данных.

По окончании передачи вызывается функция viClose для освобождения канала связи осциллографа. Собранные данные сохраняются в текстовом формате для возможности дальнейшего изучения сигналов и просмотра истории, в дальнейшем они также могут быть использованы для обучения нейронных сетей.

Так как пути к файлам будут использоваться многократно, они будут записаны в отдельную переменную. При сохранении файлов им присваиваются разные имена, отличающиеся номерами.

Функция открытия необходимого файла запускается нажатием на кнопку Open file. Данные файла записываются в массив, который впоследствии используется для анализа. Также необходимо указать номер канала.

3. Работа программы

На рисунках 6-8 изображены примеры электрических отказов.

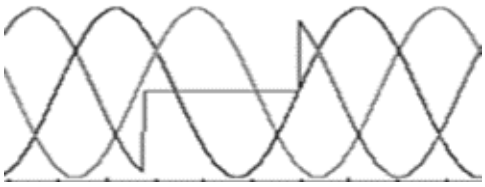


Рис. 6. Обрыв обмотки фазы

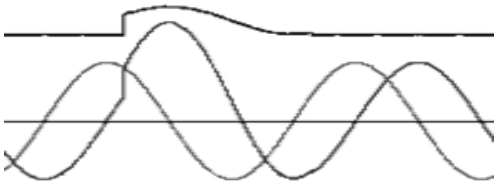


Рис. 7. Короткое замыкание обмотки фазы на корпус

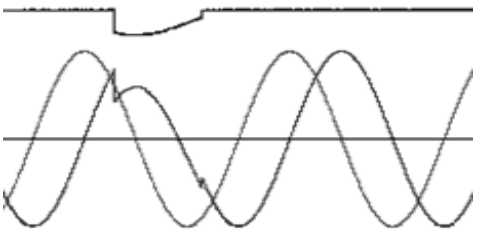


Рис. 8. Межвитковое короткое замыкание

На рисунках 9-14 представлены электрические отказы в амплитудном и фазовом спектрах.

Обрыв обмотки сопровождается резким ростом второй гармоники на амплитудном спектре и увеличенной частотой на фазовом спектре.

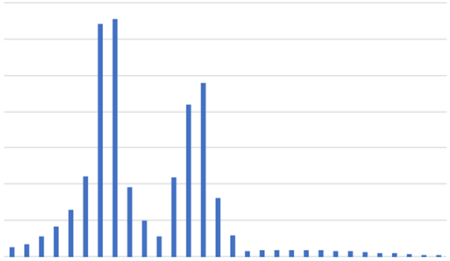


Рис. 9. Амплитудный спектр при обрыве обмотки фазы

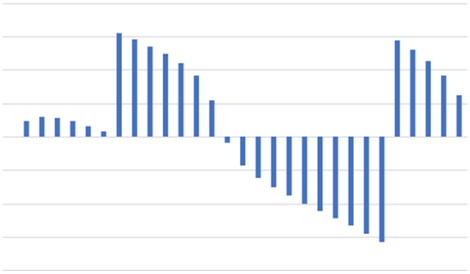


Рис. 10. Фазовый спектр при обрыве обмотки фазы

Короткое замыкание на корпус сопровождается небольшим ростом второй гармоники и изменением фазового спектра.

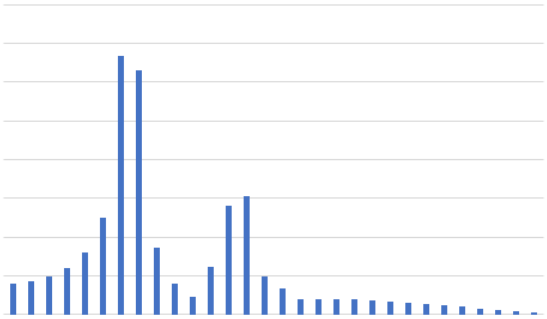


Рис. 11. Амплитудный спектр при замыкании на корпус

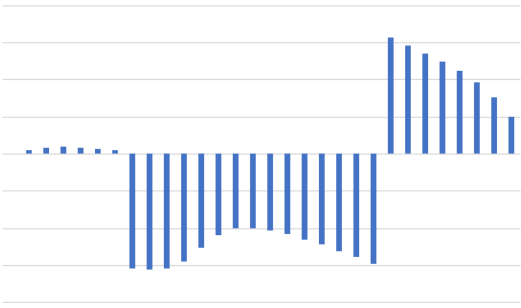


Рис. 12. Фазовый спектр при замыкании на корпус

Межвитковое короткое замыкание сопровождается уменьшением второй гармоники и изменением фазового спектра

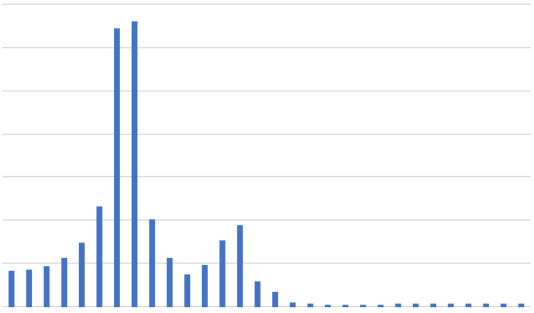


Рис. 13. Амплитудный спектр при межвитковом замыкании

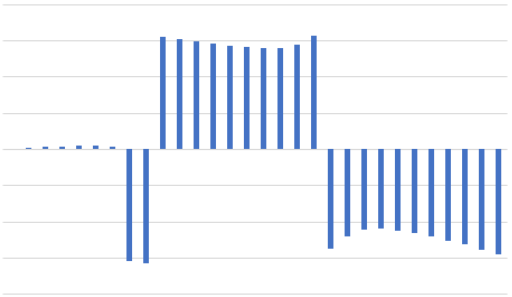


Рис. 14. Фазовый спектр при межвитковом замыкании

На рисунках 15-19 изображены блок-схемы алгоритмов.

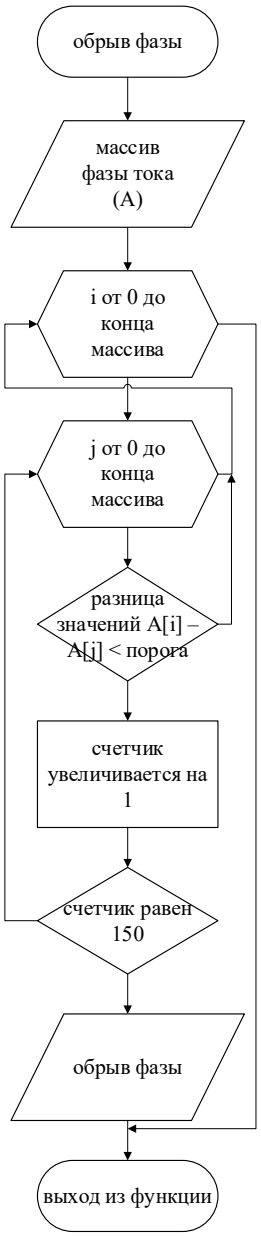


Рис. 15. Блок схема обрыва фазы

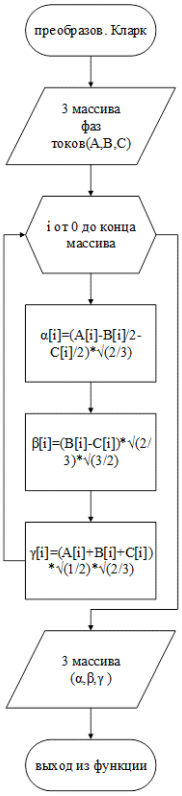


Рис. 16. Блок-схема преобразования Кларка

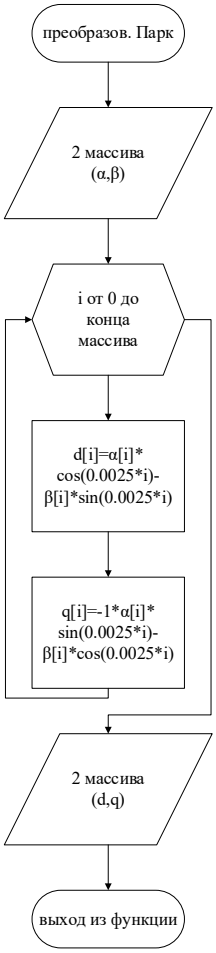


Рис. 17. Блок-схема преобразования Парка

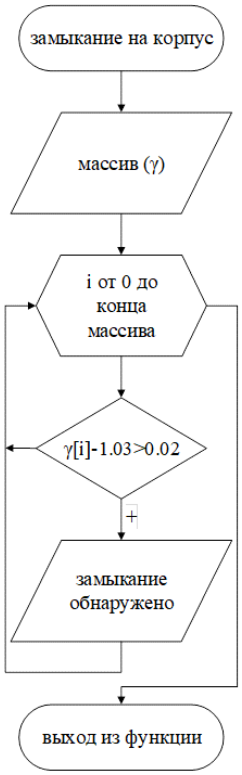


Рис. 18. Блок-схема замыкания на корпус

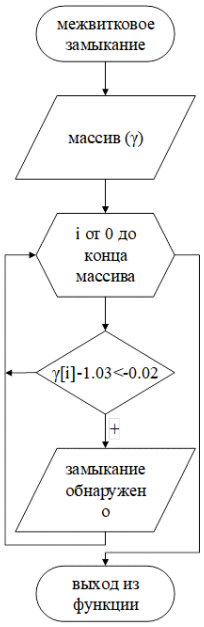


Рис. 19. Блок-схема межвиткового замыкания

4. Заключение

1. Проведены исследования электрических отказов электродвигателя
2. Разработано программное обеспечение для выявления неисправностей. Преимуществами являются: возможность его использования в реальном времени в автоматизированном режиме; работа в виде десктопного приложения; портативность.

3. Получены результаты с высокой точностью обнаружения: обрыв обмотки фазы - 98,8%, короткое замыкание обмотки фазы на корпус - 98,7%, межвитковое короткое замыкание – 100%, размагничивание магнитов – 70,4%.

Список используемых источников

1. Сидельников, Л.Г. Обзор методов контроля технического состояния асинхронных двигателей в процессе эксплуатации / Л.Г. Сидельников, Д.О. Афанасьев // Архитектура и строительство России. – 2013. – № 7. – 137 с. (статья)
2. Гаврилов, Р.С. Управление синхронными машинами с постоянными магнитами: учебное пособие / Р.С. Гаврилов, Ю.Н. Мустафаев // Балт. гос. техн. ун–т. – СПб. – 2019. – 78 с. (Учебное пособие)
3. Чигирева, О.Ю. Ряды Фурье. Преобразование Фурье // МГТУ им. Н.Э. Баумана. – 2010. – 51 с. (учебно-методическое пособие)
4. Ерохин, Е.Ю. Способ контроля обрыва фазы в трехфазной сети и устройство для его осуществления / Е.Ю. Ерохин, Е.Г. Токмаков, М.А. Ваткина // Открытое акционерное общество «Всероссийский научно–исследовательский проектно–конструкторский и технологический институт релестроения с опытным производством». – 2008. – 10 с. (патент)
5. Савицкий Р.В. Расчет амплитудного спектра по длительностям периодов сигнала/ Р.В. Савицкий, В.Л. Ким// Интернет-журнал «Науковедение», №5,2013 (статья)
6. Тен Э.Р. Алгоритмы выявления видов отказов электродвигателя / Э.Р. Тен, Р.П. Абдрахманова // Мавлютовские чтения. – Том 5. – 2024 – 991 с. (151 – 156) (статья)

Методологические аспекты применения ИИ при проектировании информационных систем

Набатов А. Н.
Уфимский университет науки
и технологий
Институт информатики, математики
и робототехники
каф. АСУ
Уфа, Россия
e-mail: nbtv@yandex.ru

Веденяпин И. Э.
Уфимский университет науки
и технологий
Институт информатики, математики
и робототехники
каф. ВМиК
Уфа, Россия
e-mail: viggi62@yandex.ru

Аннотация[1]

С ростом количества и усложнением информационных систем философские и методологические вопросы их создания и эксплуатации становятся все более актуальными. Важным моментом этого процесса является выработка и принятие единой точки зрения на информацию, на ее философское и методологическое определение. Современные тенденции в воззрениях на информационные системы говорят о возрастающем месте искусственного интеллекта, его использования при эксплуатации так и при создании (проектировании) систем. Также важно понимать схожести и отличия естественного и искусственного интеллекта для понимания тенденций развития и возможностей применения последнего. Рассматривается обобщенная классификация методов создания типовых подсистем.

Ключевые слова: информация, информационные системы, искусственный интеллект, определения информации, классификация методов создания типовых подсистем.

1. Введение

На всем жизненном пути человека окружают информационные системы, начиная от рождения и записи в учётной системе населения, в процессе взросления - ИС обучения, развлечений и пр., в процессе трудовой деятельности и т.д. Современный мир не представим без большого количества разнообразных информационных систем, которые хранят обрабатывают, передают и выдают информацию. Человечество давно, еще со времён первых философов задумалось о природе информации. Было дано огромное количество определений информации, раскрывающих тот или иной ее аспект или отношение дающего определение к этой самой информации.

Например, некоторые из определений:

- «информация – сведения (сообщения, данные) независимо от формы их представления» [1].
- «ИНФОРМАЦИЯ (от лат. informatio - разъяснение - изложение), первоначальная - сведения, передаваемые людьми устным, письменным или другим способом (с помощью условных сигналов, технических средств и т. д.); с сер. 20 в. общенаучное понятие, включающее обмен сведениями между людьми, человеком и автоматом, автоматом и автоматом; обмен сигналами в животном и растительном мире; передачу признаков от клетки к клетке, от организма к организму; одно из основных понятий кибернетики.» [2].
- «информация - сведения о предметах, процессах и явлениях окружающего мира, передаваемые людьми устно, письменно или иным способом (напр., с помощью условных сигналов, технических средств и т. д.). С сер. 20 в. информация воспринимается более широко, как общенаучное понятие, включающее обмен сведениями между людьми, человеком и автоматом (напр., роботом), автоматом и автоматом; обмен сигналами в животном и растительном мире; передачу признаков от клетки к клетке, от организма к организму (генетическая информация).» [3].

Постепенно с развитием человечества развивалось и понимание информации, ее природы и влияния на окружающую действительность. Здесь мы вынуждены, однако, констатировать, что до сих пор не существует единого, всеми принятого – ну или хотя бы большинством научных работников в мире - понятия того, что же есть такое «информация».

[1] Труды XI Международной научной конференции "Информационные технологии интеллектуальной поддержки принятия решений", 11-13 ноября, Уфа-Чандигарх-Фандербейлпарк, 2025

Часть причин определялось различием в философских воззрениях (течениях).

Согласно работам [4-6], в данный момент можно выделить три основных подхода к пониманию того, что такое «информация»:

1. субстанциональный подход. Согласно данному подходу, информация - субстанция (лат. substantia - сущность; то, что лежит в основе), т.е. самостоятельная сущность;
2. атрибутивный подход: информация есть всеобщее свойство материи, ее атрибут, т.е. существенное, неотъемлемое свойство;
3. функциональный подход: информация - свойство лишь сложных, самоорганизующихся систем, связанное с их функционированием.

Кроме того, необходимо отметить наличие работ, в которых утверждается, что понятие информации является принципиально неопределимым, следовательно – выработать единое достаточно общее понятие термина «информации» невозможно [7].

Также, кстати, следует отметить, что некоторые философы вообще отказываются считать философию наукой. В качестве примера можно привести высказывание доктора философских наук, ведущего научного сотрудника Института философии РАН А. Л. Никифорова: «Философия никогда не была, не является и, по-видимому, никогда не будет наукой» [8]

Появление информационных систем повлияло и на представление об информации. С ИТ-шной точки зрения наиболее близким кажется атрибутивное определение. Действия и взаимодействия с информацией происходят именно по атрибутивному подходу.

Принимая данное определение за основу в ИТ как в предметной области можно адаптировать представление об окружающих информационных системах к современным воззрениям и к новым тенденциям, в частности, идее искусственного интеллекта во всех его проявлениях

2. Философия - относиться к ИИ как к ЕИ или ИИ

Появление идеи, создание, повседневное применение искусственного интеллекта потребовало осмысления и на социальном и философском уровне. Постепенно появились и новые направления, и новые способы использования ИИ, также появились и продолжают появляться новые профессии и направления деятельности, связанные с ИИ. Задаются вопросы о полезности применения, легитимности использования результатов и т.д. Для дальнейшего следует дать определения ИИ. Однако, в настоящее время, как нам кажется, существует некая неопределенность в определении понятия ИИ, доказательством чему может служить множественность таких определений.

Определения ИИ

Искусственный интеллект (AI, Artificial Intelligence) - это технология, которая дает возможность компьютеру имитировать работу человеческого мозга (f.e.: принимать решения на основе неполных данных etc.). [9]

ИИ (искусственный интеллект) это собирательное название для множества подходов, методов и алгоритмов. Сейчас … под ИИ, как правило, подразумеваются нейронные сети. [13]

Искусственный интеллект простыми словами - это метод, позволяющий компьютеру, управляемому компьютером роботу или программному обеспечению мыслить разумно, подобно человеку (f.e.: программа, которая фильтрует спам в электронной почте; программа, показывающая рекламу с учетом вашего места жительства, интересов и поисковых запросов, голосовой помощник в смартфоне/колонке, ПО, управляющее беспилотным автомобилем). [10]

Искусственным интеллектом (ИИ) называют направление компьютерной науки, в рамках которого разрабатываются технологии, способные решать интеллектуальные задачи. [11]

Какое из этих определений правильное? С точки зрения авторов, каждое из них в той или иной мере является адекватным, но с учетом позиции области применения ИИ и с позиции тех, кто даёт определение.

Авторы считают, что ИИ – это совокупность методов и алгоритмов, разработанных человеческим интеллектом, позволяющих имитировать те или иные возможности человеческого мозга.

Также следует задуматься насколько ИИ близок к естественному интеллекту. С точки зрения дальнейшего развития следует ли считать, что естественный интеллект развивается параллельно искусственному. Эти и похожие вопросы волнуют не только учёных различных направлений, но и практиков. Глобально это вопрос - на сколько можно доверять ИИ и результатам его применения.

Рассмотрим вопрос эволюции искусственного интеллекта и на сколько эволюция ИИ совпадает с эволюцией ЕИ

Эволюция ИИ

При сопоставлении эволюции естественного (человеческого) и искусственного интеллектов, существуют, с точки зрения авторов, два критерия:

а) временной – время, за которое произошел тот или иной этап эволюции;

Сопоставление по этому критерию усложняется неизвестностью или, как минимум, неопределенностью нижней границы: какую точку в истории брать за начало эволюции человека (человеческого интеллекта)? Появление Homo sapiens? Или что-то более позднее? Аналогично и с искусственным интеллектом – что брать за нулевую точку? Появление понятия «искусственный интеллект»? То есть, не учитывая эволюцию алгоритмов решения тех или иных задач – основу ИИ?

б) количественный – высота эволюционного перехода.

Здесь тоже ситуация достаточно сложная, поскольку непонятно – как измерять эволюционный скачок?

Вывод: Относиться к ИИ как ЕИ полностью не корректно, но в ограниченных рамках и пределах (например, в необходимости обучения) можно. С другой стороны, предполагать на современном уровне самостоятельное развитие ИИ в нечто новое, наверное, не совсем корректно)

3. Применение ИИ для проектирования ИС [12, 19]

1. Возможности применения преимущества (для проверки предлагаемого проектного решения, для использования в рамках СППР по выбору одного из вариантов проекта, возможно в дальнейшем для генерации вариантов решения (проектов) для не больших задач/подсистем)
2. Необходимость обучения (применение ИИ требует обучения ИИ, особенно на первом, начальном этапе, причем желательно с обучателем определяющим корректность как обучающей выборки, так и результатов. Очевидно, что правильное обучение, как и в случае с ЕИ должно давать именно тот результат, который запланирован)
3. Недостатки и сложности (многовариантность проектных разработок, сложность обучения – малые выборки для статистических методов, сильное влияние человеческого фактора, необходимость высокоуровневого контроля)

Выводы: Данное направление развития ИИ является перспективным, но требующим развития человеческой мысли.

4. Классификация подходов для создания типовых подсистем или функций

В отличии от материального мира и производства вещей в нем при программировании есть возможность реализовать весь программный код одному разработчику, что можно было бы сравнить с ремесленником в древнем мире и в средние века до появления производства. Очевидно, что такой подход не выдерживает требований по скорости разработки, по использованию коллективного труда и пр. Один из подходов связанных с преодолением этих недостатков является использование библиотек и фреймворков.

Согласно [13, 14]: библиотека – это «готовый набор функций и объектов для того или иного языка программирования. Такие наборы объединяются назначением или областью применения. В качестве примера можно выделить математические библиотеки или веб-библиотеки. Такие элементы используются для более быстрой и простой разработки программного обеспечения»

Существует большое количество разнообразных классификаций библиотек, приведем некоторые из них (см. рис.) [14, 15]:

1. По правам авторства
2. По правам собственности (приоритетности) и необходимости приобретения лицензий на их использование.
3. По способу подключения проект
4. и т.д.

Фреймворк обобщенно можно воспринимать как набор различных библиотек, объединённых по тем или иным принципам, например, по месту применения, данная классификация приведена на рис. Выделены фреймворки для работы с пользователями/клиентами – фронтэнд, для обеспечения работы и поддержки взаимодействия с клиентом бэкенд, и фреймворки полного цикла фулстек. Также для фреймворков как и для библиотек можно привести аналогичные классификации по правам, подключениям и т.д.

Еще одним способом превращения ремесла в производство для программирования может служить подход, получивший в российской литературе наименование Типовые Проектные Решения (ТПР), а в иностранной – шаблоны или паттерны. Также стоит отметить, что идеологию подхода заимствовали из «материального мира» (например, типового строительства). Также связав ТПР (или шаблон) с куском готового кода на том или ином языке программирования сильно ускорили и облегчили труд программистов. [17, 18]

Классификация ТПР приведенная на рис. считается общепризнанной классификацией верхнего уровня. Хотя к ТПР также можно применить такие же способы классификации, что и к библиотекам и фреймворкам.

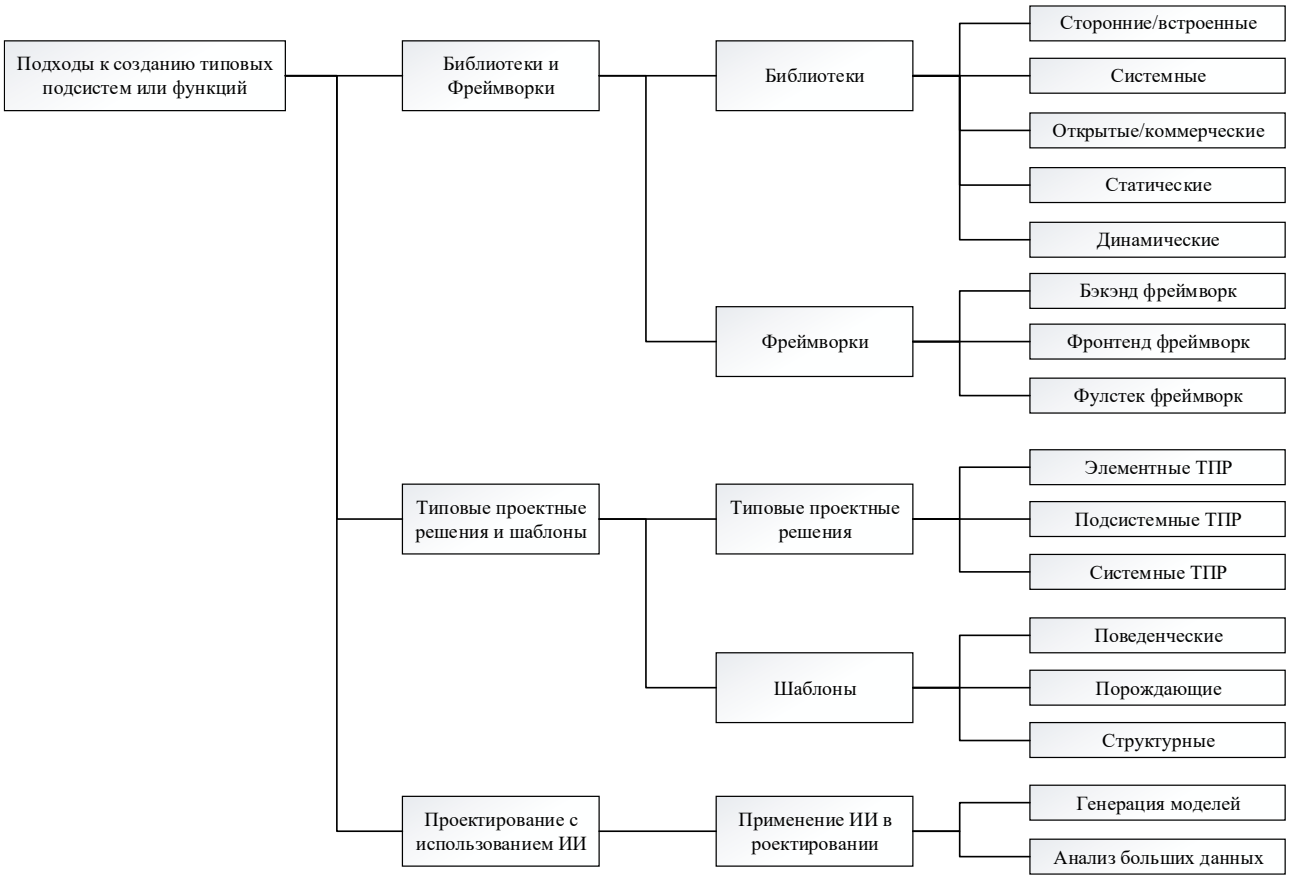


Рис. 1. Обобщенная классификация подходов к созданию типовых подсистем или функций

Шаблонный подход получил широкое применение в проектировании ИС с появлением объектно-ориентированного подхода. В глобальном масштабе появились такие решения как RUP. Согласно [16, 17, 20] «RUP – это процесс, направленный на поддержку коллективной разработки». Классификация шаблонов приведения на рис.

5. Заключение

Как и в живой природе (естественном отборе), когда в начале процесса предлагается множество вариантов, а затем в процессе эволюции какие-то из вариантов остаются и постепенно улучшаются, а какие-то отбрасываются и постепенно сходят «со сцены». Такой же процесс можно наблюдать и в области искусственного интеллекта: в последнее время появилось огромное количество направлений развития и применения ИИ в разнообразных областях деятельности.

На взгляд авторов развитие и применение ИИ для проектирования информационных систем – интересная область приложения искусственного интеллекта, в перспективе способная принести пользу. Исходя из современных воззрений на область информационных систем и их участия в жизнедеятельности человека их количество и разнообразие растет, и эта тенденция роста постепенно возрастает.

Таким образом, постоянно и все больше требуется проектирование новых ИС, модернизация и модификация существующих одним из направлений, облегчающих и упрятывающих данные процессы является применение ИИ.

Список используемых источников

1. [149-ФЗ от 27 июля 2006 г., http://www.consultant.ru/document/cons_doc_LAW_61798/c5051782233acca771e9adb35b47d3fb82c9ff1c/]
2. [Большой Энциклопедический словарь, https://dic.academic.ru/dic.nsf/enc3p/140202]
3. [Энциклопедия «Техника». — М.: Росмэн. 2006; https://dic.academic.ru/dic.nsf/enc_tech/476/%D0%B8%D0%BD%D1%84%D0%BE%D1%80%D0%BC%D0%B0%D1%86%D0%B8%D1%8F]
4. Лысак, И.В. Информация как общенаучное и философское понятие: основные подходы к определению [Text] // Философские проблемы информационных технологий и киберпространства, 2015. № 2, vol. 10. – 9–26. DOI 10.17726/philIT.2015.10.2.391
5. Отраднова. Информация как онтологическая категория // Молодой учёный №10 (45) / 2012, с.169-174

6. К.К. Колин. Природа информации и философские основы информатики // Открытое образование 2005-02, с. 43-51
7. Мелик-Гайказян И.В. Информационные процессы и реальность. – М.: Наука, Физматлит, 1998. – 192с.
8. Никифоров А.Л. Философия как личный опыт // Заблуждающийся разум. Многообразие ненаучного знания. М., 1990. С. 298
9. [Искусственный интеллект — что это и на что он способен. https://journal.sovcombank.ru/tehnologii/iskusstvennii-intellekt--chto-eto-i-na-chto-on-sposoben?ysclid=mdl43arm9w844189453]
10. [ИИ для «чайников» https://habr.com/ru/articles/930140/]
11. [Что такое искусственный интеллект и зачем он нужен https://www.eldorado.ru/blog/publications/ponyatie-iskusstvennogo-intellekta-dlya-chego-on-nuzhen-chelovechestvu-35419?utm_source=yandex&utm_medium=organic&utm_campaign=yandex&utm_referrer=yandex]
12. [Области применения искусственного интеллекта: как технологии меняют мир вокруг нас https://education.yandex.ru/knowledge/oblasti-primeneniia-iskusstvennogo-intellekta-kak-tekhnologii-meniaiut-mir-vokrug-nas]]
13. Library и framework в разработке – что это такое https://otus.ru/journal/library-i-framework-v-razrabotke-chto-eto-takoe/
14. Проектирование информационных систем https://intuit.ru/studies/courses/2195/55/lecture/1622?page=4
15. Design Patterns in Haskell https://blog.ezyang.com/2010/05/design-patterns-in-haskel/
16. Peter Norvig Design Patterns in Dynamic Programming https://efaidnbmnnnibpcajpcglclefindmkaj/https://norvig.com/design-patterns/design-patterns.pdf
17. Y Combinator Revenge of the Nerds https://www.paulgraham.com/icad.html
18. Харольд Абельсон Джеральд Сассман Структу́ра и интерпретация компьютерных программ https://efaidnbmnnnibpcajpcglclefindmkaj/https://web.mit.edu/6.001/6.037/sicp.pdf
19. Искусственный интеллект в проектировании: новые возможности https://xn--b1agapfwapgcl.xn--p1ai/iskusstvennyj-intellekt-v-proektirovanii-novye-vozmozhnosti/
20. RUP. Общие сведения https://www.informicus.ru/mps07.html

Интеллектуальная поддержка принятия решений в управлении проектами на основе оценки рисков, связанных с человеческими эмоциями

Богданова Д. Р.
Институт информатики, математики и робототехники
Уфимский университет науки и технологий
Уфа, Россия
e-mail: dianochka7bog@mail.ru

Пожидаев А. Е.
Институт информатики, математики и робототехники
Уфимский университет науки и технологий
Уфа, Россия
e-mail: pozidaev2002@gmail.com

Аннотация[1]

Современные IT-проекты реализуются в условиях высокой неопределенности, где отклонения по срокам и качеству часто связаны не только с техническими причинами, но и с человеческим фактором. Цель работы - описать подход к интеллектуальной поддержке принятия решений в управлении проектами за счет оценки рисков, связанных с состоянием команды и эмоциональным контекстом коммуникаций.

Использован обзор и систематизация подходов к риск-менеджменту и системам поддержки принятия решений (СППР), предложена формализация задачи как классификации уровня риска по признакам проекта и человеческого фактора, разработана функциональная модель процесса в нотации IDEF0. В качестве эмоциональных признаков рассматривается мультимодальная оценка эмоций по тексту, голосу и мимике с последующей агрегацией по контрольным периодам.

Сформирована структура решения, включающая сбор и интеграцию данных, формирование признаков и мультимодальный анализ эмоций, моделирование риска, интерпретацию результатов и выдачу рекомендаций. Определены управляющие воздействия (регламенты, правила предобработки, метрики качества, правила реагирования) и ожидаемые выходы (уровень риска, объяснение факторов, рекомендации). Предложенный подход обеспечивает переход от позднего реагирования к проактивному выявлению рисков и повышает обоснованность управленческих решений.

Ключевые слова: управление проектами; оценка рисков; человеческий фактор; мультимодальные эмоции; интеллектуальная СППР; объяснимый ИИ; project management; risk assessment; human factors; multimodal emotions; intelligent DSS; explainable AI

1. Введение

Современные IT-проекты нередко развиваются в ситуации, когда заранее невозможно достоверно описать требования, а изменения происходят быстрее, чем обновляются планы. На практике это приводит к тому, что руководителю проекта приходится постоянно принимать решения в условиях неполной информации, например: что делать с просроченными задачами, как перераспределить нагрузку, где усилить контроль качества и в какой момент вмешиваться в коммуникации команды.

Классический риск-менеджмент хорошо работает там, где риски можно напрямую измерять через сроки, бюджет и количество дефектов, однако значительная часть проблем в проектах проявляется через человеческий фактор: перегруз отдельных участников, рост напряжения в обсуждениях, снижение вовлеченности, конфликтность, выгорание. Эти явления часто фиксируются слишком поздно - когда проект уже перешел в критическое состояние, а стоимость исправления ошибок стала высокой.

В связи с этим актуальной становится интеллектуальная поддержка принятия решений,

[1] Труды XI Международной научной конференции "Информационные технологии интеллектуальной поддержки принятия решений", 11-13 ноября, Уфа-Чандигарх-Фандербейлпарк, 2025

которая позволяет не только собирать проектные метрики, но и учитывать косвенные признаки состояния команды. Важный аспект - мультимодальность эмоций: эмоциональные состояния проявляются не только в текстах (чаты и комментарии), но и в голосе на созвонах, а также в невербальных сигналах при видеосвязи. Следовательно, система оценки рисков может опираться на совокупность источников данных и формировать предупреждения и рекомендации до того, как отклонения по срокам и качеству станут критичными.

Цель статьи - описать структуру решения и подход к построению интеллектуальной СППР, способной оценивать риски, связанные с человеческим фактором и эмоциями, и выдавать руководителю проекта интерпретируемые результаты и практические рекомендации.

2. Обзор литературы

В документе [1] описан современный подход к управлению проектами, основанный на принципах и областях ответственности, которые подстраиваются под контекст проекта. Риск и неопределенность рассматриваются как постоянный фактор на протяжении всего жизненного цикла проекта. Описано, что управление должно быть гибким и ориентированным на результат, а не на формальное соблюдение правил. .

В стандарте [2] задана универсальная схема риск-менеджмента. Она состоит из: определение контекста, идентификация, анализ и оценка рисков, выбор мер обработки, коммуникация и мониторинг. Отмечено, что на восприятие риска влияют человеческие и культурные факторы,можно сделать вывод, что поведение людей и эмоциональные реакции можно рассматривать как часть условий, в которых риск возникает и меняется.

В документе [3] приведены общие рекомендации по управлению проектами, которые можно применить к разным стилям работы. Подчеркнута важность координации, коммуникации и управления заинтересованными сторонами. Риск рассмотрен как регулярная задача управления, а не разовая необходимость.

В источнике [4] риск-менеджмент описан как последовательный процесс работы с неопределенностью, где риски нужно структурировать и закреплять ответственность за их контроль. Показано, что без владельцев рисков и понятных процедур мониторинга управление рисками становится формальным. Также описана необходимость связать оценку рисков с конкретными действиями и решениями.

В исследовании [5] рассмотрены технические риски проектов с машинным обучением и предложена их структуризация. Для приоритизации используются экспертные методы, в результате выделяются наиболее значимые факторы. Наиболее высоко ранжируются риски качества и доступности данных, а также риски требований и архитектуры. Это даёт понять, что риски можно описывать как измеримые признаки и отслеживать их системно.

В обзоре [6] рассмотрены объяснимые СППР, где результаты моделей сопровождаются обоснованием. Отражено, что объяснимость повышает прозрачность и доверие пользователя к выводам системы. Систематизированы подходы объяснимого ИИ и направления их применения в прикладных задачах.

В источнике [7] описаны основные типы СППР и их архитектуры. Сделаны выводы, что выбор подхода зависит от доступных данных, моделей и задач пользователя. Рассмотрены системы, ориентированные на данные и на модели, что соответствует задачам прогнозирования и классификации рисков.

В статье [8] ошибки человека рассматриваются как результат системных условий работы, а не только индивидуальных действий. Выделена роль нагрузки, организации процессов и качества информационных потоков. Такой подход предполагает, что риски человеческого фактора нужно связывать с условиями и организационными причинами.

В источнике [9] описано, что люди систематически ошибаются в оценках вероятностей и последствий из-за когнитивных искажений. Рассмотрены механизмы быстрых интуитивных и медленных аналитических решений. Вывод состоит в том, что экспертные оценки рисков могут быть смещены, поэтому полезна опора на данные и формализованные процедуры оценки.

В обзоре [10] систематизированы исследования о связи аффективных состояний инженеров в гибкой разработке с производительностью и качеством. Отмечено, что измерение эмоций в реальных командах часто трудно масштабировать и стандартизировать. Отражено, что чаще изучаются негативные состояния и их связь с поведением и результатами работы.

В исследовании [11] показано применение методов анализа текста к проектным описаниям для получения количественных индикаторов. На основе текстовых данных оцениваются управленческие практики и их согласованность. В исследовании показано, что проектные тексты могут служить источником признаков для автоматизированного анализа.

В источнике [12] введена концепция эмоционального интеллекта как набора способностей к распознаванию, пониманию и регулированию эмоций. Подчеркнуто, что это отличимо от личностных черт. Концепция задаёт основу для описания эмоционального фактора через наблюдаемые проявления саморегуляции и эмпатии.

В обзоре [13] рассмотрены тримодальные подходы к распознаванию эмоций, объединяющие текст, аудио и видео. Проанализированы методы слияния модальностей, используемые данные и ограничения внедрения. Отмечены проблемы шума реальных сигналов, сложности сбора данных, а также вопросы этики и приватности.

В обзоре [14] систематизированы применения искусственного интеллекта в управлении проектами, включая прогнозирование, оптимизацию ресурсов и управление рисками. Сделан вывод, что эффективность внедрения зависит не только от алгоритмов, но и от организационных условий и интеграций. Обобщаются требования к внедрению и типовые области применения.

В работе [15] предложен метод локальных объяснений предсказаний модели, который не зависит от типа исходной модели. Объяснение строится через локальное приближение сложной модели простой интерпретируемой моделью. В исследовании показано, что это помогает пользователю понять причины предсказания и выявлять ошибочные зависимости.

В исследовании [16] предложен подход к объяснению моделей на основе значений Шепли, обеспечивающий согласованную оценку вкладов признаков. Продемонстрировано, что разные методы объяснения можно представить в единой аддитивной рамке. Метод позволяет сравнивать влияние признаков на предсказания и интерпретировать результат для конкретных наблюдений.

В работе [17] обсуждается, что интерпретируемость должна иметь строгие определения и критерии оценки. Предлагается оценивать объяснения с учетом задач пользователя и измеримых процедур. Подчеркивается необходимость проверяемости объяснений, иначе они не имеют практической ценности.

В статье [18] описана программа по объяснимому ИИ, направленная на повышение прозрачности моделей и улучшение взаимодействия с пользователем. Рассматривается сочетание моделей, методов объяснения и интерфейсов представления объяснений. Показано, что объяснимость важна для доверия и применения ИИ в прикладных системах.

В источнике [19] систематизированы интерпретируемые модели и методы объяснения сложных моделей, а также их ограничения. Выделен риск неверной интерпретации, включая смешение корреляции и причинности. Даны рекомендации по выбору методов в зависимости от задачи и типа модели.

В исследовании [20] сравниваются детерминированные методы прогнозирования длительности проектов на основе метода освоенного объема и реальных данных. Была оценена точность и своевременность прогнозов для разных подходов и их комбинаций. Выражено, что выбор метрик и техники прогнозирования существенно влияет на качество контроля проекта.

Анализ показывает, что управление рисками в проектах следует рассматривать как непрерывный процесс, который зависит не только от планов и процедур, но и от условий работы людей, коммуникации и организационной среды. Современные исследования подтверждают, что риски можно формализовать через наблюдаемые показатели проекта и цифровые следы деятельности, а затем использовать модели для раннего предупреждения о неблагоприятных тенденциях по срокам и качеству. Эмоциональные состояния связаны с поведением команды и результатом, при этом эмоции проявляются в разных каналах (текст, голос, мимика), но практическая реализуемость часто требует начинать с наиболее доступных данных. Для внедрения таких решений в управление ключевым условием является объяснимость выводов, чтобы пользователь понимал причины оценки риска и мог проверять и применять рекомендации на практике.

3. Задача оценки рисков проекта с учетом человеческого фактора и подход к ее решению

В статье рассматривается задача оценки уровня риска проекта с учетом показателей выполнения работ, организационных факторов и эмоционального состояния участников команды, которая сведена к задаче классификации уровня риска на основе анализа мультимодальных данных. На рисунке 1 показана структура решения данной задачи в виде диаграммы IDEF0.

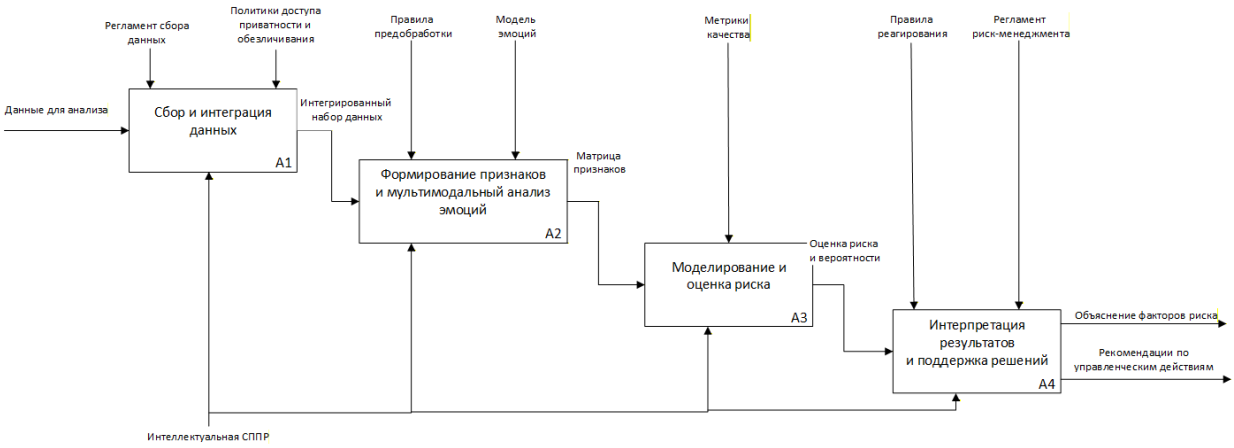


Рис. 1. Структура решения задачи оценки риска проекта

Процесс оценки риска проекта включает следующие этапы: сбор и интеграцию данных, предобработку и формирование признаков с учетом мультимодального анализа эмоциональных состояний, моделирование и оценку уровня риска, а также интерпретацию результатов и формирование рекомендаций по управленческим действиям. Входными данными (рис. 1) являются данные для анализа, характеризующие текущее состояние проекта, включая показатели выполнения задач, результаты коммуникаций участников и организационные события. Управляющими воздействиями выступают регламент сбора данных, политики доступа и обезличивания, правила предобработки, модель эмоций, метрики качества, правила реагирования и регламент риск-менеджмента. Выходными данными являются оценка уровня риска и вероятности, объяснение факторов риска и рекомендации по управленческим действиям.

На основании входных данных осуществляется сбор и интеграция информации, отражающей ход выполнения проекта. В интегрированный набор данных включаются сведения о состоянии задач, динамике выполнения работ, показателях качества, а также данные коммуникаций участников проекта. Сбор данных осуществляется в соответствии с установленным регламентом и с учетом требований к приватности и защите информации.

На этапе предобработки производится очистка и структурирование данных, а также формирование признаков, характеризующих текущее состояние проекта. В качестве признаков используются показатели выполнения работ, такие как отклонения по срокам, динамика задач и показатели качества, а также параметры, отражающие особенности организационного взаимодействия. Дополнительно выполняется мультимодальный анализ эмоционального состояния участников, основанный на обработке текстовых сообщений, анализе характеристик речи на совещаниях и анализе невербальных сигналов, наблюдаемых при видеокоммуникации. Эмоциональные состояния классифицируются по базовым категориям, что позволяет выявлять признаки повышенного напряжения, конфликтности или снижения вовлеченности команды.

Далее осуществляется моделирование и оценка уровня риска проекта на основе сформированного набора признаков. Для оценки используются методы машинного обучения, позволяющие анализировать взаимосвязи между показателями проекта и вероятностью возникновения проблемных ситуаций. Результатом данного этапа является оценка уровня риска и вероятностные показатели, характеризующие степень неопределенности.

На заключительном этапе выполняется интерпретация результатов анализа и формирование рекомендаций по управленческим действиям. На основе полученной оценки риска и с учетом правил реагирования и регламента риск-менеджмента формируются объяснения факторов, оказавших наибольшее влияние на результат, а также рекомендации, направленные на снижение вероятности возникновения неблагоприятных сценариев, включая корректировку процессов работы, распределения задач и организации коммуникаций.

Прототип системы может быть реализован с использованием современных инструментов анализа данных и машинного обучения. В качестве среды разработки целесообразно использовать язык Python и специализированные библиотеки для обработки данных и построения моделей, такие как Pandas и NumPy для подготовки данных, scikit-learn, XGBoost или LightGBM для реализации моделей оценки риска, а также библиотеки обработки естественного языка например, spaCy или модели на базе трансформеров для анализа текстовых коммуникаций. Для анализа аудио- и видеоданных могут применяться инструменты обработки сигналов и компьютерного зрения, такие как librosa и OpenCV. Методы интерпретации результатов могут быть реализованы с использованием библиотек SHAP или LIME, обеспечивающих объяснение факторов, влияющих на оценку риска. Визуализация результатов и предоставление

информации пользователю могут быть организованы с использованием панелей мониторинга или веб-интерфейса. Использование интеллектуальной СППР на базе указанных технологий позволяет автоматизировать процесс анализа и повысить обоснованность принимаемых управленческих решений.

4. Заключение

Основные результаты работы заключаются в следующем:

- рассмотрены особенности учета человеческого фактора и эмоциональных состояний при оценке рисков проектов;
- предложен подход к построению интеллектуальной СППР для оценки уровня риска на основе анализа данных проекта и мультимодальных признаков;
- разработана функциональная модель процесса оценки риска в нотации IDEF0, отражающая основные этапы обработки информации;
- показана возможность применения методов анализа данных и машинного обучения для выявления факторов, влияющих на уровень риска;
- определены направления дальнейшего развития, связанные с расширением источников данных и практической апробацией предложенного подхода.

Список используемых источников

1. A Guide to the Project Management Body of Knowledge (PMBOK® Guide): Seventh Edition; The Standard for Project Management. — Newtown Square, PA: Project Management Institute, 2021. — 274 p. — ISBN 978-1628256642.
2. ISO 31000:2018. Risk management — Guidelines. — Geneva: International Organization for Standardization, 2018. — 16 p.
3. ISO 21502:2020. Project, programme and portfolio management — Guidance on project management. — Geneva: International Organization for Standardization, 2020. — 44 p.
4. Chapman C., Ward S. Project Risk Management: Processes, Techniques and Insights. — 2nd ed. — Chichester: John Wiley & Sons, 2003. — 408 p.
5. Lin C.-T., Huang S.-J. Technical risk model of machine learning based software project development: Identifying and prioritizing key risk factors // Information and Software Technology. — 2024. — Vol. 171. — Art. 107449. — DOI 10.1016/j.infsof.2024.107449.
6. Kostopoulos G., Davrazos G., Kotsiantis S. Explainable Artificial Intelligence-Based Decision Support Systems: A Recent Review // Electronics. — 2024. — Vol. 13, no. 14. — Art. 2842. — DOI 10.3390/electronics13142842.
7. Power D. J. Decision Support Systems: Concepts and Resources for Managers. — Westport, CT: Quorum Books, 2002. — 272 p.
8. Reason J. Human Error. — Cambridge: Cambridge University Press, 1990. — 316 p. — DOI 10.1017/CBO9781139062367.
9. Kahneman D. Thinking, Fast and Slow. — New York : Farrar, Straus and Giroux, 2011. — 499 p.
10. Salido A., Insfran E., Rodríguez P., et al. Agile software engineers' affective states, their performance and software quality: A systematic mapping review // Journal of Systems and Software. — 2023. — Vol. 206. — Art. 111800. — DOI 10.1016/j.jss.2023.111800.
11. David I., Gelbard R. Using sentiment analysis to assess PMBOK knowledge areas' compatibility with agile methodology // Procedia Computer Science. — 2024. — Vol. 239. — P. 381–395. — DOI 10.1016/j.procs.2024.06.185.
12. Salovey P., Mayer J. D. Emotional intelligence // Imagination, Cognition and Personality. — 1990. — Vol. 9, no. 3. — P. 185–211. — DOI 10.2190/DUGG-P24E-52WK-6CDG.
13. Al-Saadawi H. F. T., Das B., Das R. A systematic review of trimodal affective computing approaches: Text, audio, and visual integration in emotion recognition and sentiment analysis // Expert Systems with Applications. — 2024. — Vol. 255, Part D. — Art. 124852. — DOI 10.1016/j.eswa.2024.124852.
14. Prasetyo M. L., Peranginangin R. A., Martinovic N., et al. Artificial intelligence in open innovation project management: A systematic literature review on technologies, applications, and integration requirements // Journal of Open Innovation: Technology, Market, and Complexity. — 2025. — Vol. 11, no. 1. — Art. 100445. — DOI 10.1016/j.joitmc.2024.100445.
15. Ribeiro M. T., Singh S., Guestrin C. "Why Should I Trust You?": Explaining the Predictions of Any Classifier // Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '16). — 2016. — P. 1135–1144. — DOI 10.1145/2939672.2939778.
16. Lundberg S. M., Lee S.-I. A Unified Approach to Interpreting Model Predictions // Advances in Neural Information Processing Systems. — 2017. — Vol. 30. — P. 4765–4774. — arXiv:1705.07874.
17. Doshi-Velez F., Kim B. Towards a Rigorous Science of Interpretable Machine Learning:

препринт. — 2017. — arXiv:1702.08608. — DOI 10.48550/arXiv.1702.08608.

18. Gunning D., Aha D. W. DARPA’s Explainable Artificial Intelligence (XAI) Program // AI Magazine. — 2019. — Vol. 40, no. 2. — P. 44–58. — DOI 10.1609/aimag.v40i2.2850.

19. Molnar C. Interpretable Machine Learning: A Guide for Making Black Box Models Explainable. — 3rd ed. — 2025. — URL: https://christophm.github.io/interpretable-ml-book/ (дата обращения: 18.02.2026).

20. Batselier J., Vanhoucke M. Evaluation of deterministic state-of-the-art forecasting approaches for project duration based on earned value management // International Journal of Project Management. — 2015. — Vol. 33, no. 7. — P. 1588–1596. — DOI 10.1016/j.ijproman.2015.04.003.

Использование сверточных нейронных сетей для определения состояния здоровья пчёл

Махмутов А. Д.
Уфимский университет науки и технологий
Институт информатики, математики и кибернетики
Уфа, Россия
amal-makhmutov@mail.ru

Сазонова Е. Ю.
Уфимский университет науки и технологий
Институт информатики, математики и кибернетики
Уфа, Россия
ekaterina_rassadnikova@mail.ru

Сметанина О. Н.
Уфимский университет науки и технологий
Институт информатики, математики и кибернетики
Уфа, Россия
smoljushka@mail.ru

Аннотация[1]

В работе описывается система автоматической классификации медоносных пчел (Apis mellifera) с применением сверточных нейронных сетей. Разработана оптимизированная архитектура для распознавания пяти категорий пчел. Экспериментально достигнута точность 90% на тестовой выборке. Проанализированы факторы ошибочной классификации. Результаты подтверждают перспективность метода для диагностики заболеваний пчелосемей.

Ключевые слова: компьютерное зрение, сверточные нейронные сети, классификация изображений, апиология, диагностика заболеваний пчел, машинное обучение.

1. Введение

Пчеловодство является стратегически важной отраслью агропромышленного комплекса не только Республики Башкортостан, но и всей Российской Федерации. Согласно данным Росстата, Россия входит в число мировых лидеров по производству меда, а Башкортостан традиционно занимает первое место по объемам его производства и экспорта. Однако индустрия сталкивается с серьезными проблемами, среди которых одной из наиболее значимых остается ранняя диагностика заболеваний пчелосемей. Особую угрозу представляет заражение клещом Varroa destructor (варроатоз), которое при отсутствии своевременных мер может привести к массовой гибели насекомых и значительным экономическим потерям, оцениваемым в миллиарды рублей ежегодно.

Традиционные методы мониторинга состояния пчел основываются на периодическом визуальном осмотре пасечником. Этот подход является трудоемким, требует высокой квалификации, подвержен субъективности и не обеспечивает непрерывного контроля. В условиях крупных пасек ручной осмотр становится практически невыполнимой задачей. Развитие технологий компьютерного зрения и методов глубокого обучения открывает новые возможности для автоматизации процессов диагностики и мониторинга в сельском хозяйстве, в том числе в апиологии.

В данной работе представлена разработка и оптимизация системы автоматической классификации медоносных пчел (Apis mellifera) по визуальным признакам с использованием сверточных нейронных сетей. Целью исследования является создание точного и вычислительно эффективного инструмента, способного различать пчел в различных состояниях, включая здоровых, зараженных паразитами и выполняющих специфические поведенческие акты. Предполагается, что такая система может стать ядром комплексной платформы для неинвазивного, непрерывного мониторинга состояния пчелиных семей в реальном времени.

2. Обзор существующих решений

Традиционные подходы к диагностике заболеваний пчел включают визуальный осмотр, лабораторные анализы и химические тесты. Визуальный осмотр

[1] Труды XI Международной научной конференции "Информационные технологии интеллектуальной поддержки принятия решений", 11-13 ноября, Уфа-Чандигарх-Фандербейлпарк, 2025

требует высокой квалификации специалиста и является субъективным. Лабораторные методы, такие как микроскопическое исследование образцов, точны, но требуют значительных временных затрат и специального оборудования. Химические методы, включая акарицидные обработки, эффективны, но могут негативно влиять на здоровье пчел и качество продукции.

В последние годы появились исследования по применению компьютерного зрения в апиологии. Большинство из них фокусируются на следующих направлениях: подсчет пчел на входе в улей для оценки активности семьи; распознавание пчел-воровок и чужаков; определение наличия клещей с использованием методов классического компьютерного зрения (детекция по цвету, форме).

Однако существующие решения имеют ряд ограничений: низкая точность в условиях изменчивого освещения, неспособность различать сложные поведенческие паттерны и высокая вычислительная стоимость. Предлагаемое в данной работе решение направлено на преодоление этих ограничений через использование глубокого обучения и оптимизированной архитектуры.

3. Комплексный подход к разработке интеллектуальной системы автоматической классификации медоносных пчел по визуальным признакам

В рамках исследования был реализован комплексный подход к разработке интеллектуальной системы автоматической классификации медоносных пчел по визуальным признакам. Для обеспечения методологической стройности и воспроизводимости результатов была разработана детализированная структура исследования, охватывающая все этапы — от концептуализации задачи до практической реализации и валидации модели. Системный подход к организации работы позволил не только достичь высоких метрик качества, но и обеспечить прозрачность каждого этапа разработки, что является важным аспектом как для научного сообщества, так и для потенциальных практических внедрений.

На начальном этапе проектирования была разработана обобщенная блок-схема (Рис. 1.), визуализирующая полный цикл разработки системы — от сбора исходных данных до развертывания готового решения. Данная схема служит не только инструментом планирования, но и наглядным пособием для понимания логической последовательности процессов, составляющих основу исследования.

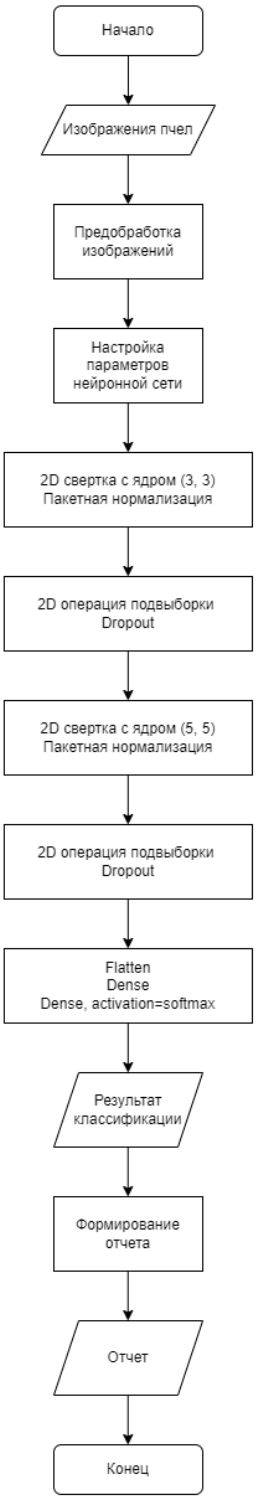


Рис. 1. Алгоритм разработки системы

Блок-схема включает следующие ключевые этапы:

- Сбор и подготовка данных – формирование репрезентативного датасета, включающего изображения пчел различных категорий, что является фундаментальным этапом для любого проекта в области машинного обучения.
- Предобработка изображений – проведение операций нормализации, изменения размеров, а

также применение методов аугментации данных для увеличения разнообразия обучающей выборки и повышения устойчивости модели к изменению условий.

- Проектирование архитектуры CNN – разработка и последующая оптимизация структуры сверточной нейронной сети, учитывающей специфику задачи и особенности визуальных данных.

- Обучение модели – итеративный процесс тренировки нейронной сети с использованием методов контроля переобучения и регулярной валидации на отдельной выборке.

- Тестирование и оценка – комплексная проверка точности и надежности модели на независимой тестовой выборке с расчетом различных метрик качества.

- Анализ ошибок и оптимизация – глубокий анализ случаев ошибочной классификации с целью выявления слабых мест модели и их последующего устранения через корректировку параметров или архитектуры.

- Внедрение и мониторинг – этап практической реализации системы в реальных условиях с организацией непрерывного мониторинга ее работы и производительности.

Для более детального понимания внутренних процессов, происходящих в системе в момент выполнения классификации, была разработана диаграмма активности (Рис. 2. Диаграмма активности). Эта диаграмма демонстрирует последовательность операций, выполняемых системой от момента получения входного изображения до выдачи конечного результата, что позволяет визуализировать логику работы алгоритма и идентифицировать потенциальные узкие места в производительности.

Диаграмма активности описывает следующий процесс:

- Начало работы и инициализация – система активируется при поступлении нового изображения, происходит загрузка предварительно обученной модели и необходимых конфигурационных параметров.

- Загрузка и предобработка изображения – исходное изображение загружается в память, после чего подвергается серии преобразований: конвертация в градации серого для уменьшения вычислительной сложности, нормализация значений пикселей к стандартному диапазону, а также изменение размеров до 100×100 пикселей для обеспечения единообразия входных данных.

- Извлечение признаков с помощью сверточных слоев – предобработанное изображение последовательно проходит через сверточные слои нейронной сети, где происходит автоматическое выделение характерных визуальных признаков, таких как текстура, форма, контуры и другие морфологические особенности.

- Классификация с использованием полносвязных слоев – выделенные признаки передаются на вход полносвязным слоям, которые осуществляют их анализ и преобразование в вероятностные оценки принадлежности к каждому из пяти целевых классов.

- Вывод и интерпретация результата – система формирует итоговый ответ, содержащий метку наиболее вероятного класса, а также дополнительные сведения о распределении вероятностей по всем категориям, что может быть полезно для анализа неоднозначных случаев.

- Логирование и аналитическая обработка результатов – результаты каждой классификации автоматически сохраняются в системном журнале, что позволяет в дальнейшем проводить статистический анализ работы системы, выявлять закономерности и осуществлять мониторинг ее стабильности во времени.

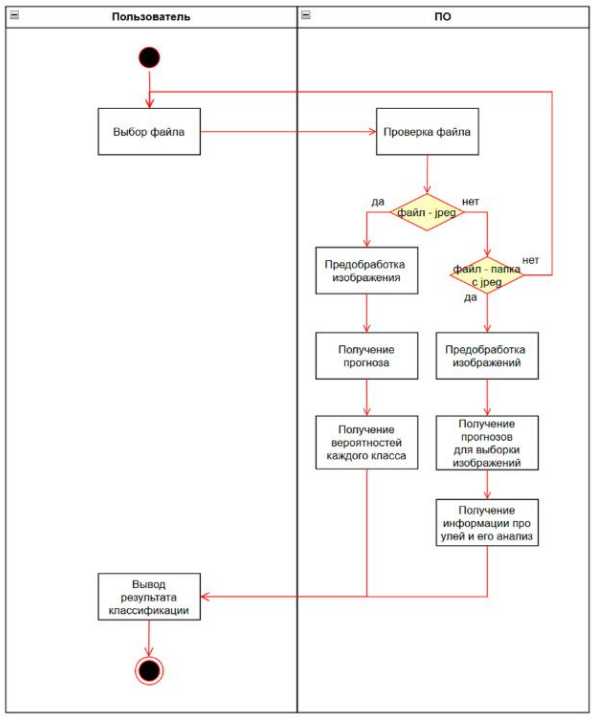


Рис. 2. Диаграмма активности

4. Обучение модели для определения состояния здоровья пчёл

- Для обучения и валидации модели был собран и размечен специализированный датасет, содержащий 5093 изображения насекомых. Сбор данных осуществлялся на учебно-опытной пасеке Башкирского государственного аграрного

университета в период с мая по август 2023 года. Все изображения были получены с использованием единого протокола съемки: камера Sony Alpha 6000 с макрокольцом фиксировалась на расстоянии 15-20 см от объекта при естественном освещении. Среднее разрешение изображений составило 3000×2000 пикселей. Датасет был сбалансирован и включал 5 семантических классов, по 800 изображений в каждом: класс 1: «Охлаждает улей» – пчела с характерно расправленными крыльями, выполняющая вентиляцию гнезда (Рис. 3а); класс 2: «Обычная пчела» – здоровая пчела в состоянии покоя или при движении без видимых аномалий (Рис. 3б); класс 3: «Несет пыльцу» – пчела с обножкой (скоплением пыльцы) на задних лапках (Рис. 3в); класс 4: «Оса» (Vespula vulgaris) – включена для дифференциальной диагностики, что критически важно для исключения ложных срабатываний в полевых условиях (Рис. 3г); класс 5: «Заражена клещом» – пчела с визуально идентифицируемым паразитом Varroa destructor на теле (Рис. 3д).



(а) (б) (в)



(г) (д)

Рис. 3. Пример изображения семантических классов пчел

Процесс разметки осуществлялся тремя независимыми экспертами-энтомологами с последующей верификацией противоречий. Датисет был случайным образом разделен на тренировочную (4087 изображений, ~80%) и тестовую (1006 изображений, ~20%) выборки с сохранением пропорций классов. Для увеличения разнообразия данных и улучшения обобщающей способности модели на тренировочной выборке применялись методы аугментации: случайные повороты на угол до 20°, горизонтальные отражения, небольшие изменения яркости/контрастности, а также добавление гауссова шума.

Предварительная обработка включала следующие этапы: приведение всех изображений к размеру 100×100 пикселей; конвертация в градации серого для уменьшения вычислительной сложности; приведение значений пикселей к диапазону [0, 1].

В ходе исследования была разработана и последовательно оптимизирована архитектура сверточной нейронной сети. Экспериментальным путем было установлено, что для данной задачи оптимальной является неглубокая, но эффективная архитектура из двух сверточных блоков. Использование одного сверточного слоя приводило к недостаточному извлечению признаков и низкой точности (около 65%), в то время как три и более слоев вызывали быстрое переобучение на небольшом датасете.

Финальная архитектура модели включает следующие компоненты:

- Входной слой: Изображение с размером 100x100 пикселей в градациях серого.
- Сверточный блок 1: сверточный слой с 32 фильтрами размером 5x5, функцией активации ReLU; слой пакетной нормализации (Batch Normalization) для ускорения сходимости; слой субдискретизации (MaxPooling) с окном 2x2.
- Сверточный блок 2: сверточный слой с 64 фильтрами размером 5x5, активация ReLU; слой пакетной нормализации; слой MaxPooling 2x2.
- Слой выравнивания (Flatten) для преобразования карт признаков в вектор.
- Полносвязный слой с 128 нейронами и активацией ReLU.
- Слой регуляризации Dropout с вероятностью 0.5 для предотвращения переобучения.
- Выходной полносвязный слой с 5 нейронами и функцией активации softmax для многоклассовой классификации.

Для компиляции модели использовался оптимизатор Adam с learning rate = 0.001, в качестве функции потерь — категориальная перекрестная энтропия. Обучение проводилось в течение 9 эпох с размером пакета (batch size) 32. Мониторинг переобучения осуществлялся с помощью отложенной валидационной выборки. Разработка и эксперименты проводились с использованием стека технологий Python: TensorFlow & Keras для построения нейросети, OpenCV для первичной обработки изображений, NumPy и Pandas для работы с данными.

TensorFlow – это комплексная платформа с открытым исходным кодом для машинного обучения. Он имеет всеобъемлющую гибкую

экосистему инструментов, библиотек и ресурсов сообщества, которая позволяет исследователям продвигать новейшие достижения в области машинного обучения, а разработчикам легко создавать и развертывать приложения на основе машинного обучения. Keras – это API глубокого обучения, написанный на Python, работающий поверх платформы машинного обучения TensorFlow. OpenCV – библиотека алгоритмов компьютерного зрения, обработки изображений и численных алгоритмов общего назначения с открытым кодом. Широко используется в таких языках как C, C++, Python и Java. NumPy – библиотека с открытым исходным кодом для языка программирования Python. Возможности: поддержка многомерных массивов; поддержка высокоуровневых математических функций, предназначенных для работы с многомерными массивами.

Обучение модели заняло 6 минут 30 секунд на графическом ускорителе NVIDIA GeForce GTX 1650, что подчеркивает вычислительную эффективность предложенной легковесной архитектуры. На тестовой выборке модель продемонстрировала общую точность (accuracy) 90.1%. Детальная оценка по классам представлена в матрице ошибок.

Наибольшую точность модель показала на классе «Оса» (97.4%), что объясняется четкими морфологическими отличиями от пчел. Наиболее критичные ошибки возникали между классами «Обычная пчела» и «Заражена клещом»: 9 изображений здоровых пчел были ложно классифицированы как зараженные.

5. Анализ эффективности распознавания

Изучив данные обученной модели, был сделан вывод, что предложенная модель обладает хорошей точностью для всех классов. Наибольшей точностью обладают классы «Матка» (97%) и «Оса» (99%), так как они больше всего отличаются от других классов. Далее по точности идет класс «Несет пыльцу» (92%). Меньшая точность может быть связана с тем, что пыльца у некоторых пчел может быть прикрыта крыльями или брюшком. Следующими по точности идут классы «Заражена клещом» (86%) и «Просто пчела» (74%). Имеются недочеты при распознавании пчел из класса «Просто пчела». Изучив распределение по другим классам, мы увидели, что некоторая часть изображений, которые должны были быть определены в класс «Просто пчела», определяются в класс «Заражена клещом».

Сравнив изображения из тренировочной выборки с метками «Просто пчела» и «Заражена клещом» (Табл. 1), была выдвинута гипотеза о том, что блики света могут быть причиной ошибочного распознавания, так как они похожи на клещей. В таблице с изображениями, красным обведена возможная причина, из-за которой модель может ошибочно относить пчелу не к тому класс.

Таблица 1. Примеры ошибочных изображений

Класс	Пример изображения
Класс «Заражена клещом»	
Класс «Ничего не делает»	

Проведенный анализ позволил выявить ключевой фактор, влияющий на ошибочную классификацию: наличие световых артефактов. Блики от влажного хитина или капель нектара на теле пчелы класса «Обычная пчела» модель в ряде случаев интерпретировала как присосавшихся клещей. Это явление, условно названное «феноменом световой интерференции», указывает на важность стандартизации условий съемки для подобных систем.

В процессе оптимизации модели были получены следующие значимые результаты:

- Количество слоев: Двухслойная архитектура оказалась оптимальным балансом между способностью к обучению и риском переобучения.
- Размер ядра: Увеличение размера ядра свертки до 5x5 на втором слое позволило лучше улавливать морфологические контексты, что повысило точность распознавания класса «Обычная пчела» с 53% (при ядре 3x3) до 74%.
- Регуляризация: Комбинация Batch Normalization и Dropout (p=0.5) снизила разрыв между точностью на тренировочной и валидационной выборках с 15%

до менее чем 3%, эффективно борясь с переобучением.

6. Заключение

В работе представлена и экспериментально обоснована система для автоматической классификации медоносных пчел по визуальным признакам с использованием легковесной сверточной нейронной сети.

Основные результаты исследования заключаются в следующем:

- Разработана и оптимизирована двухслойная архитектура CNN, показавшая точность 90.1% на тестовой выборке, что превосходит традиционные методы компьютерного зрения и сопоставимо с глубокими архитектурами при значительно меньших вычислительных затратах.

- Определены оптимальные гиперпараметры модели: два сверточных слоя с ядром 5×5, комбинированная регуляризация (Batch Normalization и Dropout), 9 эпох обучения. Абляционные исследования подтвердили значимость каждого компонента архитектуры.

- Выявлен и проанализирован ключевой источник ошибок классификации – световые артефакты на изображениях, что важно для формирования требований к качеству входных данных и разработки методов предобработки.

- Продемонстрирована практическая применимость решения для задач апиологического мониторинга, включая диагностику заболеваний, анализ поведения и дифференциацию видов.

Предложенное решение обладает значительным потенциалом для практической имплементации в пчеловодческих хозяйствах в качестве инструмента поддержки принятия решений, способствующего цифровизации и повышению эффективности отрасли.

Несмотря на высокие показатели точности, система имеет следующие ограничения: зависимость от качества данных; статичность модели, то есть текущая версия работает с отдельными кадрами, не учитывая временные зависимости и поведенческие последовательности; ограниченный датасет, хотя датасет сбалансирован, его объем может быть недостаточным для гарантированной работы в разнообразных условиях; для работы в реальном времени на пасеке требуется дополнительная оптимизация для встраиваемых систем.

Перспективы дальнейших исследований включают: расширение функционала: добавление классов для распознавания других заболеваний (нозематоз, аскосфероз), внедрение детекции роевого состояния, оценка силы семьи по визуальным признакам; технические улучшения: разработка модуля предобработки для автоматического устранения бликов и других артефактов, создание мобильного или встраиваемого прототипа системы для проведения натурных испытаний, внедрение временных моделей (LSTM, Transformers) для анализа видеопотоков; интеграционные решения: разработка API для интеграции с существующими системами управления пасекой, создание облачной платформы для сбора и анализа данных с множества пасек, разработка рекомендательной системы для пчеловодов на основе собранных данных.

Acknowledgments (благодарности)

Работа выполнена при финансовой поддержке в форме гранта, полученного автором в конкурсе «Студенческий стартап» (очередь VI). Грант предоставлен на реализацию проекта по разработке интеллектуальной системы мониторинга в пчеловодстве.

Список используемых источников

1. ГОСТ Р 56812-2015. Пчеловодство. Методы диагностики заболеваний пчел. – Введ. 2016-07-01. – М.: Стандартинформ, 2015. – 12 с.

2. ГОСТ Р 54644-2011. Мед натуральный. Технические условия. – Введ. 2013-01-01. – М.: Стандартинформ, 2011. – 8 с.

3. О пчеловодстве в Российской Федерации: федеральный закон от 30.12.2020 № 490-ФЗ // Собрание законодательства РФ. – 2021. – № 1 (часть I). – Ст. 1.

4. Ильина И. В. Современные методы диагностики заболеваний пчел // Ветеринария сельскохозяйственных животных. – 2022. – № 4. – С. 45-51.

5. Петров Р. С. Цифровизация в пчеловодстве: перспективы и вызовы // Агроинформатика. – 2023. – № 1. – С. 23-30.

6. Отчет о состоянии пчеловодства в Республике Башкортостан за 2023 год / Союз пчеловодов Республики Башкортостан. – Уфа, 2024. – 67 с.

7. Методические рекомендации по диагностике и профилактике варроатоза пчел / ФГБНУ "ВНИИ ветеринарной энтомологии и арахнологии". – Тюмень, 2022. – 45 с.

8. Сельское хозяйство России. 2023: статистический сборник / Росстат. – М., 2023. – 250 с.

9. LeCun Y., Bengio Y., Hinton G. Deep learning // Nature. – 2015. – Vol. 521. – P. 436-444.

10. Simonyan K., Zisserman A. Very Deep Convolutional Networks for Large-Scale Image Recognition // arXiv preprint arXiv:1409.1556. – 2014.

11. He K., Zhang X., Ren S., Sun J. Deep Residual Learning for Image Recognition // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. – 2016. – P. 770-778.

12. Selvaraju R.R., Cogswell M., Das A., Vedantam R., Parikh D., Batra D. Grad-CAM: Visual Explanations from Deep Networks via Gradient-based Localization // Proceedings of the IEEE International Conference on Computer Vision. – 2017. – P. 618-626.

13. Abadi M., Barham P., Chen J., Chen Z., Davis A., Dean J., Devin M., Ghemawat S., Irving G., Isard M. et al. TensorFlow: A System for Large-Scale Machine Learning // Proceedings of the 12th USENIX Symposium on Operating Systems Design and Implementation (OSDI'16). – 2016. – P. 265-283.

14. Bradski G. The OpenCV Library // Dr. Dobb's Journal of Software Tools. – 2000. – Vol. 25. – № 11. – P. 120-126.

15. McKinney W. Data Structures for Statistical Computing in Python // Proceedings of the 9th Python in Science Conference. – 2010. – P. 51-56.

16. Состояние пчеловодства в Российской Федерации в 2020-2023 гг.: аналитический отчет / Россельхознадзор. – 2023. – 89 с.

17. Развитие пчеловодства в субъектах Российской Федерации: аналитические материалы / Министерство сельского хозяйства РФ. – 2023. – 112 с.

18. Harris C.R., Millman K.J., van der Walt S.J., Gommers R., Virtanen P., Cournapeau D., Wieser E., Taylor J., Berg S., Smith N.J. et al. Array programming with NumPy // Nature. – 2020. – Vol. 585. – P. 357-362.

19. Kingma D.P., Ba J. Adam: A Method for Stochastic Optimization // arXiv preprint arXiv:1412.6980. – 2014.

20. Ioffe S., Szegedy C. Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift // Proceedings of the 32nd International Conference on Machine Learning. – 2015. – P. 448-456.

21. Srivastava N., Hinton G., Krizhevsky A., Sutskever I., Salakhutdinov R. Dropout: A Simple Way to Prevent Neural Networks from Overfitting // Journal of Machine Learning Research. – 2014. – Vol. 15. – P. 1929-1958.

22. Varroa destructor: биология, экология и методы борьбы / под ред. А. В. Смирнова. – М.: Агропромиздат, 2021. – 234 с.

23. Цифровые технологии в сельском хозяйстве: монография / под ред. В. П. Иванова. – СПб.: Лань, 2022. – 345 с.

24. Krizhevsky A., Sutskever I., Hinton G.E. ImageNet Classification with Deep Convolutional Neural Networks // Communications of the ACM. – 2017. – Vol. 60. – № 6. – P. 84-90.

25. Paszke A., Gross S., Massa F., Lerer A., Bradbury J., Chanan G., Killeen T., Lin Z., Gimelshein N., Antiga L. et al. PyTorch: An Imperative Style, High-Performance Deep Learning Library // Advances in Neural Information Processing Systems 32. – 2019. – P. 8024-8035.

Поддержка принятия решений при применении принципов кайдзен в организационных системах

Зайцева Н. Ю.
Самарский федеральный исследовательский центр РАН
Институт проблем управления сложными системами РАН
Самара, Россия
e-mail: nat.polyaeva@yandex.ru

Моисеева Т. В.
Поволжский государственный университет телекоммуникаций и информатики
Самара, Россия
e-mail: mtv-2002@yandex.ru

Аннотация[1]

Анализ современных национальных моделей управления выявил общую тенденцию ориентации на субъекта и переосмысление роли личности в организационных системах. Японская философия кайдзен признается одним из успешных методов решения проблемы вовлечения сотрудников в дела организации, делая человеческую личность ключевым звеном новой модели управления. Целью работы является поиск средств поддержки принятия решений для реализации принципов кайдзен в организационных системах. Приводится классификация типов управления в зависимости от страны происхождения. Исследуется возможность интеграции философии непрерывного улучшения (кайдзен) в процессы принятия решений в организационных системах, ориентированные на активное участие и учет мнений заинтересованных сторон (субъектов-акторов), что является научной новизной исследования. Проводится параллель между принципами кайдзен и концепцией интерсубъективного управления. При проведении исследования используются методы системного и сравнительного анализа. Предлагается применить инструменты интерсубъективного управления для поддержки принятия решений при реализации принципов кайдзен в организационных системах. Представлена структура интерсубъективно-ориентированной системы поддержки принятия решений (СППР), усовершенствованной для реализации принципов кайдзен. Практическая значимость исследования заключается в разработке архитектуры субъектно-ориентированной кайдзен-СППР.

Ключевые слова: кайдзен; субъектно-ориентированный подход; поддержка принятия решений; интерсубъективное управление; система поддержки принятия решений.

1. Введение

Принципы управления, сложившиеся в современных организационных системах, определяются культурой, политическими и экономическими факторами, доминирующими в рамках определенных географических границ. Поэтому сегодня принято говорить о существовании таких видов менеджмента, как японский, американский и европейский (рис. 1).

Несмотря на то, что в России управленческая наука и практика складываются особым образом уже около тридцати лет, понятия «российский менеджмент» пока не существует. Ученые и менеджеры, формируя собственные подходы, пытаются использовать те особенности зарубежного управления, которые наилучшим образом соответствовали бы реалиям нашей жизни. Анализируя особенности различных моделей управления, можно отметить, что главенствующим подходом в новой системе управления станет ориентация на личность человека («гуманистический подход»).

[1] Труды XI Международной научной конференции "Информационные технологии интеллектуальной поддержки принятия решений", 11-13 ноября, Уфа-Чандигарх-Фандербейлпарк, 2025

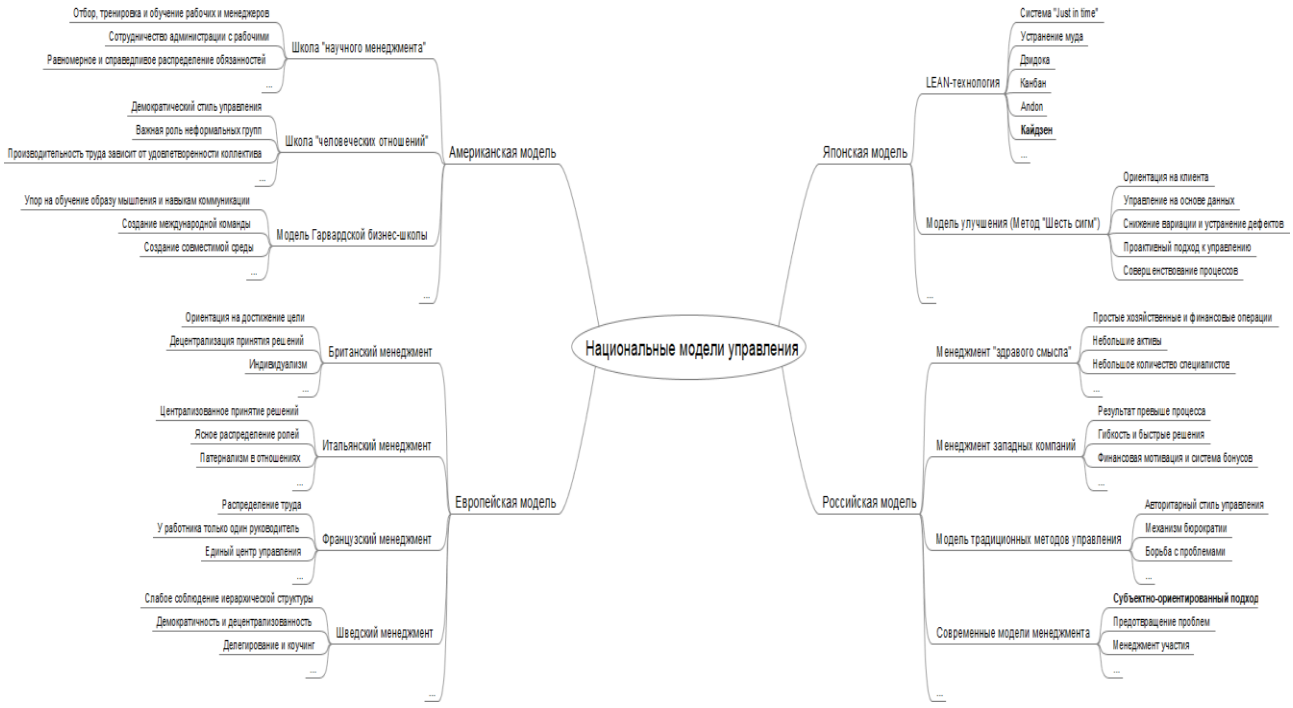


Рис. 1. Классификация видов управления «по странам»

Происходящее сегодня переосмысление роли личности в организационных системах приводит к тому, что многие традиционно инженерно-технические задачи переходят в разряд социотехнических, что означает, что методы их решения радикально меняются. Японская философия кайдзен признается одним из успешных методов решения проблемы вовлечения сотрудников в дела организации, делая человеческую личность ключевым звеном новой модели управления. Собственно термин «кайдзен» означает совершенствование, а применительно к организационным системам - это непрерывное совершенствование, вовлекающее в производственный процесс и менеджеров, и рабочих.

Поскольку философия японского менеджмента не так хорошо известна российским специалистам, кратко рассмотрим основные положения кайдзен как элемента бережливого производства.

2. Кайдзен как элемент LEAN-технологии

LEAN (бережливое производство) - это управленческая философия и набор инструментов, направленных на максимизацию ценности для клиента при минимизации отходов. В основе LEAN-технологии лежит постоянное стремление к совершенствованию процессов, устранению всего, что не добавляет ценности, и вовлечению всех сотрудников в этот процесс [1-4].

Основные принципы LEAN-технологии:

- определение ценности: четкое понимание того, что именно клиент считает ценным в продукте или услуге;
- картирование потока создания ценности: визуальное представление всех шагов, необходимых для доставки ценности клиенту, от начала до конца. Определение и устранение "узких мест" и отходов;
- обеспечение непрерывного потока: организация процессов таким образом, чтобы продукт или услуга двигались плавно и без задержек через все этапы создания ценности;
- вытягивание (Pull-система): производство только тогда, когда есть реальный спрос со стороны клиента, избегая перепроизводства и складирования запасов;
- постоянное совершенствование (Kaizen): непрерывный поиск и устранение отходов, улучшение процессов и повышение эффективности. Вовлечение всех сотрудников в процесс улучшений.

Основные инструменты LEAN-технологии представлены на рис. 2.

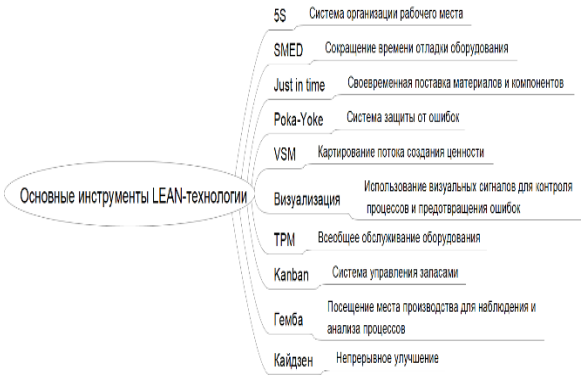


Рис. 2. Основные инструменты LEAN-технологии

Считается, что существенная часть рабочего времени руководителя отводится на разрешение проблем, которые всплывают в процессе работы. Соответственно, умение разрешать текущие проблемы в современных организационных системах ценится гораздо выше, чем предотвращение их появления.

Но не лучше ли предусмотреть возможность назревания проблемной ситуации и не допустить ее?

Следование стратегии кайдзен означает совершенствование процессов в организации и становится ключевой постоянной функцией менеджера, не допускающего «взрывов» и «пожаров». LEAN-технология, в частности кайдзен, которые направлены и сфокусированы на постоянное совершенствование, адаптацию к изменяющимся условиям, вовлечение всех сотрудников и создание ценности для клиента, хороши как идеология. Изучение философии кайдзен привело к пониманию того, что в качестве одного из инструментов ее практической реализации может быть применен субъектно-ориентированный подход к управлению разрешением проблемных ситуаций сотрудников (с тем, чтобы они не переросли в мешающий производству фактор) - теория интерсубъективного управления, разработкой которой занимаются авторы данной статьи.

3. Принципы интерсубъективного управления разрешением проблемных ситуаций в организационных системах

Идея интерсубъективного подхода к управлению разрешением проблемных ситуаций была сформулирована В.А. Виттихом [5-7].

Смыслом интерсубъективного подхода к управлению разрешением проблемных ситуаций в организации является то, что разрешение ситуации, в которой оказываются ее сотрудники, проводится коллегиально путем консенсуса акторами, ощущающими себя внутри проблемной ситуации [8]. Главная особенность подхода заключается в том, что решение принимается не обладающими властью в организации, а заинтересованными в решении людьми, определяющими, что нужно сделать для разрешения проблемной ситуации. Таким образом, субъект управления включается в объект управления (рис. 3).



Рис. 3. Пересечение субъекта и объекта управления при применении интерсубъективного подхода к управлению

4. Применение принципов кайдзен при разрешении проблемных ситуаций

Принципы кайдзен, философии непрерывного улучшения, предоставляют ценный набор методов и подходов, направленных на совместное выявление, анализ и устрашение проблемных ситуаций с учетом мнения всех заинтересованных сторон. В этом контексте применение, внедрение и обучение сотрудников организационных систем философии кайдзен может быть реализовано с использованием интерсубъективного подхода к разрешению проблемных ситуаций.

Вместо навязывания единственно правильного решения, кайдзен в рамках интерсубъективного подхода предполагает создание платформы для открытого диалога и обмена знаниями и мнениями. Инструменты, такие как кружки качества и мозговые штурмы, стимулируют активное участие всех членов команды в процессе поиска решений. Анализ коренных причин самими акторами позволяет выявить не только поверхностные симптомы проблемы, но и лежащие в ее основе системные факторы, которые могут быть упущены при рассмотрении ситуации с одной, "объективной" точки зрения руководителя. Принципы 5S, визуализация процессов и стандартизация процедур обеспечивают прозрачность и предсказуемость в организации, что способствует снижению напряженности и повышению доверия между сотрудниками организационных систем.

Важным аспектом применения интерсубъективного подхода к управлению разрешением проблемных ситуаций, созвучным идеологии кайдзен, является акцент на обучении и развитии персонала. Получение сотрудниками необходимых навыков и знаний для решения проблем и работы в команде, а также создание культуры, поощряющей инициативу и эксперименты, способствует формированию активного подхода к устранению проблем. В конечном итоге, успешное практическое применение принципов интерсубъективного подхода для реализации кайдзен не только

позволяет эффективно разрешать проблемные ситуации, но и укрепляет межличностные отношения, развивает чувство общей ответственности и способствует созданию более адаптивной организации, ориентированной на инновационное развитие.

Основные особенности, сходства, отличия, принципы каждого подхода применительно к организационным системам представлены в таблице 1.

Таблица 1 – Сравнительная таблица кайдзен и интерсубъективного подходов

Критерии	**Кайдзен**	**Интерсубъективный подход**
Основная цель	Завоевать потребителей	Разрешение проблемных ситуаций
Участники	Весь коллектив	Весь коллектив
Подход	Непрерывное совершенствование	Субъектно-ориентированный подход к разрешению проблемных ситуаций
Основные принципы	Ориентация на клиента, непрерывное улучшение, открытость к проблемам, вовлечение сотрудников и др.	Ориентация на субъекта, открытость к диалогу и разрешению проблем, консенсус, коммуникация сотрудников всех уровней и др.
Практические требования	Самодисциплина, моральный дух, командная работа, непрерывность изменений, открытость	Учет особенностей акторов, создание общего коммуникативного пространства, вовлеченность сотрудников, стремление к консенсусу
На кого/что ориентированы	Фокус на клиентах, непрерывные изменения, на работу в команде	Фокус на акторов-сотрудников, консенсус, полилог, взаимодействие и сотрудничество
Роль руководителя	Тренер	Союзник
Отношение к персоналу	Вовлеченность сотрудников	Вовлеченность сотрудников
Решение проблем	Ставится вопрос «Как это произошло?» не для того, чтобы наказать виновных, а чтобы разобраться в причинах, устранить последствия и постараться не допускать повторения	Ставится вопрос «Что делать?» для того, чтобы найти совместное, удовлетворяющее всех акторов решение за счет осознания проблемы изнутри, а не познания извне
Регламент/процедуры	Динамичны и рассчитаны на постоянные изменения	Динамичны, формируются акторами самостоятельно в процессе переговоров

В таблице представлено сравнение по следующим пунктам:

- основная цель;
- участники;
- подход;

- основные принципы;
- практические требования;
- на кого/что ориентация;
- роль руководителя;
- отношение к персоналу;
- решение проблем;
- регламент/процедуры.

В ходе анализа основных особенностей двух подходов (кайдзен и интерсубъективного управления) отмечается глобальное сходство их идеологий, что позволяет предположить, что интерсубъективное управление может стать хорошим инструментом реализации кайдзен в организационных системах. В качестве одной из главных характеристик, объединяющих два подхода, отметим особо фокус на людей с их субъективными особенностями, знаниями, опытом и уникальностью.

5. Субъектно-ориентированная система поддержки принятия решений на базе кайдзен

Авторы статьи уже несколько лет занимаются разработкой концепции интерсубъективного управления и анализом возможностей его приложения для решения задач управления в организационных системах [9-12].

Это связано с тем, что современные организации сталкиваются с возрастающей сложностью и динамичностью внешней среды, что требует от них высокой адаптивности и эффективности в принятии решений. Традиционные, централизованные подходы к управлению и принятию решений все чаще оказываются неэффективными, уступая место моделям, в которых активное участие принимали бы сотрудники разных уровней. Например, субъектно-ориентированный подход ставит в центр внимания акторов-сотрудников, признавая их роль не только как источников данных, но и как активных участников процесса принятия решений. Такой подход, акцентирующий внимание на вовлеченности, знаниях и опыте акторов, является перспективным направлением в управлении и требует современных способов информационной поддержки. Один из них – это системы поддержки принятия решений (СППР). СППР для поддержки акторов в процессе разрешения проблемных ситуаций в организационных системах, основанная на принципах интерсубъективного управления, была представлена в [13-16].

В отличие от традиционных СППР, которые, в основном, фокусируются на автоматизации процессов анализа данных и предоставлении аналитических отчетов, разрабатываемая система учитывает важные характеристики так называемого «человеческого фактора» (экспертные знания, межличностные отношения и неформальные коммуникации), что повышает вовлеченность сотрудников в производственные процессы.

Предлагается адаптация разработанной субъектно-ориентированной СППР, опирающейся на интерсубъективную модель принятия решений, для реализации принципов кайдзен-философии непрерывного улучшения.

Ключевые принципы кайдзен, представленные в СППР:

1. Принцип непрерывного улучшения: систематическое стремление к улучшению процессов, продуктов и услуг. Разрешение проблемных ситуаций в субъектно-ориентированной СППР направлено на совершенствование процессов управления в организационных системах, повышает эффективность принимаемых решений, способствует выявлению "узких мест" и разрешению противоречий.
2. Вовлечение всех сотрудников: активное участие всех членов организации в процессе улучшения. В субъектно-ориентированной СППР созданы механизмы взаимодействия сотрудников на всех уровнях управления в процессе разрешения проблемных ситуаций, актуальных для вовлеченных в них субъектов.
3. Ориентация на процесс: сосредоточение внимания на оптимизации процессов принятия решений, а не только на конечных результатах. Идеология субъектно-ориентированной СППР направлена на вовлечение субъектов, заинтересованных в результатах, в процессы принятия решений.
4. Ориентация на клиента (внутреннего и внешнего): учет потребностей и ожиданий всех заинтересованных сторон. В субъектно-ориентированной СППР это означает участие пользователей в процессе разработки и развития системы.
5. Принятие решений на основе фактов: использование данных и информации для принятия обоснованных решений. В субъектно-ориентированной СППР в процессе разрешения проблемных ситуаций акторы представляют свои точки зрения, подкрепляя их необходимой информацией, используя средства СППР. Такая информация хранится в базе данных, обеспечивая в дальнейшем доступ к релевантным данным и инструментам анализа.
6. Устранение потерь: исключение деятельности, не добавляющей ценности. В субъектно-ориентированной СППР при разрешении проблемных ситуаций, в которых кровно заинтересованы акторы, они сами выбирают, какую деятельность следует осуществлять,

оптимизируя процесс принятия решения, исключая непродуктивные элементы.

7. Стандартизация: установление и соблюдение стандартов для процессов и процедур. В субъектно-ориентированной СППР акторы, прежде чем приступить к обсуждению собственно проблемной ситуации, договариваются о регламенте принятия решения (нормативная интерсубъективность).

Ранее авторами была разработана и представлена архитектура субъектно-ориентированной СППР [13, 16]. Особенности кайдзен-идеологии повлияли на архитектуру СППР, которая была усовершенствована путем добавления новых модулей. Усовершенствованная архитектура предлагаемой субъектно-ориентированной кайдзен-СППР состоит из следующих ключевых компонентов, которые взаимосвязаны между собой (рис. 4).



Рис. 4. Архитектура субъектно-ориентированной кайдзен-СППР

Модуль сбора и агрегации информации обеспечивает сбор данных из различных источников (корпоративные базы данных, отчеты, отзывы клиентов, информация из социальных сетей и т.д.) и агрегацию в единое информационное пространство.

Модуль аналитики и визуализации данных предоставляет инструменты для анализа данных, выявления закономерностей и трендов, а также визуализации информации в удобном и понятном формате.

Модуль экспертных оценок позволяет собирать и анализировать экспертные оценки, мнения и прогнозы специалистов по различным аспектам принимаемых решений. Этот модуль может включать в себя инструменты для проведения опросов, мозговых штурмов и онлайн-конференций.

Модуль рекомендаций. На основе анализа данных и экспертных оценок генерирует рекомендации для принятия решений. Рекомендации могут быть представлены в виде сценариев, альтернативных вариантов действий или предложений по улучшению процессов.

Модуль обратной связи и оценки эффективности позволяет собирать отзывы пользователей о качестве принятых решений и эффективности работы системы. Эта информация используется для непрерывного улучшения процессов и функциональности СППР.

Модуль управления знаниями обеспечивает хранение, поиск и обмен знаниями и опытом между участниками процесса принятия решений. Этот модуль включает в себя базы знаний, форумы для обсуждения и инструменты для совместной работы над документами.

Функциональные средства кайдзен, реализованные в СППР, следующие:

- цикл PDCA (Plan-Do-Check-Act) используется для непрерывного улучшения процессов принятия решений. На этапе Plan определяются цели и задачи, на этапе Do принимаются решения и реализуются действия, на этапе Check оценивается эффективность принятых решений, а на этапе Act вносятся корректировки в процессы и процедуры;

- диаграммы Исикавы (причинно-следственные диаграммы) используются для выявления коренных причин проблем, возникающих в процессе принятия решений. Участие экспертов-акторов и анализ различных факторов позволяют выявить системные проблемы и разработать эффективные решения;

- принципы 5S (Сортировка, Соблюдение порядка, Содержание в чистоте, Стандартизация, Совершенствование) применяются для организации информационного пространства, обеспечения доступности и достоверности данных, а также для создания комфортной среды для работы с системой;

- кружки качества используются для вовлечения сотрудников в процесс улучшения. Кружки качества представляют собой группы сотрудников, которые регулярно собираются для обсуждения проблем и разработки решений (разрешения проблемных ситуаций).

6. Заключение

Предложенная архитектура СППР и механизмы реализации принципов кайдзен предоставляют основу для разработки и внедрения эффективных инструментов поддержки принятия решений в организациях различного масштаба и отраслевой принадлежности. Дальнейшие исследования авторов направлены на доработку прототипа системы и его апробацию в реальных условиях.

Список используемых источников

1. Джеффри Л. Дао Toyota. 14 принципов менеджмента ведущей компании мира. – Альпина Бизнес Букс, 2021. Том 4. – С.368.
2. Имаи М. Кайдзен. Ключ к успеху японских компаний. – «Альпина Диджитал», 1986.
3. Тихонов Н.Е., Андреасян Г.М. Что такое кайдзен? Пять правил кайдзен // Экономика и социум. 2021. № 11-2 (90). С. 439-442.
4. Бочко О.Ю. Использование элементов стратегии кайдзен в маркетинговых коммуникациях // Экономика и управление АПК. 2014. Т. 1. С. 32-36.
5. Vittikh V.A. Evergetics: Science of Intersubjective Management Processes in Everyday Life // International Journal Management Concepts and Philosophy. 2016. Vol. 9. №.2. P.63-72.
6. Vittikh V.A. Introduction to the theory of intersubjective management // Group Decision and Negotiation. 2015. Issue 1. Vol. 24. P. 67-95.
7. Vittikh V.A. Heterogeneous actor and everyday life as key concepts of evergetics // Group decision and negotiation. 2015. Vol. 24. No. 6. P. 949-956.
8. Moiseeva, T. Intersubjective Management: Model of the Problem Situation / T. Moiseeva, N. Polyaeva // Proceedings of the XXI International Conference "Complex Systems: Control and Modeling Problems" (CSCMP 2019). IEEE Xplore. P. 765 – 768.
9. Moiseeva, T. Intersubjective Management Theory: Conceptualization of the Subject Area / T. Moiseeva, N. Polyaeva // Proceedings of the 8th Scientific Conference on Information Technologies for Intelligent Decision Making Support (ITIDS 2020). Volume 483. Ufa-Stavropol, Russia. 2020. P.19-23.
10. Моисеева Т.В. Управление знаниями при интерсубъективном управлении инновационным развитием социотехнических объектов // Вестник Самарского государственного технического университета. Серия: Технические науки. 2018. № 3 (59). С. 41-55.
11. Зайцева Н.Ю., Моисеева Т.В., Мятишкин Ю.В. Система поддержки принятия решений при субъектно-ориентированном управлении социально-экономическими системами // Информационные технологии интеллектуальной поддержки принятия решений (памяти проф. Н.И. Юсуповой) ITIDS'2024. Труды X Международной научной конференции. В 2-х томах. Уфа, 2024. С. 260-267.
12. Моисеева Т.В., Поляева Н.Ю. Инфокоммуникационная поддержка взаимодействия акторов в теории интерсубъективного управления // Вестник Дагестанского государственного технического университета. Технические науки. 2018. Т. 45. № 2. С. 158-170.
13. Моисеева Т.В. Методологические основы поддержки принятия решений по управлению инновационным развитием социотехнических объектов на основе интерсубъективного подхода // Системная инженерия и информационные технологии. 2023. Т. 5. № 2 (11). С. 66-95.
14. Исаева Е.М., Кузнецов К.Д., Моисеева Т.В., Поляева Н.Ю. Анализ функциональных особенностей СППР при интерсубъективном управлении разрешением проблемных ситуаций // Цифровая трансформация социальных и экономических систем. Москва, 2023. С. 1142-1150.
15. Исаева Е.М., Моисеева Т.В., Поляева Н.Ю., Точильникова К.А. Структура систем поддержки принятия решений по разрешению проблемных ситуаций // XXIV Всероссийская студенческая научно-практическая конференция Нижневартовского государственного университета. Нижневартовск, 2022. С. 81-84.
16. Поляева Н.Ю. Разрешение проблемных ситуаций: архитектура системы поддержки принятия решений // Эргодизайн. 2023. № 2 (20). С. 117-124.

Анализ качества математических моделей-кандидатов при прогнозировании электропотребления

Дзгоев А. Э.
Кафедра цифровой трансформации
ИИТ МИРЭА – Российский технологический университет
Москва, Россия
e-mail: dzgoev@mirea.ru

Брайловский А. В.
Кафедра цифровой трансформации
ИИТ МИРЭА – Российский технологический университет
Москва, Россия
e-mail: brajlovskij@mirea.ru

Морозов Д. В.
Кафедра цифровой трансформации
ИИТ МИРЭА – Российский технологический университет
Москва, Россия
e-mail: morozov_dav@mirea.ru

Копылова Я. А.
Кафедра цифровой трансформации
ИИТ МИРЭА – Российский технологический университет
Москва, Россия
e-mail: kopylova_y@mirea.ru

Мажей Я. В.
Кафедра цифровой трансформации
ИИТ МИРЭА – Российский технологический университет
Москва, Россия
e-mail: mazhej@mirea.ru

Аннотация[1]

Анализ качества регрессионных уравнений является одним из этапов проверки пригодности моделей для прогнозирования потребления электроэнергии, что особенно актуально при работе компании на оптовом рынке электроэнергии и мощности (ОРЭМ). Ошибка прогнозирования может привести компанию к штрафам и финансовым убыткам. Целью данной работы является разработка набора различных по структуре адекватных и качественных математических моделей-кандидатов для прогнозирования электропотребления на сутки вперед. Одним из условий включения модели в набор была проверка адекватности модели. Рассчитаны коэффициенты всех моделей в наборе. Коэффициенты проверены на значимость по t-критерию Стьюдента. Проведена проверка качества всех разработанных моделей по нескольким критериям качества. В статье показана проблема выбора наиболее подходящей к имеющимся данным электропотребления математической модели, когда несколько критериев качества показывают высокие результаты. Проведены расчёты прогнозных значений электропотребления на каждый день. Вычислены ошибки прогнозирования и построена визуализация результатов. Проведена процедура двухэтапного выбора лучшей регрессионной модели из набора моделей-кандидатов с использованием комбинации критериев AIC/AICc и RMSE. Интерпретированы результаты проведенного анализа математических моделей по критериям качества, для понимания природы данных. Выбранную математическую модель аппроксимации данных можно использовать на энергетических предприятиях для прогнозирования суточного электропотребления в регионе.

Ключевые слова: качество моделей, выбор модели, прогнозирование электропотребления, адекватность модели, критерии качества,

[1] Труды XI Международной научной конференции "Информационные технологии интеллектуальной поддержки принятия решений", 11-13 ноября, Уфа-Чандигарх-Фандербейлпарк, 2025

математическая статистика, статистическое моделирование, информационные критерии AIC/AICc.

1. Введение

Электроэнергетика — стратегически важная отрасль страны, потому что является базисом для функционирования всех без исключения отраслей экономики — от горно-металлургических и обрабатывающих производств до сферы услуг и цифровой экономики. Нарушения в работе больших систем электроэнергетики (БСЭ) обладают мультипликативным эффектом, вызывая каскадные сбои в смежных отраслях экономики Государства. Следовательно, устойчивое и надежное энергоснабжение есть необходимое условие национальной экономической безопасности. Стратегическое планирование позволяет прогнозировать будущие потребности в мощности и электроэнергии, синхронизируя их с долгосрочными программами промышленного и регионального развития.

Исследование и разработка качественных математических моделей для прогнозирования потребления электроэнергии является критически важной стратегической задачей Государства, находящейся на стыке экономики, энергетики и компьютерных наук.

Управление режимами работы Единой энергетической системы (ЕЭС) России, являющейся одной из крупнейших в мире синхронных зон, требует обеспечения баланса между производством и потреблением электроэнергии в реальном времени. Несоответствие между этими величинами приводит к изменениям частоты тока, что угрожает устойчивости и надежности всей энергосистемы. В условиях нестационарности и стохастичности спроса, обусловленного множеством экономических, технологических и поведенческих факторов, точное прогнозирование потребления становится императивом для эффективного и безопасного функционирования отрасли.

Ключевые аспекты важности математического моделирования заключаются в оптимизации оперативного диспетчерского управления и режимов работы генерирующего оборудования. Краткосрочное и оперативное прогнозирование (от нескольких часов до 30 минут) позволяет Системному оператору (СО ЕЭС) формировать оптимальный суточный график нагрузки. Это минимизирует затраты на регулирование частоты тока, сокращает использование неэкономичных пиковых мощностей и повышает коэффициент использования установленной мощности (КИУМ) базовых станций (АЭС, ГЭС, ТЭС). Математические модели, учитывающие временные ряды электропотребления, метеорологические данные (температура, освещенность) и календарные факторы (дни недели, праздники), являются основой для принятия управленческих решений.

Другим важным аспектом разработки моделей является повышение экономической эффективности на оптовом и розничном рынках электроэнергии и мощности (ОРЭМ). Участники рынка — генерирующие и энергосбытовые организации, а также крупные промышленные предприятия — используют модели для формирования прогнозных заявок необходимых объёмов электроэнергии. Ошибка прогноза приводит предприятия к финансовым потерям вследствие покупки или продажи электроэнергии по невыгодным балансирующим ценам. Однако более точные математические модели прогнозирования напрямую конвертируются в снижение коммерческих рисков и повышение маржинальности.

Таким образом, выбор лучшей математической модели прогнозирования электропотребления является одной из фундаментальных задач научного исследования.

Теоретики информации не верят в понятие эталонных (идеальных, истинных) математических моделей. Модели, по определению, являются лишь приближением (аппроксимацией) к неизвестной реальности или истине — не существует истинных моделей, идеально описывающих полную реальность (имеющиеся данные) на 100%. Однако, британский статистик Джордж Бокс, известный учёный в области планирования эксперимента сделал знаменитое философское заявление в 1976 году: «Все модели неверны, но некоторые из них полезны». Поясняя фразу учёного, отметим, что модель полезна, если её прогнозы, несмотря на неизбежную ошибку, позволяют принимать более эффективные решения, чем те, что были бы приняты без неё. Например, модель со средней абсолютной ошибкой в 3% может быть бесполезной для оперативной диспетчеризации, но чрезвычайно полезной для стратегического планирования мощностей.

В процессе разработки прогнозных систем исследователь сталкивается с ситуацией, когда множество разработанных моделей-кандидатов

{M1,M2,...,Mn}, основанных на различных математических парадигмах (от классических статистических до современных алгоритмов машинного обучения), претендуют на решение одной и той же задачи прогнозирования электропотребления. Проблема выбора наилучшей модели (model selection) прогнозирования является нетривиальной задачей оптимизации, где целевой функцией выступает не просто точность, оцененная на имеющихся исторических данных, а обобщающая способность модели на независимых тестовых выборках (т.е. на совершенно новых данных для модели).

Актуальность анализа качества моделей заключается не в поиске «безошибочной» (идеально точной) модели, а в строгой и многокритериальной оценке их «полезности» для конкретных задач в

условиях энергетического перехода. Мы оцениваем, насколько та или иная «ошибочная» модель помогает минимизировать риски, оптимизировать затраты и обеспечить надежность энергосистемы, постоянно помня о её принципиальных ограничениях.

2. Обзор литературы

В настоящее время поток публикаций о проблеме выбора лучшей модели, а также о способах оценки качества моделей не уменьшается.

В статье [1] авторы подробно обсуждают проблемы выбора модели (model selection), подчеркивая важность воспроизводимости и объективного сравнения моделей. Авторы выделяют, что оценка качества должна выходить за рамки точечных прогнозов и стандартных метрик (вроде MAPE или RMSE) и включать вероятностные прогнозы, оценку неопределенности и операционную ценность модели для конечного пользователя. В данной работе исследователи предлагают методологию, к которой необходимо стремиться при оценке качества моделей, однако для полноценной вероятностной оценки и проверки воспроизводимости требуются обширные, безупречно аннотированные и публично доступные наборы данных. В реальности же данные энергетических организаций и крупных промышленных предприятий часто являются коммерческой тайной, которая запрещает публикацию результатов исследования, содержат пропуски и аномалии, а их предобработка (feature engineering, разведочный анализ данных) сама по себе попадает под Соглашение о неразглашении информации (NDA), которое компании не раскрывают.

В научном труде [2] исследователь систематизирует принципы оценки и выбора моделей, которые в полной мере применимы и к прогнозированию электропотребления. Автор критикует так называемые «конкурсы» математических моделей, где «побеждает» алгоритм с наименьшей ошибкой на одном единственном наборе данных, и настаивает на строгих процедурах статистического тестирования, включающих тесты на остатках (Residuals) модели (например, на автокорреляцию остатков). Исследование [3] предлагает углубленное эмпирическое сравнение традиционных статистических методов и моделей глубокого обучения для прогнозирования спотовых цен на электроэнергию. В центре внимания — проблема выбора между сложными и простыми моделями. Авторы демонстрируют, что тщательная предварительная подготовка данных, включающая анализ сезонности, обработку выбросов и тестирование на множестве периодов, показывает, что часто правильно настроенные традиционные модели могут быть столь же эффективны, что и глубокие нейронные сети, что ставит под сомнение необходимость всегда выбирать самую сложную архитектуру модели.

В работе [4] рассматривается проблема выбора модели в контексте больших данных. Авторы предлагают новую модель, учитывающую лаги в значениях переменных, и сравнивают ее с традиционными подходами. Статья подчеркивает, что при использовании больших данных критерии для выбора модели смещаются в сторону её способности находить сложные, неочевидные зависимости, а оценка качества модели должна проводиться на репрезентативных и достаточно больших тестовых наборах, чтобы гарантировать, что улучшения являются статистически значимыми, а не случайными.

Корейские ученые из Университета Инха и команда по разработке авионики и радиолокационных систем технологической компании Hanwha Systems в своей научной работе [5] провели сравнительный анализ трёх однослойных (CNN, GRU, LSTM) и семи гибридных многослойных моделей глубокого обучения для прогнозирования нагрузки в энергосистемах, используя данные по потреблению и погоде с острова Чеджу и систематически изменяя гиперпараметры (эпохи, функции активации). Предложенный авторами комплексный подход с проведением 180 экспериментов позволил получить модель LSTM-CNN с ошибкой 0,231%. Однако, ограниченность набора моделей только CNN, GRU, LSTM и их гибридами, а также фиксированная скорость обучения, может снижать обобщающую способность выводов.

Нидерландские учёные в статье [6] предложили модификацию байесовского предсказательного распределения, которая позволяет динамически переключаться между моделями в процессе последовательного прогнозирования. Этот подход призван решить «дилемму AIC–BIC», сочетая в себе преимущества обоих методов. Однако, в работе авторов нет глубокого обсуждения того, насколько результаты робастны к изменению априорной информации.

В статье [7] исследователи предложили гибридную математическую модель для долгосрочного прогнозирования нагрузки (LTLF) в Иордании, сочетающую параметрический метод множественной линейной регрессии и модель авторегрессии — проинтегрированного скользящего среднего (ARIMA) на основе исторических данных за 1990–2022 годы. Модель учитывает влияние ключевых факторов, таких как ВВП, население, цены на нефть и генерацию возобновляемых источников энергии. Предложенная авторами модель демонстрирует высокую точность, особенно для ARIMA с ошибками (MAPE) 2,55%. Однако, ключевым недостатком является зависимость прогнозов от гипотетических сценариев (оптимистичного,

пессимистичного и нормального), что ограничивает устойчивость модели в условиях неопределённости будущих макроэкономических и политических изменений. Кроме того, использование дамми-переменных (фиктивных переменных) и нормализация данных могут упрощать сложные нелинейные взаимосвязи, потенциально снижая адаптивность модели к резким структурным сдвигам.

Исследователи из Китая и Тайваня в своей статье [8] предлагают мощный теоретический инструмент для анализа расстояния Куллбак-Лейблера и информации Фишера в сложных моделях с зависимостями, что открывает путь к обоснованному использованию информационного критерия Акаике (AIC) и других методов. Однако практическое применение требует преодоления вычислительных и теоретических сложностей.

Ученые из Греции в своем научном труде [9] предлагают метод прогнозирования энергопотребления дата-центров (DC) на основе погодных условий с использованием линейной регрессии и данных с IoT-сенсоров платформы FIESTA-IoT. Прогноз позволяет оптимизировать работу систем охлаждения и электропотребления, снижая операционные затраты (OPEX). Использование платформы FIESTA-IoT позволяет автоматически собирать и обрабатывать данные в реальном времени. Однако, линейная регрессия не учитывает нелинейные зависимости, которые могут существовать в реальных условиях электропотребления в датацентрах. Авторы используют данные за короткий период, что может снижать надёжность долгосрочных прогнозов. Следует также отметить, что модель, обученная на данных одного дата-центра, может не подходить для других географических зон или различных типов датацентров.

Исследователи из Австралии и Японии в своем научном труде «Критерий AIC и симметризация расстояния Кульбака–Лейблера» [10] расширили теоретическую базу AIC и предложили новые пути для построения критериев выбора модели - три новых симметричных функции для сравнения моделей на основе симметризации расстояния Кульбака–Лейблера (K-L) с помощью: среднего арифметического; среднего геометрического; среднего гармонического. Новая скорректированная версия критерия Акаике AIC*c для линейной регрессии в малых выборках, основанная на симметричной дивергенции K-L (среднем арифметическом), которая отличается от известной формулы AICc. K-L-расстояние измеряет, насколько распределение P отличается от Q, или количество информации, теряемой при аппроксимации P с помощью Q.

В статье [11] автор утверждает, что, несмотря на внешнее сходство формул, AIC и BIC преследуют разные философские цели и основаны на разных предположениях. Главное предложение заключается в том, что выбор между AIC и BIC должен определяться исследовательским вопросом:

если цель — прогноз, то есть найти модель, которая лучше всего предскажет новые, будущие данные из той же выборки данных, следует использовать AIC; если цель — установление истины, то есть найти истинную генеративную модель (при условии, что она существует и включена в набор кандидатов), то следует использовать BIC. Однако, оба критерия позволяют выбрать лучшую модель только из предложенного исследователем набора моделей. Если истинная модель или хорошая прогнозирующая модель не включена в этот набор, результат будет неоптимальным. «Лучшая из плохих моделей» все равно остается плохой моделью. Остается неизвестным, адекватны ли модели, которые попали в набор.

Выбор лучшей модели и критерия качества из потенциально большого класса моделей-кандидатов и критериев - вопрос, который занимает центральное место в задачах регрессии и моделировании временных рядов электропотребления [12, 13].

Решение задач регрессии на основе метода наименьших квадратов (МНК) представляют собой набор статистических тестов для оценки взаимосвязей между зависимой переменной (электропотребление) и несколькими независимыми переменными (факторами). МНК в обратных задачах регрессии широко используются для прогнозирования, поскольку по имеющимся данным можно провести аппроксимацию и интерпретировать результаты.

В своей статье [14] авторы предлагают сравнить линейные и нелинейные модели для анализа данных профиля нагрузки электроэнергии с использованием трёх методов оценки: ANOVA, AIC и BIC. Они применяют шесть моделей, включая полиномиальные регрессии 2-й, 3-й и 4-й степени, линейную регрессию, множественную линейную регрессию и модели с взаимодействием переменных. Однако низкое значение скорректированного R2 = 0.76% говорит о слабой объясняющей способности даже лучшей модели.

Сложность определения наилучшего критерия качества для оценки регрессионной модели прогноза электропотребления заключается в многокритериальности и контекстно-зависимой природе самой задачи, которая находится на стыке статистического моделирования и прикладной электроэнергетике.

3. Разработка набора адекватных математических моделей-кандидатов

Для моделирования и прогнозирования электропотребления на сутки вперёд с помощью классического МНК были разработаны 7 априорных адекватных регрессионных уравнения различной

структуры (полиномиальные и линейная) по данным процесса потребления электроэнергии предприятия (количество строк в выборке данных (N) = 25). Структура и вид математических моделей-кандидатов представлены в таблице 1.

Структура моделей:

Таблица 1. – Структура моделей набора

№	Структура адекватных математических моделей в сете
1	$X := \text{augment}(X0, X1, X2, X3)$
2	$X := \text{augment}\left[X0, X1, X2, X3, X1^2, X2^2, \overrightarrow{(X1 \cdot X2)}, \overrightarrow{(X1 \cdot X3)}, \overrightarrow{(X2 \cdot X3)}\right]$
3	$X := \text{augment}\left[X0, X1, X2, X3, X1^2, \overrightarrow{(X1 \cdot X2)}, \overrightarrow{(X1 \cdot X3)}, \overrightarrow{(X2 \cdot X3)}\right]$
4	$X := \text{augment}\left[X0, X1, X2, X3, X2^2, \overrightarrow{(X1 \cdot X2)}, \overrightarrow{(X1 \cdot X3)}, \overrightarrow{(X2 \cdot X3)}\right]$
5	$X := \text{augment}\left[X0, X1, X2, X3, X1^2, X1^3, \overrightarrow{(X1 \cdot X2)}, \overrightarrow{(X1 \cdot X3)}, \overrightarrow{(X2 \cdot X3)}\right]$
6	$X := \text{augment}\left[X0, X1, X2, X3, X2^2, X2^3, \overrightarrow{(X1 \cdot X2)}, \overrightarrow{(X1 \cdot X3)}, \overrightarrow{(X2 \cdot X3)}\right]$
7	$X := \text{augment}\left[X0, X1, X2, X3, X1^2, X1^3, X2^2, \overrightarrow{(X1 \cdot X2)}, \overrightarrow{(X1 \cdot X3)}, \overrightarrow{(X2 \cdot X3)}\right]$

Разработанные регрессионные уравнения с оценёнными коэффициентами по формуле (1), согласно структуре в таблице 1 представлены ниже:

$$Y1 = 3.122 \cdot 10^3 + 60.308 \cdot X1 + 5.692 \cdot X2 - $$
$$-499.251 \cdot X3; _(Линейная_Модель_№1_k = 4).$$

$$Y2 = 7.341 \cdot 10^3 + 105.108 \cdot X1 - 309.164 \cdot X2 -$$
$$-2.734 \cdot 10^3 \cdot X3 - 3.343 \cdot X1^2 + 5.395 \cdot X1^3 + 0.423 \cdot X1 \cdot X2 +$$
$$+34.643 \cdot X1 \cdot X3 + 85.978 \cdot X2 \cdot X3; _(Неполный_полином$$
$$_3й_степени_Модель_№2_k = 9).$$

$$Y3 = 4.619 \cdot 10^3 + 100.345 \cdot X1 - 66.841 \cdot X2 -$$
$$-2.442 \cdot 10^3 \cdot X3 - 3.818 \cdot X1^2 + 1.126 \cdot X1 \cdot X2 +$$
$$+35.666 \cdot X1 \cdot X3 + 72.273 \cdot X2 \cdot X3; _(Неполниый_полином_$$
$$_2й_степени_Модель_№3_k = 8).$$

$$Y4 = 1.426 \cdot 10^4 + 2.832 \cdot X1 - 849.935 \cdot X2 -$$
$$-3.815 \cdot 10^3 \cdot X3 + 16.482 \cdot X2^2 + 1.107 \cdot X1 \cdot X2 +$$
$$+35.187 \cdot X1 \cdot X3 + 130.7 \cdot X2 \cdot X3; _(Неполный_полином_$$
$$_2й_степени_Модель_№4_k = 8).$$

$$Y5 = 4.547 \cdot 10^3 + 123.723 \cdot X1 - 68.245 \cdot X2 -$$
$$-2.429 \cdot 10^3 \cdot X3 - 6.496 \cdot X1^2 + 0.069 \cdot X1^3 + 1.496 \cdot X1 \cdot X2 +$$
$$+31.073 \cdot X1 \cdot X3 + 73.948 \cdot X2 \cdot X3; _(Неполный_полином_$$
$$_3й_степени_Модель_№5_k = 9).$$

$$Y6 = -4.234 \cdot 10^4 - 3.929 \cdot X1 + 7.087 \cdot 10^3 \cdot X2 -$$
$$-3.579 \cdot 10^3 \cdot X3 - 350.754 \cdot X2^2 + 5.603 \cdot X2^3 + 1.401 \cdot X1 \cdot X2 +$$
$$+38.478 \cdot X1 \cdot X3 + 121.621 \cdot X2 \cdot X3; _(Неполный_полином_$$
$$_3й_степени_Модель_№6_k = 9).$$

$$Y7 = 7.175 \cdot 10^3 + 122.289 \cdot X1 - 300.224 \cdot X2 -$$
$$-2.712 \cdot 10^3 \cdot X3 - 5.354 \cdot X1^2 + 0.051 \cdot X1^3 + 5.173 \cdot X2^2 +$$
$$+0.727 \cdot X1 \cdot X2 + 31.27 \cdot X1 \cdot X3 + 86.658 \cdot X2 \cdot X3; _$$
$$(Неполный_полином_3й_степени_Модель_№7_k = 9).$$

где Y1, Y2, Y3, Y4, Y5, Y6, Y7 – значения зависимых переменных электропотребления, кВт·ч; X1– признак, характеризующий время, сутки; X2– признак - температура воздуха, ^{0}C; X3– признак – день недели (1-рабочие дни, 0-выходные дни); k – количество коэффициентов в уравнении регрессии; N – количество строк в матрице признаков X и матрице-столбце зависимой переменной Y (исходная выборка с фактическими данными электропотребления).

Коэффициенты регрессии рассчитаны по формуле (1) [15, 16]

$$B = (X^T \cdot X)^{-1} \cdot X^T \cdot Y; \quad (1)$$

где B – коэффициенты регрессии; X^T – транспонированная матрица признаков; Y – исходные (фактические) данные потребления электроэнергии.

Оценка адекватности всех семи регрессионных уравнений в наборе проверена по следующим формулам.

Расчётные значения целевой переменной (*YR*), дисперсия адекватности (*Dad*), средняя арифметическая целевой переменной (*YSR*), дисперсия целевой переменной (*DY*), расчётное значение целевой переменной (*FR*) были рассчитаны по формулам (2 - 6):

$$YR := X \cdot B, \quad (2)$$

$$Dad := \frac{\sum (Y - YR)^2}{N - k}. \quad (3)$$

$$YSR := \frac{\sum Y}{N}. \quad (4)$$

$$DY := \frac{\sum (Y - YSR)^2}{N - 1}. \quad (5)$$

$$FR := \frac{DY}{Dad}. \quad (6)$$

Расчетное значение *F*-критерия Фишера сравнивали с табличным значением (*F*) (7):

$$F := qF(0.95, N - 1, N - k)., \quad (7)$$

В результате проведённой проверки на адекватность разработанных уравнений регрессии был сделан следующий вывод: в связи с тем, что **во всех семи моделях** $FR > F$, следовательно уравнения регрессии признаны адекватными исходным

фактическим данным электропотребления на уровне значимости 0,05, что соответствует доверительной вероятности
p = (1-0,05)·100=95,0 %.

4. Анализ качества разработанных математических моделей и выбор лучшей

Обязательным условием в процессе моделирования электропотребления является проверка разработанных моделей на качество.

Проведён тест на значимость коэффициента корреляции с использованием *t*-критерия Стьюдента.

Вычислено расчётное значение *t*-критерия Стьюдента по формуле (8), которое затем сопоставлено с табличным значением (9)

$$tr := |r| \cdot \sqrt{\frac{N-k}{1-r^2}}. \quad (8)$$

$$t := qt(0.975, N-k) \quad (9)$$

Далее проведён расчёт критического значения коэффициента корреляции (*rcr*) по формуле (10)

$$rcr := \sqrt{\frac{t^2}{t^2+N-k}} \quad (10)$$

В результате проведённых расчётов можно сделать вывод: в связи с тем, что $tr > t$ и $r > rcr$, все коэффициенты корреляции моделей признаны статистически значимыми с 95-% надёжностью. На этом же основании все уравнения регрессии в наборе могут быть признаны качественными и адекватными фактическим значениям электропотребления *(Y).*

В настоящем исследовании использовались следующие критерии качества моделей для проведения анализа: коэффициент корреляции (*r*), коэффициент детерминации (R^2), скорректированный $R^2_{adjusted}$, среднеквадратическая ошибка (*MSE*), квадратный корень из среднеквадратической ошибки (*RMSE*), стандартная ошибка (*SD*) или квадратный корень из дисперсии, средняя абсолютная ошибка моделирования (*MAE*), стандартная ошибка (*SE*), максимальное правдоподобие (Max Likelihood, *MLE*), максимальное логарифмическое правдоподобие (Max Log Likelihood, *MLLE*), а также информационные критерии Акаике (*AIC, AICc*), информационный критерий Байеса (*BIC*), Mallows'Cp (*Cp*).

4.1. Обоснование выбора критерия для определения лучшей модели прогнозирования

Мы разделили критерии на три группы: 1. Критически важные для прогнозирования (информация Фишера). 2. Для углубленного анализа на этапе моделирования электропотребления (информационно-теоретический подход в сравнении с реальностью). 3. Неподходящие для оценки качества моделей, но полезные при проверке адекватности.

Группа 1: Критически важные для прогнозирования электропотребления (основа для выбора модели). Эти критерии напрямую оценивают ошибку прогноза на данных, которые модель не использовала на этапе обучения.

Среднеквадратическая ошибка (MSE) и ее производные (RMSE, MAE)

Это прямое измерение того, насколько точно прогнозы соответствуют реальным значениям. Единица измерения электропотребления *кВт·ч.*, следовательно, *MSE* (или его корень *RMSE*) рассчитывает численную оценку средней ошибки в этих единицах. Лучшая модель — та, у которой MSE на тестовой выборке минимальна. RMSE (Root MSE): полезнее при интерпретации результатов модели, так как выражается в тех же единицах, что и исходные данные (*кВт·ч.*,). Критерий MAE (Mean Absolute Error) менее чувствителен к выбросам, чем MSE.

Средняя абсолютная процентная ошибка (MAPE - Mean Absolute Percentage Error)

Электропотребление имеет ярко выраженные суточные и недельные циклы. Отклонение прогноза от факта в 100 МВт в ночное время — это большая ошибка, а в часы пиковой нагрузки — допустимая. Критерий MAPE показывает ошибку в процентах, что позволяет оценить относительную точность модели в разные периоды. Критерий MAPE легко интерпретируется руководством компаний и диспетчерами (например, «наша средняя ошибка прогноза составляет 2%»).

Среди недостатков можно отметить тот факт, что критерий MAPE может быть неопределенным или давать бесконечно большие значения, когда реальное потребление близко к нулю (маловероятно для электропотребления всей сети, но возможно для отдельного малого объекта).

Итог по Группе 1: Наш основной метод выбора модели должен основываться на сравнении MSE/RMSE и MAPE на тестовой (валидационной) выборке. Это прямо отвечает на вопрос: «Какая модель даст более точный прогноз на сутки вперед?».

Группа 2: Для углубленного анализа на этапе моделирования электропотребления.

Эти критерии полезны для понимания модели и выбора между моделями близкой точности, но они не должны быть основными.

Информационные критерии Акаике и Байеса (AIC, AICc, BIC)

Это критерии, которые оценивают компромисс между качеством подгонки модели (правдоподобие) и её сложностью (количество параметров). Сложная модель может идеально подойти к фактическим данным, но плохо прогнозировать будущее (что означает переобучение). AIC/BIC штрафуют модель за сложность. Данные критерии являются отличными инструментами для предварительного отбора моделей. Если у нас есть 7 моделей-кандидатов, мы можем быстро отсеять явно переобученные модели, выбрав модель с наименьшим значением AIC/BIC. AICc (скорректированный AIC) лучше для небольших выборок. AIC стремится выбрать более сложные модели, чем BIC. Для прогнозирования часто предпочтительнее AIC, так как он асимптотически эквивалентен выбору модели с наилучшим прогнозом. AIC/BIC вычисляются на обучающих данных.

Mallow's Cp

Данный критерий используется в основном в регрессионном анализе для выбора переменных. Для задач прогнозирования временных рядов AIC/BIC применяются чаще и являются более стандартными.

Правдоподобие (Likelihood, MaxLikelihood, Max Log-Likelihood)

Являются основой для вычисления AIC/BIC. Само по себе значение правдоподобия не интерпретируемо при сравнении моделей с разным числом параметров. Более сложная модель всегда будет иметь большее правдоподобие, что ведет к переобучению. Log-Likelihood — это то, что напрямую используется в формулах AIC/BIC.

Группа 3: Неподходящие для оценки качества моделей, но полезные при проверке адекватности.

Адекватность по F-критерию Фишера

Этот критерий из области проверки статистических гипотез. Он отвечает на вопрос «Значима ли наша модель в целом?» или «Значима ли группа коэффициентов?». Почти любая сложная модель для электропотребления будет признана «значимой» по F-критерию. Он не говорит ничего о прогнозной точности и не помогает выбрать между двумя адекватными моделями.

Разработанные статистические аппроксимирующие уравнения регрессии, использующиеся для представления процесса потребления электроэнергии, который сгенерировал данные, не могут быть эталонными (описывающие фактические данные с точность 100%). Следовательно, некоторый объем информация будет потерян. В таком случае, оценить относительный объём информации, потерянной конкретной аппроксимирующей моделью можно с помощью информационного критерия Акайке (AIC) (11) который является важнейшим в теоретико-информационном подходе Кульбак-Лейблера:

$$AIC = N \cdot (\ln(\sigma^2) + 1) + 2 \cdot (k + 1). \quad (11)$$

где N – количество строк в матрицах данных; σ^2 – дисперсия адекватности, k – число коэффициентов в математической модели.

Интерпретируя результаты расчета критерия AIC можно отметить, что чем меньше информации теряет модель, тем выше качество анализируемой аппроксимирующей функции. Учитывая набор моделей-кандидатов, разработанных на основе данных, сгенерированных процессом электропотребления, предпочтительной является модель с минимальным значением критерия Акаике AIC.

Результаты оценки адекватности и качества уравнений регрессии в наборе представлены в таблице 2 и 3, где М1 – М7 разработанные модели.

Таблица 2. – Расчет критериев оценки адекватности и качества

№	MSE	RMSE	MAE	FR>F	r	R^2	R^2adj
M1	76940	277.383	204.754	+	0.865	0.749	0.698
M2	33810	**183.874**	109.84	+	0.943	0.89	0.823
M3	34820	186.588	116.267	+	0.941	0,886	0.818
M4	48370	219.93	161.015	+	0.918	0.842	0.747
M5	34610	186.029	111.807	+	0.942	0.887	0.819
M6	41170	202.911	137.648	+	0.93	0.866	0.785
M7	33700	183.565	107.833	+	0.943	0.89	0.824

Таблица 3. – Расчет информационных критериев

№	MLE	MLLE	AIC	AICc	BIC	Mallow Cp
M1	76940	-176.108	291.270	294.428	297.30	25.005
M2	33810	-165.83	280.712	296.427	636.84	25.004
M3	34820	-166.196	**279.445**	**291.445**	361.36	25.004
M4	48370	-170.306	287.665	299.662	369.58	25.003
M5	34610	-166.121	281.295	297.009	293.48	25.004
M6	41170	-168.293	285.638	301.353	297.82	25.003
M7	33700	-165.788	282.629	302.936	366.98	25.004

В таблицах 2 и 3 можно заметить одинаковые значения при расчёте среднеквадратической ошибки и максимального правдоподобия. Почему

так произошло? Это происходит потому, что при нормальном распределении ошибок минимизация MSE (Mean Squared Error) математически эквивалентна максимизации правдоподобия (Maximum Likelihood Estimation, MLE). Мы получили тот же результат, используя два разных, но в данном контексте эквивалентных, принципа. Это косвенно подтверждает, что все наши модели в сете корректны.

Для регрессионных задач с непрерывной целевой переменной минимизация MSE является стандартом т.к. она является MLE-оценкой для нормального шума.

Однако нам необходимо больше доказательств, для выбора лучшей модели из сета, когда все модели-кандидаты в сете конкурентоспособные, т.е. адекватные и качественные, согласно таблицам 2 и 3.

Учитывая, что информационный критерий Акаике ничего не говорит об абсолютном качестве уравнения регрессии, а только об **относительном качестве** по сравнению с другими уравнениями из разработанного сета, проведён расчет относительного различия в значения AIC.

4.2. Относительные различия в значениях AIC (Δ_i)

Для того, чтобы проверить уравнения регрессии на абсолютное качество, необходимо рассчитать различия в остатках значений ΔAIC и ΔAICc по формуле (12)

$$\Delta_i = AIC_i - AIC_{\min} \tag{12}$$

где Δ_i – относительное различие значений критерия AIC; AIC_i – значение AIC модели, AIC_{min} – самое минимальное значение AIC в наборе моделей-кандидатов. Уравнение регрессии, оценённое как лучшее в наборе моделей-кандидатов, будет иметь $\Delta_i = 0$ или, другими словами, чем больше значение Δ_i, тем меньше вероятность, что подобранная модель регрессии является лучшей моделью в информационно-теоретическом подходе Кульбак-Лейблера (K-L расстояние).

С. Куллбек и Р. А. Лейблер в 1951 году опубликовали работу [9], в которой количественно определили значение понятия «информация» в связи с концепцией статистики Фишера. Их знаменитый научный труд, названный *информацией Кульбака-Лейблера (K-L)*, является фундаментальной величиной в науке и связан с концепцией энтропии Больцмана. Информацию *K-L* можно представить как дивергенцию (расстояние, разность) между полной реальностью (фактическими данными зависимых и независимых переменных) и функцией аппроксимации [17, 18].

В нашей задаче энтропия характеризует сложность профиля потребления электроэнергии, а *K-L-расстояние* измеряет, насколько хорошо наша разработанная регрессионная модель эту сложность воспроизводит. Минимизация *K-L-расстояния* – это поиск модели, которая наилучшим образом описывает форму суточной кривой электропотребления.

Нам необходимо выбрать из числа разработанных и адекватных семи моделей-кандидатов такую модель, которая минимизирует расстояние Кульбак-Лейблера, или другими словами, имеет минимальную потерю информации. Модель №3 имеет самое минимальное значение в наборе моделей по критериям AIC и AICc. Поэтому рассчитаем Δi по каждой модели (13)

По критерию AICc:

$$\begin{aligned}
\Delta_1 &= 294.428 - 291.445 = 2.983_Модель_1\\
\Delta_2 &= 296.427 - 291.445 = 4.982_Модель_2\\
\Delta_3 &= 291.445 - 291.445 = 0_Модель_3\\
\Delta_4 &= 299.665 - 291.445 = 8.22_Модель_4\\
\Delta_5 &= 297.009 - 291.445 = 5.564_Модель_5\\
\Delta_6 &= 301.353 - 291.445 = 9.908_Модель_6\\
\Delta_7 &= 302.936 - 291.445 = 11.491_Модель_7
\end{aligned} \tag{13}$$

Модель №3 имеет значение $\Delta_i = 0$.

По критерию AIC:

$$\begin{aligned}
\Delta_1 &= 291.27 - 279.445 = 11.825_Модель_1\\
\Delta_2 &= 280.712 - 279.445 = 1.267_Модель_2\\
\Delta_3 &= 279.445 - 279.445 = 0_Модель_3\\
\Delta_4 &= 287.665 - 279.445 = 8.22_Модель_4\\
\Delta_5 &= 281.295 - 279.445 = 1.85_Модель_5\\
\Delta_6 &= 285.638 - 279.445 = 6.193_Модель_6\\
\Delta_7 &= 282.629 - 279.445 = 3.184_Модель_7
\end{aligned}$$

Модель №3 имеет значение $\Delta_i = 0$.

Особенностью информационно-теоретического подхода является то, что он ранжирует уравнения в сете, позволяя, также, сделать некоторые выводы о других моделях, которые тоже могут иметь ценность для исследователей [19]. Мы упорядочили Δ_i от наименьшего к наибольшему, и такой порядок моделей нам показал, насколько функции полезны при аппроксимации в информационно-теоретическом подходе Кульбака-Лейблера [20]. Ранжирование моделей представлено в таблицах 4 и 5.

Таблица 4 .– Ранжирование моделей по Δ_i *no AICc*

Ранг	Δ_i	№ модели
1	0	3
2	2.983	1
3	4.982	2
4	5.564	5
5	8.22	4
6	9.908	6
7	11.491	7

Таблица 5. – Ранжирование моделей по Δ_i *no AIC*

Ранг	Δ_i	№ модели
1	0	3
2	1.267	2
3	1.85	5
4	3.184	7
5	6.193	6
6	8.22	4
7	11.825	1

В таблице 6 представлены значения Δ_i и уровни эмпирической поддержки модели.

Таблица 6. – Различия AIC (Δ_i) и уровни эмпирической поддержки модели

Δ_i	Уровни эмпирической поддержки модели
0-2	Существенный уровень
4-7	Значительно низкий уровень
>10	Нет поддержки модели

Согласно уровням эмпирической поддержки функций аппроксимации авторами решено было выбрать для второго этапа оценки качества модели и выбора лучшей только такие модели, у которых значение Δ_i от 0 до 2 – это Модель №2, Модель №3 и Модель №5, так как именно они сильно конкурируют за позицию лучшей аппроксимирующей модели [19].

Мы предлагаем выбирать лучшую математическую модель на первом этапе по относительным различиям критериев Акайке AIC и AICc, а на втором этапе по критерию RMSE.

Обратившись к результатам в таблице 3 мы видим, что наименьшее значение RMSE у Модели №2 (183.874), следовательно именно эта модель будет отобрана для дальнейшего прогнозирования электропотребления. В доказательство выбора лучшей модели, отметим, что Модель №2 имеет самые минимальные значения по MAE и высокий коэффициент корреляции, а также одно из лучших значений MSE и MLLE. Расчетное значение F-критерия Фишера-Снедекора в 2,7 раза больше табличного значения F-статистики, что говорит об адекватности выбранного уравнения регрессии. Отметим, что Модель №7 имеет значения критериев MSE, RMSE и MAE лучше, чем Модель №2, однако, Модель №7 не была отобрана на первом этапе анализа качества.

В результате проведенной дополнительной проверки на адекватность и качество разработанной математической модели (табл. 2 и 3) можно сделать вывод о том, что прогнозным оценкам моделей можно доверять.

5. Прогнозирование электропотребления на сутки вперед и обсуждение результатов

Для расчёта оценок прогнозирования электропотребления на сутки вперед мы воспользовались разработанной регрессионной моделью №2, которую выбрали по результатам анализа качества на двух этапах.

Ошибка прогнозирования и доверительный коридор интервала ошибок были рассчитаны по формулам (14 и 15)

$$D := X \cdot (X^T \cdot X)^{-1} \cdot X^T. \tag{14}$$

$$S_{i,i} := t \cdot \sqrt{Dad \cdot (1 + D_{i,i})}. \tag{15}$$

Мы создали вектор значений признаков в прогнозной точке (26 сутки), представленный на рисунке 1.

τ= 26 (X1 – Время, сутки) Следующая прогнозная точка.

A= 23.09 (X2 – Температура окружающей среды, 0C).

W=0 (X3 – День недели: 0 -рабочий, 1-выходной).

$\mathbf{X} := \mathrm{augment}\left[\mathbf{X0}, \mathbf{X1}, \mathbf{X2}, \mathbf{X3}, \mathbf{X1}^2, \mathbf{X2}^2, \overrightarrow{(\mathbf{X1}\cdot\mathbf{X2})}, \overrightarrow{(\mathbf{X1}\cdot\mathbf{X3})}, \overrightarrow{(\mathbf{X2}\cdot\mathbf{X3})}\right]$ Структура Модели №2

$\mathbf{XP} := \left(1 \;\; \tau \;\; A \;\; W \;\; \tau^2 \;\; A^2 \;\; \tau\cdot A \;\; \tau\cdot W \;\; A\cdot W\right)$ Вектор независимых переменных в прогнозной точке

$$\mathbf{XP}^{\mathbf{T}} = \begin{pmatrix} 1 \\ 26 \\ 23.09 \\ 0 \\ 676 \\ 533.148 \\ 600.34 \\ 0 \\ 0 \end{pmatrix} \qquad \mathbf{B} = \begin{pmatrix} 7.341 \times 10^3 \\ 105.108 \\ -309.164 \\ -2.734 \times 10^3 \\ -3.343 \\ 5.395 \\ 0.423 \\ 34.643 \\ 85.978 \end{pmatrix}$$

Рис. 1. Вектор независимых переменных в прогнозной точке

На рисунке 1 представлен вектор *XP* – вектор факторов в прогнозной точке, согласно выбранному виду математической модели №2, а также оценённые коэффициенты выбранной регрессионной модели.

Подставляя значения факторов в прогнозной точке (новые данные) и оценённые коэффициенты в выбранную структуру математической модели №2, рассчитываем прогнозную оценку на сутки вперёд. Прогнозное значение (*YP*) электропотребления составило 3806 кВт.·ч. (16)

Выбранная регрессионная модель (16)

$$YP := B_1 + B_2 \cdot \tau + B_3 \cdot A + B_4 \cdot W + B_5 \cdot \tau^2 + B_6 \cdot A^2 + B_7 \cdot \tau \cdot A + B_8 \cdot \tau \cdot W + B_9 \cdot A \cdot W$$

Провели расчёт максимального и минимального значения прогнозной оценки по формулам (17) и (18)

$$YP\max := YP + S_{25,25} = 4404_кВт \cdot ч. \quad (17)$$

$$YP = 3806\ кВт.\cdot ч.$$

$$YP\max := YP - S_{25,25} = 3208_кВт \cdot ч. \quad (18)$$

Вычислены абсолютная и относительная ошибки прогнозирования:

- абсолютная ошибка прогнозирования равна 154 кВт.·ч.;
- относительная ошибка прогнозирования равна 3,64%.

Визуализация моделирования на сутки вперёд представлена на рисунке 2.

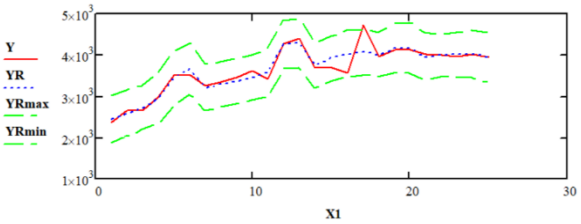


Рис. 2. Моделирование электропотребления на сутки вперёд и доверительный интервал коридора ошибок

Из рисунка 7 следует, что 95% всех фактических значений электропотребления попадают в доверительный интервал коридора ошибок. Модель корректно оценивает неопределенность. Выход одной точки (одного фактического значения электропотребления) за 95% доверительного коридора интервала ошибок соответствует нашим статистическим ожиданиям и подтверждает, что интервалы авторами не завышены.

6. Заключение

В заключении необходимо отметить, что в результате проведенного исследования и эксперимента:

- разработаны полезные, адекватные и качественные регрессионные модели, которые вошли в исходный набор моделей-кандидатов;
- проведен анализ расчетов оценок критериев качества моделей;
- показаны этапы выбора лучшей модели из сета, комбинируя информационные критерии Акаике и метрику RMSE;
- проведены расчеты прогнозного значения электропотребления на сутки вперед, с относительной ошибкой 3,64 %.

Таким образом, определение «наилучшего» критерия оценки качества модели — это не поиск единственного статистического показателя, а процесс выбора или конструирования метрики (или их ансамбля), который адекватно отражает экономические, технические и операционные требования конкретной задачи прогнозирования в электроэнергетике.

Список используемых источников

1. Hong T, Pinson P., Wang Y., Weron R., Yang D., Zareipour H. Energy forecasting: A review and outlook // IEEE Open Access Journal of Power and Energy, vol. 7, pp. 376-388, 2020.
2. Weron R. Electricity price forecasting: A review of the state-of-the-art with a look into the future // International Journal of Forecasting, vol. 30, no. 4, pp. 1030-1081, 2014.
3. Lago J., De Ridder F., & De Schutter B. Forecasting spot electricity prices: Deep learning approaches and empirical comparison of traditional algorithms// Applied Energy, vol. 221, pp. 386-405, 2018.
4. Wang P., Liu B., & Hong T. Electric Load Forecasting with Recency Effect: A Big Data Approach // International Journal of Forecasting, vol. 32, no. 3, pp. 585-597, 2016.
5. Jang J., Kim B., Kim I. Comparative Analysis of Deep Learning Techniques for Load Forecasting in Power Systems Using Single-Layer and Hybrid Models // Wiley. International Transactions on Electrical Energy Systems, Volume 2024, Article ID 5587728, 22 pages https://doi.org/10.1155/2024/5587728.
6. Van Erven T., Grünwald P., De Rooij S. Catching up faster by switching sooner: a predictive approach to adaptive estimation with an application to the AIC–BIC dilemma.// Journal of the Royal Statistical Society: Series B (Statistical Methodology) Volume 74, Issue 3, Jun 2012, Pages 361-622.
7. Momani M., Tashtush S.A., Shahrour R.J., Al Satari A.M. Modeling of Long-Term Load Forecast in Jordan Based on Statistical Techniques // Wiley, Journal of Electrical and Computer Engineering.

Volume 2024, Article ID 8255513, 16 pages https:// doi.org /10.1155/2024/8255513.

8. Fuh C., Kao M., Pang T. Kullback-Leibler Divergence and Akaike Information Criterion in General Hidden Markov Models // Reseacrh Gate, Preprint, DOI: 10.48550/arXiv.2303.07673, March 2023, ulr: https://www.researchgate.net/publication/369233662_Kullback-Leibler_Divergence_and_Akaike_Information_Criterion_in_General_Hidden_Markov_Models, дата посещения: 19.10.2025 г.
9. Smpokos G., Elshatshat M.A., Lioumpas A., Iliopoulos I. On the Energy Consumption Forecasting of Data Centers Based on Weather Conditions: Remote Sensing and Machine Learning Approach // Conference: 2018 11th International Symposium on Communication Systems, Networks and Digital Signal Processing (CSNDSP), July 2018, DOI: 10.48550/arXiv.1804.01754.
10. Seghouane A., Amari S. The AIC Criterion and Symmetrizing the Kullback–Leibler Divergence // IEEE TRANSACTIONS ON NEURAL NETWORKS, VOL. 18, NO. 1, JANUARY 2007.
11. Scott I. Vrieze. Model Selection and Psychological Theory: A Discussion of the Differences Between the Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC)//February 2012. Psychological Methods 17(2):228-243. DOI: 10.1037/a0027127.
12. McQuarrie A., Tsai C. Regression and Time Series Model Selection. 1998.
13. Hurvich C., Tsai C.. Regression and time series model selection in small samples. Biometrika (1989), 76, 2, pp. 297-307.
14. Keka I., Cico B.. Comparing Linear and Nonlinear Models for Load Profile Data Using ANOVA, AIC, and BIC// SAGE Open 15(12):22. March 2025. DOI: 10.1177/21582440251326389.
15. Кремер Н.Ш. Теория вероятностей и математическая статистика: учебник для студентов вузов, обучающихся по экономическим специальностям / Н.Ш. Кремер. – 3-е изд., перераб. и доп. – М.: ЮНИТИ-ДАНА, 2010. – 551 с. – (Серия «Золотой фонд российских учебников») с.439 – 441.
16. Дзгоев А.Э. Методы обработки и анализа данных для разработки предиктивных моделей [Электронный ресурс] : учебное пособие / А.Э. Дзгоев. — М.: РТУ МИРЭА, 2024 . — Электрон. опт. диск (ISO).ISBN 978-5-7339-2179-2.
17. Seghouane A., Amari S. Variants of the Kullback-Leibler divergence and their role in model selection / 14th IFAC Symposium on System Identification, pp. 826-831. Newcastle, Australia, 2006.
18. Kullback S. and Leibler R.A. On Information and Sufficiency. The Annals of Mathematical Statistics, 22, 79-86. 1951.
19. Burnham, Kenneth P. Model selection and multimodel inference: a practical information-theoretic approach/ Kenneth P. Burnham, David R. Anderson.—2nd ed. 2002.
20. Kenneth P. Burnham, David R. Anderson. Multimodel Inference. Understanding AIC and BIC in Model Selection //SOCIOLOGICAL METHODS & RE-SEARCH, Vol. 33, No. 2, November 2004, 261-304. DOI: 10.1177/0049124104268644.

Обзор современного состояния исследований в области анализа медицинских данных болезней органов дыхания на основе методов искусственного интеллекта

Шакирзянов А. А.
Уфимский университет науки
и технологий
Институт информатики, математики
и робототехники
Уфа, Россия
e-mail: aydar102@gmail.com

Аннотация[1]

Представлен аналитический обзор современных исследований в области анализа медицинских данных с использованием методов искусственного интеллекта. Рассмотрены подходы к обработке изображений, текстовых и аудиоданных, а также данных непрерывного мониторинга состояния пациентов. Проведено сопоставление зарубежных и российских решений, выделены их сильные и слабые стороны. Установлено, что существующие системы характеризуются узкой специализацией и отсутствием межмодальной интеграции, что ограничивает их клиническую применимость. Обоснована необходимость создания новой системы поддержки принятия клинических решений (СППКР), основанной на принципах мультимодальности, пациентоориентированности и непрерывного анализа медицинских данных. Предложены направления дальнейших исследований и принципы архитектуры интегрированной интеллектуальной платформы медицинской аналитики.

Ключевые слова: система поддержки принятия клинических решений (СППКР), система поддержки принятия врачебных решений, автоматизированный мониторинг состояния пациента, непрерывный мониторинг состояния пациента, искусственный интеллект (ИИ), машинное обучение (МО), рентгенография, компьютерная томография (КТ), магнитно-резонансная томография (МРТ), аускультация легких, медицинская информатика; мультимодальные системы; автоматизированный мониторинг; анализ медицинских изображений; обработка биомедицинских данных; персонализированная медицина; цифровая диагностика; превентивная медицина.

1. Введение

Распознавание и диагностика респираторных заболеваний играют ключевую роль в их успешном лечении и улучшении прогноза. Использование информационных технологий и анализа медицинских изображений, таких как компьютерная томография, магнитно-резонансная томография и рентгенография, становится все более популярным среди специалистов в области медицины и обработки данных для повышения точности диагностики и определения оптимального лечения для каждого пациента [1]. Эти технологии могут помочь ускорить процесс диагностики, повысить точность и надежность результатов.

Проблема дифференциальной диагностики очаговых образований в легких на протяжении многих лет остается актуальной. Современный уровень диагностического оборудования таков, что детекция очаговых образований в легочной ткани не представляет трудностей, однако сохраняются сложности в интерпретации данных изменений [2]. От 20 до 40% случайно выявленных образований в легких являются злокачественными [3]. Рак легкого (РЛ) является самым распространенным из злокачественных новообразований во всех странах мира [4]. Ежегодно в мире регистрируется около миллиона новых случаев РЛ, причем большинство выявляемых случаев приходится на индустриально развитые страны (54%) [5] РЛ является основной причиной смертности от злокачественных заболеваний. В России РЛ также занимает первое место среди онкологических заболеваний, и его доля составляет 12%. В 15% всех случаев у умерших онкологических больных был выставлен диагноз РЛ. Каждая четвертая злокачественная опухоль, выявляемая у мужчин, и каждая двадцатая у женщин являются новообразованиями трахеи, бронхов, легкого [6]. Необходимость динамического

[1] Труды XI Международной научной конференции "Информационные технологии интеллектуальной поддержки принятия решений", 11-13 ноября, Уфа-Чандигарх-Фандербейлпарк, 2025

наблюдения очагов, подозрительных на злокачественные, а также развивающиеся скрининговые программы приводят к ежегодному увеличению количества рентгеновских исследований, включая компьютерную томографию грудной клетки (КТ) [7] Этот факт усугублен исследований приводит к увеличению количества обрабатываемой и хранимой информации о пациентах, что требует увеличения [9] скорости обработки со стороны врача. В связи с такой возрастающей нагрузкой особую важность приобретает применение автоматизированных и интеллектуальных программ обработки данных [10]. Разработка автоматизированных средств принятия решений в рентгенологии началась в 50-х годах прошлого века [11], и сегодня можно с уверенностью сказать, что основным элементом автоматизации в этой области становится искусственный интеллект (ИИ) и его основная часть – машинное обучение (МО). В настоящее время под ИИ принято понимать одно из направлений информационных технологий, которое связано с изучением и разработкой систем, моделирующих возможности человеческого интеллекта, таких как способность к обучению, логическому рассуждению, выводу и принятию решений [12]. Разработка интеллектуальных автоматизированных систем диагностики (ИАСД), как одного из направлений развития ИИ в медицине, является одним из таких подходов, способствующих качественной обработке данных лучевых исследований. В целом следует подчеркнуть, что актуальность работ по разработке ИАСД неоспорима как с точки зрения мировых тенденций, так и с позиции государственного вектора развития страны, так как в настоящее время ИИ является одним из показателей прорывного индустриального и информационно-технологического роста в мире [13]. В связи с этим в работе уделено внимание не только подходам разработки ИАСД, но и вопросам внедрения их в практику специализированного медицинского учреждения (СМУ). Наряду с внедрением систем ИИ в работу современного отделения лучевой диагностики (ОЛД) актуальным становится вопрос соответствия уровня технического и программного обеспечения и нормативно-законодательной базы [14]. Качественные и количественные изменения в производимом диагностическом оборудовании и потоках обследуемых пациентов значительно опережают разработки новой нормативно-правовой базы в отношении модернизации процессов лучевой диагностики. Существующая ситуация не позволяет адаптировать традиционный стиль организации ОЛД к условиям, диктуемым постоянно обновляющимся техническим и программным обеспечением. Этот факт обусловливает развитие новых подходов и диагностических алгоритмов, способствующих увеличению скорости и качества обработки данных врачами-рентгенологами. Актуальность преобразований диагностического ситуацией пандемии новой коронавирусной инфекции COVID-19, когда одним из основных диагностических критериев является выявление изменений в легких методом КТ [8]. В целом цифровизация диагностических

процесса, формирование новых алгоритмов диагностики с использованием систем ИИ, обусловлена несколькими государственными программами. Так, в соответствии с утвержденной указом № 203 от 9 мая 2017 года Президента РФ Стратегией развития информационного общества, а также утвержденной указом № 490 от 10 октября 2019 года Национальной стратегией развития искусственного интеллекта на период до 2030 года (далее Стратегия 2), развитие ИИ в целом и в медицине в частности может быть рассмотрено как мера реализации и обеспечения национальных интересов и приоритетов как политической, так и экономической направленности. Разработка плана инновационной стратегии по внедрению систем ИИ на примере ИАСД очаговых образований в легких может послужить существенным шагом к реализации основных принципов Стратегий 1 и 2. Частный пример разрабатываемой ИАСД очаговых образований в легких на уровне конкретного медицинского учреждения может рассматриваться в качестве дополнения к работе врача-диагноста, как «помощника» врача для улучшения качества оказания медицинской помощи.

В последние десять лет наблюдается прогрессирующее увеличение количества публикаций, посвященных ИИ в диагностике. Большая часть из них опубликована в технических журналах, освещает вопросы разработки ИАСД с точки зрения математического моделирования и IT-технологий [15]. В этих публикациях роль рентгенолога в разработке этих систем опускается. В изданиях медицинской направленности тематика создания ИАСД освещена в меньшем объеме. При этом описания ИАСД посвящены сравнительным характеристикам автоматизированных систем диагностики с большим или меньшим потенциалом информативности, а также сравнению ИАСД с врачами, однако не отражают этапы и алгоритмы построения ИАСД [16]. В целом анализ публикаций англоязычной базы данных PubMed и национальной библиографической базы РИНЦ (Российский индекс научного цитирования) показал, что исследования по разработке ИАСД, учитывающие «логику врача» в виде применения радиомических критериев формы, структуры очагов, отсутствуют. Представленные в литературе и в сети интернет, базы медицинских данных не являются структурированными, не отражают единой концепции их построения [17,18,19,20]. Все это объясняет необходимость и актуальность выполнения диссертационной работы по разработке системы поддержки принятия клинических решений.

2. Анализ существующих исследований

Таблица 1. – Обзор решений в области распознавания снимков функциональной диагностики легких

Сервис	Достоинства	Недостатки
Lunit Южнокорейская компания, специализируется на разработке решений искусственного интеллекта для анализа медицинских изображений, включая рентген и маммографию.	1) высокая точность (97–99%); 2) адаптация к новым заболеваниям; 3) анализ множества заболеваний.	Работает только с рентген снимками.
Aidoc Израильская компания, разрабатывающая ИИ-инструменты для анализа медицинских изображений в радиологии. Основной фокус — ускорение диагностики острых патологий (например, кровоизлияний, пневмоторекса).	Работает с рентгеном.	Распознает только пневмоторекс.
LungAI Проект с открытым исходным кодом, разработанный независимыми исследователями на основе глубоких нейросетей для анализа КТ грудной клетки. Основное направление — выявление COVID-19 и заболеваний легких.	1) работает с кт; 2) формирует отчет; 3) гарантия анонимности данных; 4) быстрое время работы; 5) интерактивный рабочий процесс.	Высокая стоимость.
Qure.ai Индийская компания, специализирующаяся на ИИ для медицинской визуализации. Разрабатывает решения для анализа рентгенов, КТ и МРТ, активно внедряет мобильные решения для телемедицины.	Наличие мобильного приложения.	Нет указания местоположения узлов. Низкая скорость по заявлению авторов.

Таким образом, можно сделать вывод о том, что в настоящее время существуют эффективные решения распознавания болезней органов дыхания методами КТ, рентген снимков, однако высокая точность распознавания обременена недостатками:

- ограниченность распознаваемых признаков;
- высокая стоимость доступа программному обеспечению;
- невысокая скорость проведения исследования, что исключает использование в режиме реального времени.

Несмотря на описанные ограничения, присущие существующим инструментам решения задачи распознавания снимков ФД, стоит отметить наличие инструментов в данной области, однако ввиду недоступности набора данных, использованных в описанных выше исследованиях и зачастую высокой стоимости использования данных решений, не представляется возможным воспроизведение эксперимента по обучению модели, а значит сложно верифицировать результаты и подтвердить их высокую точность (таблица 2).

Таблица 2. – Анализ исследований в области распознавании рака легких по снимкам ФД

Авторы	Задача	Набор данных	Архитектура	Метрики	Примечание
Daniel Korat	Предсказание вероятность рака по КТ	NLCT	3D Inception	AUC = 0.892	Требуются большие вычислительные ресурсы
Университет Айовы Стэндфордский университет	Предсказание выживаемости при раке легкого не мелкоклеточного типа	Данные предоставлены Стэнфордским институтом рака	UNet	AUC = 0.88	Данные закрытые. Эксперимент не воспроизводим
Арканзасский государственный университет Университет Мичиган	Скрининг рака легких с КТ на низкой дозе облучения	LIDC/IDRI LUNA16 Challenge Данные соревнования DSB2017 Подмножество данных NLST	Градиентный бустинг	AUC = 0.858	Ставится узкая задача, рассматривается более частный случай
Wei Zhao, Jiancheng Yang, Yingli Sun, Cheng Li, Weilan Wu, Liang Jin, Zhiming Yang, Bingbing Ni, Pan Gao, Peijun Wang, Yanqing Hua and Ming Li1	Предсказание инвазивности опухоли субцентиметровых аденокарцином легких	Данные предоставлены больницей Huadong	DenseSharp	AUC = 0.88	Данные закрытые. Эксперимент не воспроизводим

Распознавание рака легких по снимкам функциональной диагностики является в целом решенной задачей с рядом ограничений:

- закрытые исходные данные;
- относительно невысокая точность в пределах 90 процентов;
- рассмотрение частных случаев рака легких;
- высокие требования к вычислительным ресурсам.

Тем не менее, задача распознавания рака легких была решена независимо друг от друга рядом исследователей, использовавших в своих исследованиях различные наборы данных. Таким образом, можно считать задачу распознавания рака легких по снимкам ФД решенной.

Ввиду недоступности набора данных, использованных в описанных выше исследованиях и зачастую высокой стоимости использования данных решений, не представляется возможным воспроизведение эксперимента по обучению модели, а значит сложно верифицировать результаты и подтвердить их высокую точность.

Таблица 3. – Отечественные разработки в области анализа медицинских изображений

№ Иссл.	Описание решения	Достоинства	Недостатки
1	**SberMedAi.** Платформа, целью которой является объединение прорывных технологий на стыке медицины и алгоритмов машинного обучения в области диагностики заболеваний [21].	1) наличие внедренных решений в ряде поликлиник москвы; 2) большой набор данных для обучения моделей; 3) для сппкр применяется анализ всей истории болезни; 4) широкий спектр диагностируемых заболеваний.	1) принятие решения основывается на текстовых данных, изображениях или аудиофайлах обособленно друг от друга; 2) привязка к конкретному пациенту возможна только по тексту из истории болезни.
2	**Celsus.** Система, которая позволяет выявлять патологические очаги в рентгенографии и компьютерной томографии [22].	1) Цельс прошел этап проверки заявленной точности на реальных верифицированных данных; 2) первый сервис, перешедший к этапу промышленной эксплуатации по направлениям: маммография, КТ легких, флюрография, КТ головного мозга.	1) решения работают обособлено; 2) модели не взаимодействуют друг с другом ни напрямую ни опосредовано как в случае выше.
3	**Третье мнение.** Платформа включает ИИ-алгоритмы для автоматической обработки МРТ, КТ, рентген, маммографии, цифровых мазков крови и костного мозга, снимков глазного дна и челюстной системы, а также анализа видеопотока из медицинских учреждений [23].	1) реализованы модели, выделяющие семиотические признаки в лучевой диагностике: рентгенах, кт и мрт; 2) разработана система видеоаналитики действий пациента в рамках лечебного учреждения; 3) в лабораторной диагностике реализован метод для выделения и классификации типов клеток крови с	1) отсутствует привязка к конкретному пациенту; 2) нет анализа текстовых данных, аускультации.

№ Иссл.	Описание решения	Достоинства	Недостатки
		помощью глубоких сверточных нейронных сетей.	
4	**Botkin.ai**. Платформа предназначена для автоматического выявления патологических проявлений в рентгенологических исследованиях, КТ и МРТ, а также маммограмм [24].	1) интеграция платформы с медицинским учреждением; 2) автоматизированный забор данных, деперсонализация, анализ и отображение результатов исследований без дополнительной нагрузки на врачей.	Ограниченность в виде анализа исключительно изображений
5	**Care Mentor AI.** Система искусственного интеллекта для интерпретации результатов лучевых исследований (рентгенологических, КТ, МРТ и маммографии) с целью оптимизации обнаружения различных патологических состояний на ранней стадии [25].	Большая научная база выполненных исследований в области анализа медицинских изображений	Ограниченность в виде анализа исключительно изображений
6	**CHEST- IRA.** Сервис комплексного анализа компьютерной томографии органов грудной клетки с помощью технологий искусственного интеллекта [26].	1) выявляет12 патологий: легочные узлы, нарушения воздушности легочной ткани, эмфизема, плевральный выпот, аневризмы аорты, расширение легочного ствола, коронарный кальций, паракардиальный жир, лимфаденопатия, утолщения надпочечников, переломы рёбер, остеопороз; 2) для анализа КТ ОБП на предмет наличия 6 патологий: образования и стеатоз печени, утолщения надпочечников, образования почек, МКБ, аневризмы аорты, остеопороз	Точный инструмент исключительно для врача-рентгенолога
7	**RADLogics.** Основанная на ИИ программная платформа анализа медицинских изображений, существенно повышающая скорость и точность работы врачей-радиологов [27].	1) поддержка клинического решения автоматическими расчетами с привязкой к клиническим рекомендациям; 2) формирование базы данных	Отсутствие на документации и ссылок на публикации по теме представленного решения, верификация информации затруднительна..

№ Иссл.	Описание решения	Достоинства	Недостатки
8	**Pirogov.AI.** Диагностика по ото, рино и ларинго эндоскопическим фото и видеоизображениям [28].	1) работает с видео; 2) автоматически делит на изображения и анатомические структуры; 3) дает клинические рекомендации по лечению.	Отсутствие работы с текстом и аудио
9	**Sciberia.** Комплекс программных модулей для анализа медицинских изображений с помощью ИИ [29].	1) подробное описание архитектуры и протоколов; 2) каждый вид изображения обрабатывается в отдельном модуле.	Отсутствие работы с текстом и аудио

Анализ отечественных исследований и разработок в области распознавания медицинских изображений показывает, что существующие решения можно охарактеризовать как обособленные, даже конкурирующие между собой, отличающиеся отсутствием учета базы клинических рекомендаций и прочих медицинских данных о пациенте при обработке медицинских изображений. Потребность в пациентоориентированном подходе удовлетворена частично, анализ медицинских изображений решается как отдельная задача, данные, полученные путем мониторинга текущего состояния пациента анализируются и обрабатываются также обособленно, а в большинстве случаев удаленный мониторинг и вовсе не проводится либо не учитывается в представленных в Таблице 3 решениях. Таким образом, создание СППКР, включающей в себя блок анализа данных мониторинга текущего состояния пациентов является актуальной задачей на сегодняшний день (таблица 4).

Совокупный анализ отечественных решений показывает, что российский рынок медицинского ИИ развивается по направлению локальных специализированных платформ, каждая из которых решает узкий набор задач (рентген, КТ, МРТ, эндоскопия и т. п.) в рамках определённой области диагностики.

Ключевые тенденции:

1. Преобладание анализа изображений.

Почти все рассмотренные решения (Botkin.ai, Celsus, Care Mentor AI, RADLogics, CHEST-IRA) концентрируются на визуальных данных (рентген, КТ, МРТ, эндоскопия). Текстовые данные, аудиозаписи (аускультация) и показатели мониторинга физиологических параметров пациента, как правило, не интегрируются в общий контур анализа.

2. Отсутствие единой мультимодальной архитектуры.

Все существующие платформы работают обособленно — как отдельные программные модули без обмена данными и без единого хранилища, объединяющего результаты диагностики разных типов (текст, изображение, аудио, и др).

3. Недостаточная пациентоориентированность.

Анализ чаще проводится «в вакууме», без учёта клинической истории, динамики состояния пациента и сопутствующих данных. Исключение — частичная реализация этого подхода в SberMedAI, где учитывается история болезни, однако даже там отсутствует комплексный анализ мультимодальных данных.

4. Фокус на автоматизации работы врача-рентгенолога.

Многие платформы (RADLogics, CHEST-IRA, Care Mentor AI) направлены на повышение скорости и точности интерпретации изображений, на снижение нагрузки на специалистов, но не на построение комплексной системы поддержки принятия клинических решений на уровне пациента или клиники.

5. Фрагментарная интеграция с медицинскими учреждениями.

Некоторые решения (Botkin.ai, SberMedAI) уже внедрены в клиническую практику, но остаются замкнутыми системами, работающими в рамках конкретных региональных инфраструктур или отдельных клиник, без национальной интеграции.

6. Отсутствие открытых данных и публикаций.

Для ряда решений (например, RADLogics, Sciberia) отсутствует достаточная открытая научная верификация, что затрудняет оценку достоверности

заявленных результатов и воспроизводимость исследований.

Современные отечественные ИИ-системы медицинской диагностики находятся на стадии функциональной зрелости отдельных компонентов, но не системы в целом. Следующим шагом в развитии должно стать создание интегрированных СППКР нового поколения, объединяющих:

1) анализ изображений, текстовых и звуковых данных;

2) данные непрерывного мониторинга состояния пациентов;

3) единое клиническое хранилище данных, связанное с ЭМК;

4) блоки интерпретации с опорой на клинические рекомендации.

Такие системы смогут не только повышать точность диагностики, но и обеспечивать персонализированное сопровождение пациента на всех этапах лечения — от первичного осмотра до динамического наблюдения.

Таблица 4. – Отечественные разработки в области цифровой диагностики

№ Иссл.	Описание решения	Достоинства	Недостатки
1	**CVL CVisionRad.** Облачный сервис для анализа радиологических исследований [30].	Предоставляет инструменты анализа КТ органов грудной клетки и брюшной полости, РГ исследований, позволяя оперативнее выявлять признаки различных патологий.	Отсутствие работы с текстом и аудио
2	**Анализ флюорограм.** Сервис умеет анализировать цифровые флюорографические снимки и выявлять в них патологические очаги [31].	Позволяет выявлять признаки и каталогизировать патологии органов грудной клетки	Работает только с флюорографическими снимками
3	**OneCell**. Телемедицинский комплекс с ИИ для патологоанатомических лабораторий, позволяет ускорить процесс диагностики онкозаболеваний, сохраняя высокий уровень достоверности результатов за счет собственного современного оборудования [32].	Комплексная оценка по всем имеющимся данным конкретного пациента	СППКР осуществляется не ИИ, а сообществом специалистов, т.е. людей
4	**iVenus.AI.** Позволяет провести предварительную диагностику заболеваний вен нижних конечностей на основе анализа фотографических изображений с целью определить наличие и класс заболевания [33].	Получение общей информации о проявлениях варикозного расширения вен	Обработка исключительно специфических изображений

№ Иссл.	Описание решения	Достоинства	Недостатки
5	**Celly.AI**. Технологическое решение на основе смартфона, которое подключается к окуляру микроскопа и помогает анализировать мазки крови [34].	1) автоматизированное выполнение ряда специфических тестов в гематологии; 2) исследование малярии; 3) пап тест; 4) анализ мазков в ветеринарии с помощью смартфона и микроскопа.	Обработка исключительно специфических изображений
6	**PathVision.ai**. Поддержка принятия решений во время проведения патоморфологической диагностики на основе ИИ [35].	1) наличие на сайте подробной документации о работе в платформе; 2) помощь ассистента при патоморфологической диагностике.	Обработка исключительно специфических изображений
7	**Retina.Ai**. Цифровая облачная офтальмологическая платформа [36].	1) функционирующее облачное решение по анализу изображений глазного дна; 2) выявление признаков заболевания и определение его стадии	1) платформа не прошла клинические испытания; 2) обработка исключительно специфических изображений
8	**Coronarography.Ai.** Платформа для определения патологии коронарных артерий с технологиями искусственного интеллекта [37].	Работа с ЭКГ и текстовыми данными	Отсутствие на документации и ссылок на публикации по теме представленного решения, верификация информации затруднительна.
9	**Polyptron.** Система на основе ИИ для анализа видеоданных эндоскопии и поиска подозрений на колоректальный рак [38].	1) анализ видео в режиме реального времени; 2) низкие требования к вычислительным мощностям; 3) обработка на периферийном устройстве без отправки на сервер	1) поиск ограниченного числа первичных признаков рака; 2) отсутствие анализа иных данных
10	**Dentomo.** Анализ стоматологических изображений [39].	Для улучшения результатов обработки изображений система обучается не только на результатах обработки больших количеств данных, но и благодаря переобучению модели: с помощью исправлений и примечаний, введённых стоматологом	Обработка исключительно специфических изображений

№ Иссл.	Описание решения	Достоинства	Недостатки
11	**BrainPhone.** Диагностика и мониторинг болезни Паркинсона по голосу на основе ИИ [40].	Использование семантического, фонетический и других типов анализа речи	1) обработка исключительно аудио; 2) отсутствие документации и ссылок на публикации по теме представленного решения, верификация информации затруднительна
12	**Oxygen Technologies.** ИИ-технология анализа рентгенограмм органов грудной клетки, позволяющая автоматически распознавать патологии [41].	1) ручная разметка снимков несколькими врачами при формировании набора данных; 2) высокая точность (выше 92%)	Отсутствие работы с текстом и аудио
13	**MultiVox**. PACS/RIS-система с применением ИИ для анализа изображений [42].	1) объединяет в единую систему регистрирующие изображения аппараты (рентген, ангиограф, узи, маммограф, кт, мрт, пэт, эндоскопы, электронные микроскопы и т.д.); 2) наличие оборудования для документирования результатов обследований; 3) арм врачей, предназначенные для работы с изображениями и доступа к диагностической радиологической информации.	Отсутствие работы с текстом и аудио
14	**Ocuscreen.** Распознавание глазных патологий с помощью алгоритмов машинного обучения [43].	1) возможность выделения областей для анализа; 2) генерация pdf отчета с предполагаемым диагнозом; 3) способность определять пыль и «битые пиксели» на оптике камер и исключать их из диагноза; 4) фильтрация непригодных изображений для анализа	Обработка исключительно специфических изображений, прошедших фильтрацию.
15	**Ai Radiology CXR.** Сервис автоматического распознавания патологий легких с помощью ИИ [44].	1) наличие в свободном доступе подробной документации о работе по; 2) доступ к исходным наборам данных.	Работа исключительно по рентгеновским изображениям в прямой проекции
16	**Digital Pathology.** Программное	1) синтез гистологического атласа и микроскопа,	отсутствие документации и

№ Иссл.	Описание решения	Достоинства	Недостатки
	обеспечение для морфологической диагностики [45].	позволяющий дифференцировать диагноз на основании объективных морфологических и клинических особенностей; 2) пользователь заполняет анкету с выбором параметров, число вероятных нозологических единиц сужается, что упрощает диагностическую работу патологоанатома в конкретном диагностическом случае.	ссылок на публикации по теме представленного решения, верификация информации затруднительна
17	**EyeMove.** Программа по выявлению признаков болезни Паркинсона по движению глаз [46].	Работа приложения с видео в реальном времени	1) недоступность для пользователей android и пк; 2) отсутствие документации и ссылок на публикации по теме представленного решения, верификация информации затруднительна
18	**NTechMed CT Brain**. ПО для диагностики пациентов с острым нарушением мозгового кровообращения (ОНМК) [47].	Высокое быстродействие: обработка нативного КТ-исследования головы (650 срезов) менее 4 минут, возможность проведения до 200 исследований параллельно	Обработка исключительно специфических изображений.

Последующий анализ существующих разработок подтверждает гипотезу ограниченности и узкой практической применимости существующих решений ввиду отсутствия согласованности в работе представленных моделей с моделями анализа текстовых медицинских данных и медицинских аудиофайлов, таких как данные спирографии и аускультации пациента, что является важным аспектом в поддержке принятия клинического решения. Кроме того, изучение результатов исследований в смежных областях анализа медицинских данных также демонстрирует неудовлетворенность потребности в индивидуализированном подходе в лечении и диагностики заболеваний различного характера. По-прежнему остается актуальной задача организации автоматизированного мониторинга пациентов, находящихся на учете в профильной организации, а также задача учета взаимного влияния заболеваний различного типа, протекающих у отдельно взятого пациента. Однако для успешного решения столь амбициозной и обширной задачи, следует разработать такую СППКР, которая была бы способна учесть все медицинские факторы, влияющие на конкретного пациента в рамках хотя бы одного заболевания. Функциональность и доступность такой системы должна быть проверена на практике в действующих медицинских учреждениях, а ее полезность и удобство должны быть подтверждены не только высокими метриками качества, но и положительной обратной связью как от врачей, так и их пациентов.

Проведённый анализ демонстрирует, что развитие систем искусственного интеллекта для медицинской диагностики в России и мире продолжает двигаться по пути сегментации и узкой специализации. Большинство решений ориентировано на определённый тип исследования или орган, что

обеспечивает высокую точность внутри ограниченной области применения, но не позволяет выстраивать сквозные диагностические маршруты.

Основные выявленные тренды и закономерности:

1. Узконаправленная функциональность.

Более 80 % представленных систем (например, PathVision.ai, Retina.AI, Dentomo, iVenus.AI, Ocuscreen) специализируются на анализе изображений конкретных органов — глазного дна, вен, ротовой полости, кожи и т. д. Это повышает качество диагностики в пределах конкретной задачи, но делает систему непригодной для интеграции в единое клиническое пространство.

2. Рост мобильных и периферийных решений.

Появление таких инструментов, как Celly.AI, iVenus.AI, Polyptron, свидетельствует о тенденции к децентрализации диагностических технологий. Диагностика всё чаще проводится не в крупных центрах, а «на местах» — с помощью смартфонов, периферийных устройств и облачных сервисов. Это особенно актуально для телемедицины и сельских регионов.

3. Развитие крайне ограниченной мультимодальности.

Исключения, такие как Coronarography.AI, демонстрируют попытки объединить визуальные и текстовые данные (например, ЭКГ + описания), однако подобные примеры единичны. Полноценная мультимодальная интеграция остаётся нерешённой задачей.

4. Отсутствие открытой научной верификации.

Для значительной части решений (BrainPhone, EyeMove, Digital Pathology, Coronarography.AI) характерно отсутствие публичной документации и рецензируемых исследований, что препятствует репликации результатов и внедрению в клиническую практику.

5. Акцент на вспомогательную роль ИИ.

Многие сервисы (OneCell, PathVision.ai, MultiVox, Ai Radiology CXR) позиционируются как помощники врача, а не автономные диагностические системы. Это подчёркивает переход к модели человек-в-контуре — врач остаётся главным субъектом принятия решений, а ИИ выполняет функцию анализа и фильтрации данных.

6. Недостаточная интеграция с клиническими рекомендациями и ЭМК

Ни одна из рассмотренных платформ не демонстрирует полноценной связи с базой клинических рекомендаций, стандартами лечения или электронными медицинскими картами. Отсутствует контекстная персонализация решений под конкретного пациента.

Совокупный анализ отечественных разработок в области цифровой диагностики подтверждает наличие серьёзного разрыва между точечными ИИ-модулями и полноценными системами поддержки принятия клинических решений (СППКР). Современные сервисы эффективно решают частные задачи, но не обеспечивают когнитивной непрерывности данных между различными этапами диагностики и лечения.

Дальнейшее развитие медицинского ИИ должно быть направлено на:

1) - создание единой мультимодальной архитектуры, объединяющей изображения, текстовые и аудиоданные, а также показатели непрерывного автоматизированного мониторинга состояния пациента;

2) - внедрение механизма обратной связи между врачом и системой, обеспечивающего обучение моделей ИИ на реальных клинических кейсах;

3) - интеграцию с национальными регистрами и ЭМК для учёта динамики заболеваний;

4) - разработку открытых протоколов верификации и сертификации решений на основе ИИ;

5) - ориентацию на пациентоцентричный подход, где анализ не изолирован, а проводится в контексте полного медицинского профиля пациента.

Внедрение подобного подхода позволит перейти от набора разрозненных диагностических сервисов к единой интеллектуальной экосистеме медицинской аналитики, способной формировать персонализированные клинические рекомендации и повышать эффективность здравоохранения в целом

.

Таблица 5. – Отечественные разработки в области профилактики и лечения состояний, заболеваний и осложнений.

№ Иссл.	Описание решения	Достоинства	Недостатки
1	**Webiomed.** Платформа прогнозной аналитики и управление рисками в здравоохранении на основе машинного обучения. Первая в России система искусственного интеллекта, зарегистрированная Росздравнадзором как программное медицинское изделие [48].	Вся информация из ЭМК автоматически обезличивается и анализируется, врачу не нужно ничего вносить вручную.	Работает только с текстовыми данными с использованием NLP технологии.
2	**MeDiCase.** Система доврачебной диагностики острых и хронических заболеваний с применением методов ИИ, помогающая проводить первичное обследование пациента, принятие решений о необходимости его очного обследования, вызова скорой помощи, мониторинга течения хронических болезней [49].	Удобная онлайн система опросов в виде проверки симптомов	Анализ только текста в виде ответов на ограниченный список вопросов
3	**Lexema-Medicine.** Специализированная СППВР для назначения персонализированной терапии с использованием алгоритмов искусственного интеллекта [50].	СППВР является частью большей медицинской информационной системы	Анализируются только текстовые и табличные данные
4	**MDinc.** Онлайн сервис для восстановления и развития мозга на основе ИИ [51].	Удобство использования онлайн-опросников в виде тестов	1) ии используется только для анализа текста; 2) выполняет роль сопоставления с научной базой, т.е. лишь автоматизирует процесс проверки тестов, а не является полноценной СППКР
5	**iCognito**. Приложения для самостоятельного прохождения диагностики ментального здоровья и обучения навыкам срочной психологической самопомощи [52].	Удобные для пользователей алгоритмы сбора и анализа данных для диагностики состояния психического здоровья и построения предиктивных моделей в целях профилактики психических расстройств	ИИ используется только для анализа текста, выполняет роль сопоставления с научной базой, т.е. лишь автоматизирует процесс проверки тестов, а не является полноценной СППКР

№ Иссл.	Описание решения	Достоинства	Недостатки
6	**Onqueta.** Сервис для выявления людей с высоким онкологическим риском для последующего генетического тестирования [53].	Доступный опросник, реализованный в виде чат-бота Telegram	ИИ используется только для анализа текста, выполняет роль сопоставления с научной базой, т.е. лишь автоматизирует процесс проверки тестов, а не является полноценной СППКР
7	**Novel.** Набор решений для генетиков [54].	Удобный инструмент для работы со специфическим и для генетиков форматами данных	Работает только со специфическими данными
8	**Stop-Hemangioma.** Система дистанционного наблюдения за детьми с инфантильными (младенческими) гемангиомами [55].	Автоматическая оценка размера пятна и яркости	1) модель работает только с изображениями и не является СППКР; 2) изображения по-прежнему вручную рассматривает врач.
9	**Muscles.Ai.** Программно-аппаратный комплекс, повышающий безопасность фитнес-тренировок и помогающий сформировать паттерн физиологических движений человека [56].	1) точность распознавания движений 96,2%; 2) латентность распознавания 8 мс	Необходимость использования костюма – портупеи с последующим анализом электрических сигналов в виде текстовых данных.

Анализ смежных областей в распознавании медицинских данных также показывает обособленность применяемых моделей при обработке данных. Не учитываются системно важные факторы из других источников, в том числе разнородные неструктурированные медицинские данные пациента. Следует отметить высокую заявленную эффективность представленных данных, однако в большинстве случаев в погоне за коммерческой привлекательностью и уникальностью представленного решения разработчики моделей не представляют данные для воспроизведения эксперимента для подтверждения заявленных метрик качества, что оставляет за пользователем данных систем только надежду на добросовестность авторов представленных решений и алгоритмов. В таблице 6 представлено множество инструментов для реализации частично автоматизированного мониторинга текущего состояния пациентов, однако ни одно из представленных в таблице 6 решений не является СППКР, хотя некоторые авторы позиционируют и представляют результаты своих работ именно как СППКР. Тем не менее, реализованные инструменты организации автоматизированного мониторинга состояния пациентов являются обязательным и необходимым шагом на пути создания полноценной целостной СППКР, в которую блок анализа результатов мониторинга состояния пациентов должен быть включен наряду с блоками распознавания иных разнородных, в том числе неструкутирированных, медицинских данных.

Рассмотренные решения демонстрируют важный сдвиг в развитии отечественных систем медицинского ИИ — переход от анализа

медицинских изображений к работе с текстовыми и структурированными данными из электронных медицинских карт (ЭМК), анкет и опросников. Вместе с тем, данный переход выявил новые ограничения, связанные с низкой степенью интеграции между различными типами данных и отсутствием комплексного подхода к пациенту.

Основные тренды и особенности развития:

1. Переход от диагностических к прогностическим и профилактическим задачам

Такие системы, как Webiomed, MeDiCase, Onqueta и iCognito, фокусируются не только на выявлении заболеваний, но и на прогнозировании рисков и раннем предупреждении патологий. Это отражает общемировую тенденцию перехода к превентивной медицине, где ИИ используется для оценки вероятности развития заболеваний до их клинического проявления.

2. Акцент на текстовом анализе и опросных данных.

Большинство решений (Webiomed, MeDiCase, Lexema-Medicine, MDinc, iCognito) работают исключительно с текстом и табличными данными, получаемыми из анкет, ЭМК и опросов. Хотя такие подходы упрощают сбор информации, они лишены возможности учитывать физиологические параметры, аудиозаписи или визуальные данные пациента, что ограничивает точность прогнозов.

3. Рост количества самообучающихся и пользовательских решений

Приложения MDinc, iCognito, Onqueta и Muscles.AI отражают стремление к вовлечению пользователя в процесс самодиагностики и профилактики. Этот тренд указывает на персонализацию и демократизацию медицинских технологий, где пациент становится активным участником мониторинга своего состояния.

4. Ориентация на коммерческую применимость и потребительский рынок

Многие решения, особенно в области психического здоровья и фитнеса (iCognito, MDinc, Muscles.AI), развиваются как B2C-продукты, что способствует популяризации технологий, но одновременно снижает уровень научной прозрачности. Отсутствие открытых данных и клинических верификаций делает невозможной независимую оценку заявленных метрик.

5. Отсутствие интеграции и единых стандартов

Все решения действуют в рамках собственных экосистем и не взаимодействуют с другими модулями или медицинскими базами. Даже зарегистрированные системы, такие как Webiomed, остаются ограниченными по охвату данных — анализ осуществляется лишь по текстовой информации, без учёта мультимодальных факторов состояния пациента.

6. Появление комплексных, но фрагментарных решений

Некоторые проекты (Lexema-Medicine, Novel, Muscles.AI) показывают попытки объединить элементы диагностики, терапии и мониторинга, однако по сути остаются модульными компонентами, не обладающими свойствами полноценных СППКР.

Отечественная медицинская ИИ-индустрия постепенно переходит от анализа изображений к контекстному и поведенческому анализу — текстов, анкет, генетических данных и физиологических сигналов. Однако отсутствует единая архитектура, способная объединить все эти данные в рамках одного пациента и формировать обоснованные клинические решения.

Следующим этапом развития должно стать создание интегрированных систем поддержки профилактики и персонализированного лечения, в которых ИИ будет:

1) синтезировать данные из ЭМК, анкет, физиологических сенсоров и медицинских изображений;

2) формировать прогнозы с учётом динамики состояния и генетических предрасположенностей;

3) адаптировать рекомендации под конкретного пациента и его контекст;

4) обеспечивать прозрачность и клиническую воспроизводимость результатов.

Таким образом, современные разработки можно рассматривать как этап перехода от узкоспециализированных модулей к будущим гибридным системам, которые объединят диагностическую, прогностическую и профилактическую аналитику в рамках единой интеллектуальной платформы здравоохранения.

3. Выводы

На рынке поддержки принятия клинических решений с использованием технологий ИИ в диагностике и мониторинге бронхолегочных заболеваний не обнаружено решений, позволяющих осуществить комплексный пациентоориентированный подход. Ряд решений, таких как Цельс, Третье мнение, Botkin AI, Pirogov AI отличаются высокой надежностью и эффективностью в распознавании специальных медицинских изображений по ограниченному числу признаков. Однако моделей, способных функционировать с учетом наличия текстовых и/или аудио медицинских данных о пациентах, либо работающих в связке с такими моделями, не обнаружено. Исключением является технология, предложенная и частично реализованная SBER

MED AI, в которой учитываются текстовые данные в виде жалоб пациента, однако модели распознавания медицинских изображений работают, как и в других случаях, обособленно, что вынуждает лечащего врача учитывать информацию, получаемую от разных моделей обособленно и самостоятельно проводить верификацию данных по принадлежности к конкретному пациенту, что в условиях высокой загруженности медицинского персонала маловероятно позволит назвать предложенное решение практически широко применимым. В связи проведенным анализом, необходима разработка такой СППКР, которая бы не отвлекала медицинский персонал от выполнения обязанностей, а напротив, упрощала трудовую деятельность лечащего врача, отличалась комплексным пациентоориентированным подходом в поддержке принятия клинических решений, учитывала результаты автоматизированного и неавтоматизированного мониторинга текущего состояния пациента в виде обработки результатов осмотра пациента и прочих биомедицинских данных, а также концептуально интегрировала модели распознавания разнородных медицинских данных, в том числе нечетко определенных и слабо структурированных.

Обобщая результаты проведённого анализа отечественных и зарубежных решений, можно заключить, что современное состояние исследований и разработок в области анализа медицинских данных с применением искусственного интеллекта характеризуется высокой степенью технологической зрелости частных направлений (распознавание изображений, анализ ЭМК, обработка аудио и генетических данных), но недостаточной степенью интеграции между ними.

Основными выявленными проблемами являются:

- фрагментарность существующих решений и отсутствие единой мультимодальной платформы, объединяющей различные типы медицинских данных;
- низкая воспроизводимость и недостаточная открытость данных, что затрудняет верификацию результатов исследований;
- ориентация на коммерческую эффективность, а не на клиническую воспроизводимость и доказательную базу;
- отсутствие механизмов обмена данными между ИИ-системами и электронными медицинскими картами, что препятствует построению когнитивной цепочки «диагностика — мониторинг — лечение».

Тем не менее, наблюдается устойчивая тенденция к пациентоцентричной трансформации медицинских ИИ-систем. Это выражается в развитии:

- интеллектуальных сервисов раннего выявления и профилактики заболеваний (Webiomed, MeDiCase, Onqueta);
- мобильных и телемедицинских платформ, ориентированных на самообслуживание и регулярный мониторинг здоровья пациента;
- инициатив по созданию национальных медицинских дата-платформ, способных аккумулировать и стандартизировать данные из разных источников.

Таким образом, можно заключить, что следующим стратегическим этапом развития отечественных систем поддержки принятия клинических решений (СППКР) должны стать:

1. Разработка интегрированной мультимодальной архитектуры, объединяющей анализ изображений, аудиосигналов, текстовых и сенсорных данных в единую когнитивную модель пациента.
2. Формирование единого клинико-диагностического контура, обеспечивающего сквозную обработку информации от регистрации данных до выработки рекомендаций.
3. Создание национальной платформы верификации и сертификации ИИ-решений, обеспечивающей воспроизводимость и достоверность заявленных метрик.
4. Проведение пилотных внедрений СППКР нового поколения в учреждениях республиканского и федерального уровней для оценки их клинической, экономической и организационной эффективности.
5. Разработка такой системы станет основой перехода от локальных диагностических модулей к интеллектуальной экосистеме медицинской аналитики, обеспечивающей персонализированное лечение и непрерывный мониторинг состояния пациентов с учётом их уникальных физиологических и социальных параметров.

Первым шагом в разработке такой СППКР является организация в действующем лечебном заведении системы автоматизированного мониторинга текущего состояния пациентов, находящихся на учете в профильном медицинском учреждении. Такая работа должна вестись под шефством практикующий врачей, а для оценки эффективности функционирования системы автоматизированного мониторинга должны быть приняты во внимание отзывы как врачей, так и пациентов, пользующихся системой мониторинга на регулярной основе.

СППКР должна включать в себя блок анализа результатов проводимого мониторинга наряду с

блоками анализа иной биомедицинской информации.

Идеальная СППКР должна отличаться не только пациентоориентированным подходом, но и обязана учитывать взаимное влияние различных одновременно протекающих заболеваний отдельного пациента, однако для решения такой амбициозной и широкой задачи следует сначала организовать СППКР, функционирующей в рамках отдельно взятой предметной области, но отличающейся индивидуализированным подходом. Примером такой системы может быть СППКР, специализирующаяся на бронхолегочных заболеваниях с учетом симптоматики пациента, проанализированной в блоке анализа результатов непрерывного мониторинга.

Список используемых источников

1. Haug C. J., Drazen J. M. Artificial intelligence and machine learning in clinical medicine, 2023 //New England Journal of Medicine. – 2023. – Т. 388. – №. 13. – С. 1201-1208.
2. Kufel J. et al. What is machine learning, artificial neural networks and deep learning? — Examples of practical applications in medicine //Diagnostics. – 2023. – Т. 13. – №. 15. – С. 2582.
3. Daidone M., Ferrantelli S., Tuttolomondo A. Machine learning applications in stroke medicine: Advancements, challenges, and future prospectives //Neural Regeneration Research. – 2024. – Т. 19. – №. 4. – С. 769-773.
4. Woodman R. J., Mangoni A. A. A comprehensive review of machine learning algorithms and their application in geriatric medicine: present and future //Aging Clinical and Experimental Research. – 2023. – Т. 35. – №. 11. – С. 2363-2397.
5. Chakraborty C. et al. From machine learning to deep learning: An advances of the recent data-driven paradigm shift in medicine and healthcare //Current Research in Biotechnology. – 2023. – С. 100164.
6. Albahra S. et al. Artificial intelligence and machine learning overview in pathology & laboratory medicine: A general review of data preprocessing and basic supervised concepts //Seminars in Diagnostic Pathology. – WB Saunders, 2023. – Т. 40. – №. 2. – С. 71-87.
7. Nilius H. et al. Machine learning applications in precision medicine: overcoming challenges and unlocking potential //TrAC Trends in Analytical Chemistry. – 2024. – С. 117872.
8. Sánchez Fernández I., Peters J. M. Machine learning and deep learning in medicine and neuroimaging //Annals of the Child Neurology Society. – 2023. – Т. 1. – №. 2. – С. 102-122.
9. Qureshi H. et al. Machine learning investigation of tuberculosis with medicine immunity impact //Diagnostic Microbiology and Infectious Disease. – 2024. – Т. 110. – №. 3. – С. 116472.
10. Araújo A. L. D. et al. Machine learning concepts applied to oral pathology and oral medicine: a convolutional neural networks' approach //Journal of Oral Pathology & Medicine. – 2023. – Т. 52. – №. 2. – С. 109-118.
11. Swarnakar R., Yadav S. L. Artificial intelligence and machine learning in motor recovery: A rehabilitation medicine perspective //World Journal of Clinical Cases. – 2023. – Т. 11. – №. 29. – С. 7258.
12. Mirijello A. et al. Identifying predictors associated with risk of death or admission to intensive care unit in internal medicine patients with sepsis: A comparison of statistical models and machine learning algorithms //Antibiotics. – 2023. – Т. 12. – №. 5. – С. 925.
13. Sim J. Z. T. et al. Machine learning in medicine: what clinicians should know //Singapore medical journal. – 2023. – Т. 64. – №. 2. – С. 91-97.
14. Бритвина П. В. Машинное обучение в медицине: революция диагностики, лечения и персонализации //Вестник науки. – 2024. – Т. 4. – №. 1 (70). – С. 442-444.
15. Кузнецов А. Д. и др. Искусственный интеллект и машинное обучение в медицине: применение для диагностики и лечения заболеваний //фундаментальная и прикладная наука: состояние и тенденции развития. – 2023. – С. 93-97.
16. Хрипунов Д. Д. Методы машинного обучения в анализе медицинских изображений. История и обзор актуальных проблем //Научные исследования 2023: актуальные теории и концепции: сборник материалов XXIX-ой международной очно-заочной научно-практической конференции, в 4 т., Том 3, 24 мая, 2023–Москва: Издательство НИЦ «Империя», 2023.–148с. – 2023. – С. 141.
17. Нурмаммедова О., Батырова О., Айдогдыев М. Усовершенствование методов машинного обучения для анализа клинической информации //Вестник науки. – 2024. – Т. 2. – №. 10 (79). – С. 520-523.
18. Бухвалов П. С., Воронов А. В. Исследование возможности использования методов машинного обучения при оценке корректности медицинских дисторсиограмм //завалишинские чтения 23. – 2023. – С. 18-22.
19. Разумников С. В., Мишин И. В. Искусственный интеллект и машинное обучение: тенденции развития и применение в различных областях

//Инновационные технологии в машиностроении: сборник трудов XV Международной научно-практической конференции, г. Юрга, 23-25 мая 2024 г. – Томский политехнический университет, 2024. – С. 74-76.

20. Лоскин В. А. Технологии искусственного интеллекта в современной медицине //Цифровые, компьютерные и информационные технологии в науке и. – 2023. – С. 298.
21. Искусственный интеллект для здравоохранения // SBER MED AI: сайт. 2025. URL: https://sbermed.ai// (дата обращения: 30.05.2025).
22. AI-платформа анализа медицинских изображений //ЦЕЛЬС: сайт. 2025. URL: https://celsus.ai/// (дата обращения: 30.05.2025).
23. Инновации для здравоохранения //Третье мнение: сайт. 2025. URL: https://thirdopinion.ai// (дата обращения: 30.05.2025).
24. Искусственный интеллект для медицинских изображений //Botkin AI: сайт. 2025. URL: http://botkin.ai// (дата обращения: 30.05.2025).
25. Система искусственного интеллекта для интерпретации результатов лучевых исследований //Care Mentor AI: сайт. 2025. URL: http://carementor.ru/// (дата обращения: 30.05.2025).
26. Искусственный интеллект в помощь рентгенологу //Ira Labs: сайт. 2025. URL: https://ira-labs.ru// (дата обращения: 30.05.2025).
27. Цифровые инструменты динамического наблюдения онкопациентов //RadLogics: сайт. 2025. URL: https://radlogics.ru/// (дата обращения: 30.05.2025).
28. Повысьте точность диагностики заболеваний уха, горла, носа //Pirogov.AI: сайт. 2025. URL: https://pirogov.ai// (дата обращения: 30.05.2025).
29. Комплекс программ //Sciberia: сайт. 2025. URL: https://sciberia.io/ru// (дата обращения: 30.05.2025).
30. Облачный сервис на основе искусственного интеллекта//CvisionRad: сайт. 2025. URL: http://cvl-ct.cvisionlab.com/// (дата обращения: 30.05.2025).
31. Анализ флюрограмм //Phthisisbiomed: сайт. 2025. URL: http://www.ftizisbiomed.ru/// (дата обращения: 30.05.2025).
32. Цифровая патология с ИИ //Onecell: сайт. 2025. URL: https://www.onecell.ai// (дата обращения: 30.05.2025).
33. Сервис для флебологов //Ivenus.ai: сайт. 2025. URL: https://ivenus.ai// (дата обращения: 30.05.2025).
34. Набор автоматизации микроскопа //Celly: сайт. 2025. URL: https://celly.ai// (дата обращения: 30.05.2025).
35. Платформа цифровой патоморфологии с искусственным интеллектом//PathVision.ai: сайт. 2025. URL: https://pathvision.ai/ru// (дата обращения: 30.05.2025).
36. Программное обеспечение для работы с цифровыми медицинскими изображениями при диагностике офтальмологических патологий путем анализа фотографий глазного дна и сканов структурной ОКТ//Retina.ai: сайт. 2025. URL: https://www.screenretina.ru// (дата обращения: 30.05.2025).
37. Платформа для определения патологии коронарных артерий с технологиями искусственного интеллекта // Coronarography.ai: сайт. 2025. URL: https://coronarography.ai// (дата обращения: 30.05.2025).
38. Система на основе ИИ для анализа видеоданных эндоскопии и поиска подозрений на колоректальный рак //Polyptron: сайт. 2025. URL: https://evalab.ai/ru// (дата обращения: 30.05.2025).
39. Анализ стоматологических изображений //Dentomo: сайт. 2025. URL: https://www.dentomo.ru/// (дата обращения: 30.05.2025).
40. Диагностика и мониторинг болезни Паркинсона по голосу на основе ИИ// Brainphone.ru: сайт. 2025. URL: https://brainphone.ru// (дата обращения: 30.05.2025).
41. ИИ-технология анализа рентгенограмм органов грудной клетки // Oxygen Technologies: сайт. 2025. URL: https://oxytech.io/ru// (дата обращения: 30.05.2025).
42. Масштабируемые PACS/РИС системы //MultiVox: сайт. 2025. URL: https://multivox.ru// (дата обращения: 30.05.2025).
43. Постановка диагноза диабетической ретинопатии //Ocuscreen: сайт. 2025. URL: https://ocuscreen.ru// (дата обращения: 30.05.2025).
44. Сервис автоматического распознавания патологий легких с помощью ИИ// Ai Radiology CXR: сайт. 2025. URL: https://mosmed.ai/service_catalog/ai-radiology-cxr// (дата обращения: 30.05.2025).
45. программное обеспечение для морфологической диагностики//Digital Pathology: сайт. 2025. URL: https://dpathology.ru// (дата обращения: 30.05.2025).

46. Мониторинг болезни Паркинсона по движению глаз//Eyemove: сайт. 2025. URL: https://eyemove.ru// (дата обращения: 30.05.2025).
47. ПО для диагностики пациентов с острым нарушением мозгового кровообращения // NTechMed CT Brain: сайт. 2025. URL: https://ntechlab.ru // (дата обращения: 30.05.2025).
48. Платформа прогнозной аналитики//Webiomed: сайт. 2025. URL: https://webiomed.ru// (дата обращения: 30.05.2025).
49. Оценка вероятностей заболеваний по симптомам//Medicase: сайт. 2025. URL: http://medicase.newdiamed.ru// (дата обращения: 30.05.2025).
50. Специализированная система управления ресурсами для медицинских холдингов //Lexema-medicine: сайт. 2025. URL: https://lexema.ru/medicine// (дата обращения: 30.05.2025).
51. ИИ сервис повышения стрессоустойчивости //via.md: сайт. 2025. URL: https://mdinc.ru// (дата обращения: 30.05.2025).
52. Мобильные приложения психологической самопомощи //iCognito: сайт. 2025. URL: https://icognito.app// (дата обращения: 30.05.2025).
53. Цифровые технологии для борьбы с наследственным раком // Onqueta: сайт. 2025. URL: https://onqueta.com// (дата обращения: 30.05.2025).
54. Комплекс программ /Novel: сайт. 2025. URL: https://novel-soft.ru/products// (дата обращения: 30.05.2025).
55. Система дистанционного наблюдения за детьми с инфантильными гемангиомами//Stop-Hemangioma: сайт. 2025. URL: https://dzen.ru/a/ZHDB3iIC6wQarwm1// (дата обращения: 30.05.2025).
56. Программно-аппаратные системы измерения и анализа движений тела человека// Musclesai: сайт. 2025. URL: www.musclesai.ru// (дата обращения: 30.05.2025).

Автоматизация первичного интервью с помощью ИИ-ассистента

Хисматов И. Э.
Уфимский университет науки и технологий
Институт информатики, математики и робототехники
Уфа, Россия
e-mail: khismatov.ildar@mail.ru

Аннотация[1]

В статье рассматривается проблема субъективности и низкой эффективности процессов первичного скрининга кандидатов в IT-рекрутинге. Предлагается решение на основе голосового интеллектуального ассистента, использующего модель GPT-4 для проведения адаптивных интервью, мультимодального анализа ответов (текст, голос) и формирования объективной оценки. Система сокращает ручные процессы HR менеджеров. Центральным элементом работы является разработка математической модели оценки эффективности системы и детальное описание ее архитектуры.

Ключевым элементом исследования является разработанная архитектура голосового интеллектуального рекрутера на базе GPT-4, включающая модуль речевого взаимодействия, анализатор смыслового содержания ответов, блок оценки эмоционального окраса речи и систему генерации инсайтов.

В статье предложена методика интеграции системы в процессы IT-рекрутинга с описанием этапов адаптивного интервьюирования и формирования итоговой оценки кандидатов. Особое внимание уделено решению проблем обработки мультимодальных данных и обеспечению корректности работы голосового интерфейса. Проведенное экспериментальное исследование демонстрирует повышение объективности первичного отбора и сокращение операционных затрат при сохранении качества найма. Результаты работы представляют практический интерес для HR-специалистов и разработчиков интеллектуальных систем поддержки принятия кадровых решений

Ключевые слова: голосовой ассистент, искусственный интеллект, рекрутинг, принятие решения

Введение

В сфере информационных технологий наблюдается острая борьба за привлечение талантливых кадров. Существующая практика отбора претендентов, осуществляемая рекрутерами, часто отличается трудоемкостью и склонна к субъективным оценкам, таким как эффект первого впечатления или предпочтение «своих». В качестве решения предлагается внедрение голосового ИИ-ассистента, который обеспечит беспристрастный анализ как профессиональных знаний, так и личностных характеристик соискателей. Основная задача данного исследования — создание формализованной модели подобной системы и детальное описание архитектуры ее основных компонентов.

1. Роль нейросетей в современной практике управления персоналом

Цифровая перестройка процессов вынуждает сферу персонала активно внедрять инструменты на основе искусственного интеллекта и нейронных сетей. В результате современный эйчар превратился в мультифункционального эксперта, которому необходимы навыки маркетолога, психолога, автора контента и стратега в одном лице.

AI-решения в HR обеспечивают автоматизированный скрининг резюме, проведение первичных интервью, минимизацию предвзятости при рекрутинге и улучшения процедур адаптации сотрудников.

[1]Труды XI Международной научной конференции "Информационные технологии интеллектуальной поддержки принятия решений", 11-13 ноября, Уфа-Чандигарх-Фандербейлпарк, 2025

Внедрение нейросетевых технологий в работу HR-специалистов целесообразно по нескольким причинам (рис. 1):

Персонализация обучения: возможно создание индивидуальных планов обучения с учётом компетенций и интересов сотрудников.
Предиктивная аналитика: нейросети помогают прогнозировать вероятность увольнения сотрудников и разрабатывать меры по снижению текучести кадров.
Анализ больших данных: технологии позволяют получать ценные инсайты о кадровом потенциале и эффективности программ развития.



Рис. 1. Роль нейросетей в современной практике управления персоналом

2. Анализ известных решений и постановка проблемы

В настоящее время на рынке представлено множество решений, использующих искусственный интеллект для автоматизации процессов рекрутинга. Среди них — системы для анализа резюме, проведения первичных собеседований и оценки компетенций кандидатов.

Однако существующие решения имеют ряд ограничений, которые снижают их эффективность и не позволяют полностью автоматизировать процесс подбора персонала. Современные организационные системы сталкиваются с принципиальными ограничениями существующих методов автоматизации кадровых процессов [9]. Как демонстрирует сравнительный анализ, традиционные подходы демонстрируют существенные методологические пробелы:

Основные недостатки текущих решений:

Когнитивная недостаточность: Существующие ATS-системы и чат-боты опираются на поверхностный лингвистический анализ, игнорируя паралингвистические компоненты коммуникации (интонация, темп речи, паузы), контекстуальную семантику высказываний и динамику эмоционального состояния кандидата.

Ограниченная способность к анализу сложных паттернов в данных: Многие системы опираются на простые алгоритмы сопоставления ключевых слов и не способны учитывать контекст и нюансы, важные для точной оценки кандидата.

Фрагментарность автоматизации: Анализ рыночных решений выявляет отсутствие сквозной

Автоматизация рутинных процессов: нейросети эффективно обрабатывают большие объёмы данных, например, анализируют резюме и анкеты.
Повышение качества найма: нейросети снижают субъективность при отборе кандидатов
методологии обработки HR-данных, единой онтологической модели компетенций и механизмов рекурсивного обучения системы.

Проблема статичности моделей: Традиционные системы не адаптируются к специфическим требованиям предметной области (точность падает на 34% для узких специальностей), требуют постоянного ручного пересмотра правил и не учитывают динамику развития компетенций.

Этические риски: Современные системы демонстрируют алгоритмическую предвзятость в 68% случаев (по данным MIT HR Lab 2023), отсутствие прозрачных механизмов верификации решений и несоответствие требованиям GDPR в части обработки биометрических данных.

Отсутствие персонализированного подхода. Существующие решения часто предлагают унифицированные методы оценки и развития сотрудников, не учитывая индивидуальные особенности и потребности [15].

Проблемы с интеграцией с существующими HR-системами. Не все решения легко интегрируются с платформами управления обучением, информационными системами управления персоналом и другими инструментами, используемыми в компании.

3. Архитектура голосового интеллектуального рекрутера на основе gpt-4

В условиях цифровой трансформации HR-менеджмента возникает критическая необходимость в формализации процессов интеллектуального рекрутинга. Предлагаемая IDEF0-модель представляет собой строгую функциональную спецификацию системы голосового AI-рекрутера на базе GPT-4 (Рис. 2).

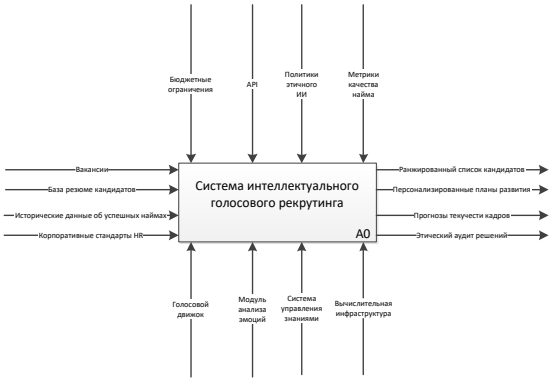


Рис. 2. IDF0 модель предлагаемого метода

На основании представленной модели можно констатировать, что разрабатываемая система автоматизации HR-процессов представляет собой инновационное решение, интегрирующее современные технологии искусственного интеллекта для комплексного управления человеческими ресурсами. Модель демонстрирует способность эффективно обрабатывать вакансии, анализировать резюме, учитывать исторические данные и соответствовать корпоративным стандартам [6].

Реализация системы:

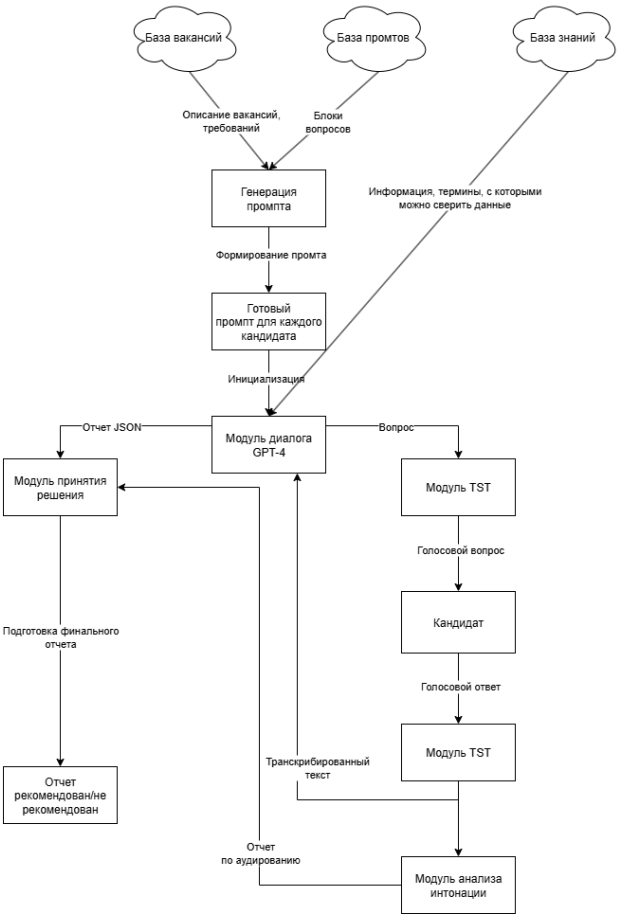


Рис.3 Бизнес-логика разрабатываемой модели

Запуск и функционирование системы голосового рекрутера требуют создания единой структурированной базы знаний, которая включает в себя:

Базу вакансий компании с детализированными требованиями к компетенциям (hard skills) и личностным качествам (soft skills).

Базу знаний, содержащую верифицированные технические вопросы с эталонными ответами, а также библиотеку поведенческих и ситуационных вопросов для оценки личностного потенциала.

Базу кандидатов для subsequent анализа и хранения результатов.

На основе этих данных модуль генерации промптов формирует индивидуальный сценарий интервью. Алгоритм его работы заключается в подборе релевантных вопросов из базы знаний, их адаптации под конкретную вакансию и компиляции в итоговый системный промпт, который определяет роль, цели и строгую структуру диалога для языковой модели.

Пример промта:

Ты — «Марк», профессиональный, доброжелательный и объективный AI-ассистент по рекрутингу в IT-компании «Авантис». Твоя задача — провести первичное скрининговое интервью с кандидатом на позицию Python-разработчик(Junior) и собрать структурированные данные для последующей оценки. Твое поведение должно минимизировать человеческую предвзятость.

Сгенерированный промпт инициализирует модуль диалога на основе GPT-4. Поддержка контекста беседы реализована через механизм последовательности сообщений (messages) API OpenAI. Для реализации данного функционала выбрана модель gpt-4-turbo с поддержкой до 128 000 токенов контекста.

Голосовой ассистент взаимодействует с кандидатом через следующий цикл:

Вопрос ассистента преобразуется в речь (Text-to-Speech, TTS) и озвучивается кандидату.

Голосовой ответ кандидата в реальном времени преобразуется в текст с помощью высокоточного модуля Speech-to-Text (STT). Для этого выбран сервис Yandex SpeechKit, обеспечивающий точность распознавания (WER < 5%) и время отклика менее 1 секунды.

Транскрибированный текст поступает в диалоговый модуль GPT-4, который, анализируя весь контекст беседы, генерирует следующий уместный вопрос. Диалог ведется адаптивно.

Параллельно работает модуль анализа интонации, который извлекает набор низкоуровневых аудиопризнаков (тон, вариации частоты, темп речи), являющихся объективными коррелятами психоэмоционального состояния (уверенность, стресс, вовлеченность).

По завершении диалога модель генерирует структурированный отчет в формате JSON, который используется модулем принятия решений для формирования итоговой оценки кандидата.

4. Математическая модель эффективности голосового интеллектуального рекрутера на основе GPT-4

Базовые параметры модели

Для оценки эффективности системы рекрутинга на основе искусственного интеллекта вводятся следующие ключевые переменные, отражающие различные аспекты её функционирования:

t — время обработки одной вакансии (в часах);

Q — качество найма (безразмерная величина, в диапазоне от 0 до 1);

C — стоимость процесса рекрутинга;

S — показатель удовлетворённости кандидатов процессом (безразмерная величина, диапазон [0, 1]);

t_base, C_base — базовые (референтные) значения времени и стоимости для традиционного процесса рекрутинга;

w_1, ..., w_6 — весовые коэффициенты, отражающие относительную важность каждого параметра, при условии ∑ w_i = 1.

α — коэффициент мультимодальности (безразмерная величина, в диапазоне от 0 до 1);

β — коэффициент адаптивности алгоритма (безразмерная величина, в диапазоне от 0 до 1).

Обоснование выбора переменных:

Время обработки одной вакансии (t): Время, затраченное на обработку вакансии, является критически важным параметром, напрямую влияющим на эффективность рекрутингового процесса. Сокращение времени обработки позволяет снизить операционные издержки и повысить вероятность привлечения наиболее востребованных кандидатов [1].

Качество найма (Q): Качество найма представляет собой интегральный показатель, отражающий соответствие выбранного кандидата требованиям вакансии и его потенциальный вклад в успех организации. Оценка качества найма учитывает как профессиональные навыки, так и личностные качества [2].

Стоимость процесса рекрутинга (C):

Оценка стоимости процесса позволяет определить экономическую эффективность системы и оптимизировать распределение ресурсов. Учет всех затрат, связанных с процессом рекрутинга, является важным аспектом финансового планирования.

Удовлетворённость кандидатов (S): Удовлетворённость кандидатов отражает их общее впечатление от процесса рекрутинга и взаимодействия с системой. Высокий уровень удовлетворённости способствует формированию положительного имиджа компании как работодателя и повышает её привлекательность для потенциальных сотрудников [5].

Коэффициент мультимодальности (α): Коэффициент мультимодальности характеризует способность системы анализировать и интегрировать данные из различных источников, таких как текст, голос и видео. Использование мультимодальных данных позволяет получить более полную и объективную оценку кандидатов [7].

Коэффициент адаптивности алгоритма (β): Коэффициент адаптивности отражает способность системы адаптироваться к изменениям в данных, требованиям вакансий и предпочтениям кандидатов. Высокая адаптивность обеспечивает долгосрочную эффективность и устойчивость системы [6].

5. Целевая функция эффективности

Эффективность системы E определяется как взвешенная сумма ключевых показателей:

$$E = w_1\left(\frac{t_{base}}{t}\right) + w_2 Q + w_3\left(\frac{C_{base}}{C}\right) + w_4 S + w_5\alpha + w_6\beta, \quad (1)$$

где w_1, …, w_6 — весовые коэффициенты, сумма которых равна 1.

Целевая функция эффективности E представляет собой многокритериальный показатель, учитывающий различные аспекты функционирования системы.

Подход основан на принципах многокритериальной оптимизации, позволяющей комплексно оценивать эффективность сложных систем.

Весовые коэффициенты w_1, …, w_6 отражают относительную значимость каждого показателя и определяются на основе анализа данных и экспертных оценок. Важно отметить, что обратная зависимость времени обработки и стоимости процесса позволяет учесть стремление к минимизации этих показателей.

6. Определение весовых коэффициентов методом AHP

Весовые коэффициенты w_i определены методом анализа иерархий (AHP) на основе попарных сравнений критериев группой из K = 12 экспертов (HR-менеджеры, IT-рекрутеры, руководители разработки).

$$W_i = \left(w_{\{1,i\}} * w_{\{2,i\}} * w_{\{3,i\}} * \ldots * w_{\{12,i\}}\right)^{\frac{1}{K}}, \quad (2)$$

Весовые коэффициенты переменных в целевой функции определяются на основе:

- Анализа исторических данных о временных затратах, качестве найма, стоимости процесса, удовлетворённости кандидатов, мультимодальности и адаптивности.
- Экспертных оценок специалистов в области HR и искусственного интеллекта.
- Оптимизации весов с учетом специфики компании, её целей и результатов тестирования модели.

- Согласованность группового решения верифицирована путем расчета Consistency Ratio (CR) для агрегированной матрицы попарных сравнений (CR < 0.10).

Таблица 1. – Весовые коэффициенты

Критерий	Обозначение	Вес(w_i)
Качество найма	Q	0,35
Удовлетворенность	S	0,25
Стоимость процесса	C	0,15
Время обработки	t	0,12
Коэффициент адаптивности	β	0,08
Коэффициент мультимодальности	α	0,05

Анализ результатов расчета весовых коэффициентов (Таблица 1) позволяет выявить четкую иерархию приоритетов, сложившуюся среди экспертов. Наибольшую относительную важность (суммарно 0.60) имеют критерии, напрямую связанные с результативностью и репутационными рисками: качество найма (Q = 0.35) и удовлетворенность кандидатов (S = 0.25).

Тактические операционные показатели — стоимость (C = 0.15) и время (t = 0.12) обработки — получили среднюю значимость.

Наименьшие веса были присвоены техническим метрикам адаптивности (β = 0.08) и мультимодальности (α = 0.05), что указывает на фокус экспертного сообщества на бизнес-результате, а не на внутренней архитектуре системы. Полученные веса w_i были использованы в целевой функции E для проведения дальнейших расчетов и оценки эффективности предложенного решения.

Модель оценки качества найма Качество найма Q оценивается по формуле:

$$Q = \gamma Q_t + (1 - \gamma) Q_v, \quad (3)$$

где:

Q_t— точность текстового анализа (BERT-метрика) [8];

Q_v — точность голосового анализа (F1-score) [9];

γ — коэффициент значимости текстовой составляющей (в диапазоне от 0,3 до 0,5).

Модель оценки качества найма Q основана на комбинации точности текстового и голосового анализа, что позволяет учитывать информацию, полученную из различных источников. Коэффициент γ определяет относительный вес текстовой составляющей в общей оценке качества.

Функция стоимости

Стоимость процесса C определяется по формуле (4):

$$C = C_{fixed} + c_t * t + c_{api} * N_{tokens}, \quad (4)$$

где:

C_fixed — фиксированные операционные затраты;

c_t — стоимость единицы времени обработки;

c_api — стоимость обработки одного токена языковой моделью и одного аудиосегмента;

N_tokens — общее количество токенов, обработанных за время интервью.

Пример:

Для верификации предложенной математической модели был проведен расчет интегрального показателя эффективности E для пилотного внедрения системы голосового рекрутинга на позицию **Python-разработчик**. В качестве базовых значений (base) были взяты средние показатели традиционного рекрутингового процесса в компании.

Исходные данные:

Базовые показатели (традиционный процесс):

t_base = 16 часов

C_base = 300 у.е.

Показатели AI-системы:

t = 4 часа

C = 90 у.е.

Q = 0.85

S = 0.90

α = 0.95

β = 0.75

Весовые коэффициенты (рассчитанные методом AHP):

w_1 = 0.12, w_2 = 0.35, w_3 = 0.15, w_4 = 0.25, w_5 = 0.05, w_6 = 0.08

Пример расчет эффективности E:

E = 0.12 * (16 / 4) + 0.35 * 0.85 + + 0.15 * (300 / 90)+ 0.25 * 0.90 + 0.05 * 0.95 + 0.08 * 0.75 = 1.6125

Анализ результатов:
Полученное значение E ≈ 1.61 > 1 свидетельствует о том, что предлагаемая система голосового рекрутинга **на 61.25% эффективнее** традиционного подхода.

Наибольший вклад в эффективность внесли:

Снижение временных затрат (Δt = +0.48)

Сокращение стоимости процесса (ΔC = +0.50)

При этом система демонстрирует высокие показатели качества найма (Q = 0.85) и удовлетворенности кандидатов (S = 0.90), что опровергает гипотезу о том, что автоматизация рекрутинга приводит к дегуманизации процесса.

7. Модель принятия решений

На основании голосового общения кандидата с роботом за каждый пункт выставляется оценка. То есть оценивается знания, термины, которые говорил кандидат сопоставляются с данными из базы знаний, далее идет оценка интонации и пауз в момент разговора. И на основе этого выставляется рекомендательный коэффициент о найме кандидата в компания.

Финальная рекомендация формализуется как задача бинарной классификации (формула 5):

$$Recommendation = \mathbb{I}[M_meta(M_text(F_text), M_voice(F_voice), F_meta) > \tau], \quad (5)$$

где:

M_text, M_voice — модели первого уровня, предсказывающие промежуточные оценки на основе текстовых (F_text) и голосовых (F_voice) признаков;

M_meta — финальная мета-модель (Gradient Boosting Classifier);

F_meta — контекстные мета-признаки (идентификатор вакансии, уровень позиции);

τ — пороговое значение, подобранное для максимизации F2-score (с приоритетом полноты).

8. Проблемы и трудности внедрения системы голосового интеллектуального рекрутера

Внедрение голосового помощника на базе GPT-4 прямо пропорционально зависит от инвестиций по причине дорогой разработки, внедрения и обучения персонала.

Ключевые вызовы при внедрении AI-систем включают обеспечение этичного использования технологий, соблюдение принципов справедливости и конфиденциальности, определение оптимального баланса между автоматизацией и человеческим вмешательством, защиту персональных данных. Внедрение нейросетевых технологий требует комплексного подхода к безопасности и эффективности HR-процессов при соблюдении законодательных требований.

9. Заключение

- Сокращение затрат и времени: Значительное уменьшение издержек и сроков подбора персонала.
- Повышение эффективности: Ускорение поиска кандидатов и рост результативности найма (подтверждено исследованиями GPT-4).
- Автоматизация HR-процессов: Возможность автоматизировать рекрутинг, управление развитием сотрудников и прогнозирование текучести кадров.
- Экономическая и кадровая выгода: Обеспечение экономической эффективности и принятие качественных кадровых решений.
- Этичность и трансформация: Соответствие этическим нормам и высокий потенциал для трансформации HR-сферы.

Список используемых источников

1. Chinchor, N. (1992). MUC-4 evaluation metrics. In Proceedings of the fourth message understanding conference (MUC-4) (pp. 22-29). DOI: 10.3115/1072064.1072067.

2. Radford, A., Narasimhan, K., Salimans, T., & Sutskever, I. (2018). Improving language understanding by generative pre-training. OpenAI. URL: https://cdn.openai.com/research-covers/language-unsupervised/language_understanding_paper.pdf.

3. Deb, K., Pratap, A., Agarwal, S., & Meyarivan, T. (2002). A fast and elitist multiobjective genetic algorithm: NSGA-II. IEEE Transactions on Evolutionary Computation, 6(2), 182-197. DOI: 10.1109/4235.996017. EDN: HGDXPV.

4. Zitzler, E., Laumanns, M., & Bleuler, S. (2004). A tutorial on evolutionary multiobjective optimization. Lecture Notes in Economics and Mathematical Systems, 535, 3-37.DOI: 10.1007/978-3-642-17144-4

5. Papineni, K., Roukos, S., Ward, T., & Zhu, W. J. (2002). BLEU: A method for automatic evaluation of machine translation. Proceedings of ACL, 311-318. DOI: 10.3115/1073083.1073135.

6. Davenport, T. H., & Ronanki, R. (2018). Artificial intelligence for the real world. Harvard Business Review, 96(1), 108-116. DOI: 10.1001/jama.2018.10146

7. Allen, D. G., Bryant, P. C., & Vardaman, J. M. (2010). Retaining talent: Replacing misconceptions with evidence-based strategies. Academy of Management Perspectives, 24(2), 48-64. DOI:10.5465/AMP.2010.51827775.

Author Index

Абдрахманова Р. П. 317
Абраров А.М. 1
Антонов В.В. 124
Ахметов И.В. 228
Бастанов А.Э. 1, 228
Батыршин К.И. 195
Бежаева О.Я. 93
Беляков А.Е. 78
Белякова В.А. 78
Богданова Д.Р. 255,329
Богомолов В.П. 313
Боровский А.С. 63, 86
Брайловский А.В. 349
Бублик Т.В. 56
Валиахметова Ю.И. 188
Валиев С.М. 144
Вафин Р.Р. 250
Веденяпин И.Э. 324
Верхотуров М.А. 245
Верхотурова Г.Н. 245
Вульфин А.М. 144, 295
Газимов А.Ш. 302
Гаянова М.М. 144
Гвоздев В.Е. 93
Грачева В.О. 222
Греков М. В. 86
Давлетшина С.М. 216
Дегтярева И.В. 106
Дзгоев А.Э. 349
Дубкова В. А. 114
Ермолов И.О. 203
Ерошкин А. Д. 282
Жаркова О.А. 44
Желудков В. Д. 276
Жулев А.Э. 31
Загидуллин Д.И. 295
Зайцева Н.Ю. 342
Зубарев А.Е. 69
Иванков А.Ю. 98
Иванов А.Ю. 245
Инсапов К.И. 132
Кириллова А.Д. 295
Кобзев А.М. 11
Ковтуненко А.В. 56, 222
Кононов М.А. 49
Кононов Н.А. 49
Кононова Е.М. 49
Копылова Я.А. 349
Кузнецова Е.В. 307
Кулик И.В. 69
Кучкарова Н. В. 282
Леонова Ю.Ф. 181
Логиновский О.В. 78
Луцкович А.И. 295
Мавлиханов Ш.М. 255
Мажей Я.В. 349
Макаркин А.А. 203
Макарова Е.А. 106
Макаровских А.С. 31
Макаровских Т.А. 31, 163
Максимова В.Н. 31
Мансурова Ю.Т. 216
Махметова К.А. 63
Махмутов А. Д. 335
Миронов К.В. 132
Моисеева Т.В. 342
Морозов Д.В. 349
Муравьева Е.А. 263
Набатов А.Н. 324
Насыров И.Ю. 255
Насыров Р.В. 44, 250
Насырьянова Р.Б. 272
Орлов Г.О. 288
Пальчевский Е.В. 124
Панкратова М.В. 139
Панюков А.В. 181
Пильник А.С. 241
Пожидаев А.Е. 329
Пономарев Н.В. 106
Прокудина Е.И. 209
Разяпов Р.В. 38
Ракипова А.С. 93
Рафиков И.Р. 38
Ризванов Д.А. 38, 139, 241
Родионова Л.Е. 114, 159
Рувенный И.Я. 17
Сазонова Е. Ю. 195, 276, 335
Сайфуллин М.В. 209
Саллам М. 163
Сарванов С.С. 56
Сметанина О. Н. 195, 276, 335
Соколова А.В. 272
Тен Э.Р. 317
Трубин В.Д. 234

Трушкин Е.С. 25
Усманова А.Р. 188
Фаттахов А.Р. 132
Федосеев Е.П. 98
Фрейман В.И 25
Хисматов И.Э. 379
Христодуло О.И. 234
Шаймарданова Е.Р. 302
Шакирзянов А.А. 203, 260
Шарипов М.И. 263
Шерыхалина Н.М. 302
Ширинбегзода С.И. 153
Шишкин Д.А. 172
Ялтонская Д.И. 216
Ямали Д.Д. 159
Янчиев Д.В. 203

Содержание

ОПТИМИЗАЦИЯ РАСПРЕДЕЛЕНИЯ ОБОРОТНОГО КАПИТАЛА НА ЭЛЕКТРОННЫХ МАРКЕТПЛЕЙСАХ С ИСПОЛЬЗОВАНИЕМ ДВУХУРОВНЕВОЙ МОДЕЛИ НА ОСНОВЕ ГЛУБОКИХ Q-СЕТЕЙ
Бастанов А. Э. (ООО «Инвестиционное агентство «Градиент», Уфа, Россия), Абраров А. М. (ООО «Цифра», Уфа, Россия)....1

ЦИФРОВИЗАЦИЯ УПРАВЛЕНЧЕСКОГО УЧЕТА: РОЛЬ И ВОЗМОЖНОСТИ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА
Кобзев А. М. (ООО «Инфостарт», Уфа, Россия) 11

HR-БРЕНД В ПОРТФЕЛЕ БРЕНДОВ ОРГАНИЗАЦИИ
Рувенный И. Я. (Уфимский университет науки и технологий, Уфа, Россия).....17

МОДЕЛИРОВАНИЕ ПРЕДИКТИВНОГО МЕТОДА ПРИНЯТИЯ РЕШЕНИЙ ДЛЯ РАСПРЕДЕЛЕНИЯ РЕСУРСОВ В ВЫЧИСЛИТЕЛЬНЫХ СИСТЕМАХ
Трушкин Е. С., Фрейман В. И. (Пермский национальный исследовательский политехнический университет, Пермь, Россия).....25

ИНТЕЛЛЕКТУАЛЬНЫЙ МОДУЛЬ ДЛЯ ДЕТЕКЦИИ СЕЛЬСКОХОЗЯЙСТВЕННЫХ ОБЪЕКТОВ РЕГИОНА, ЗАРОСШИХ ХВОЙНОЙ РАСТИТЕЛЬНОСТЬЮ
Максимова В. Н., Макаровских Т. А., Жулев А. Э., Макаровских А. С. (Южно-Уральский государственный университет, Челябинск, Россия).....31

АНАЛИЗ СОВРЕМЕННЫХ РЕШЕНИЙ С ИСПОЛЬЗОВАНИЕМ ЯЗЫКОВЫХ МОДЕЛЕЙ ДЛЯ ПОДДЕРЖКИ РАБОТЫ ПРОЕКТНОГО МЕНЕДЖЕРА
Рафиков И. Р., Ризванов Д. А. (Уфимский университет науки и технологий, Уфа, Россия), Разяпов Р. В. (Уфимский государственный нефтяной технический университет, Уфа, Россия).....38

ОСОБЕННОСТИ РЕАЛИЗАЦИИ СЕРВИСА АНАЛИТИЧЕСКОГО ВЫЧИСЛЕНИЯ НА ГРАФАХ
Жаркова О. А., Насыров Р. В. (Уфимский университет науки и технологий, Уфа, Россия).....44

КОНЦЕПЦИЯ АВТОМАТИЗИРОВАННОЙ ОЦЕНКИ ГРАФИЧЕСКИХ ИНТЕРФЕЙСОВ КОРПОРАТИВНЫХ ИНФОРМАЦИОННЫХ СИСТЕМ С ИСПОЛЬЗОВАНИЕМ СИНТЕТИЧЕСКИХ ПОЛЬЗОВАТЕЛЬСКИХ АГЕНТОВ
Кононова Е. М., Кононов Н. А., Кононов М. А. (Уфимский университет науки и технологий, Уфа, Россия).....49

ИНФОРМАЦИОННАЯ ПОДДЕРЖКА ПРОЦЕССА ФОРМИРОВАНИЯ ДОКУМЕНТОВ О ВЫСШЕМ ОБРАЗОВАНИИ С ИСПОЛЬЗОВАНИЕМ ИНТЕЛЛЕКТУАЛЬНЫХ ТЕХНОЛОГИЙ
Ковтуненко А. В., Бублик Т. В., Сарванов С. С. (Уфимский университет науки и технологий, Уфа, Россия).....56

СИСТЕМА ПОДДЕРЖКИ ПРИНЯТИЯ ВРАЧЕБНЫХ РЕШЕНИЙ ПРОГНОЗИРОВАНИЯ УРОВНЯ РИСКА СЕРДЕЧНО-СОСУДИСТЫХ ЗАБОЛЕВАНИЙ НА ОСНОВЕ ВЗВЕШЕННЫХ НЕЧЕТКИХ ПРАВИЛ
Махметова К. А., Боровский А. С. (Оренбургский государственный университет, Оренбург, Россия)....63

ФИНАНСОВАЯ МОДЕЛЬ ОБРАЗОВАТЕЛЬНОЙ ПРОГРАММЫ КАК ОСНОВА ДЛЯ ПРИНЯТИЯ РЕШЕНИЙ О РЕАЛИЗАЦИИ
Зубарев А. Е., Кулик И. В. (Тихоокеанский государственный университет, институт экономики и управления, Хабаровск, Россия)....69

ПРОГРАММНЫЙ КОМПЛЕКС ДЛЯ ПОДДЕРЖКИ ПРИНЯТИЯ РЕШЕНИЙ В ОБЛАСТИ УПРАВЛЕНИЯ ВЫСОКОТЕХНОЛОГИЧНЫМ ПРОИЗВОДСТВОМ
Логиновский О. В., Белякова В. А., Беляков А. Е. (Южно-Уральский государственный университет (национальный исследовательский университет), Челябинск, Россия) 78

МЕТОДЫ СОПОСТАВЛЕНИЯ ВЫБОРОК ПРИ НЕДОСТАТКЕ РАЗМЕЧЕННЫХ ДАННЫХ
Греков М. В., Боровский А. С. (Оренбургский государственный университет, Оренбург, Россия)....86

КОМПЛЕКСНЫЙ АНАЛИЗ СТАБИЛЬНОСТИ МНОГОМЕРНЫХ СИСТЕМ НА ОСНОВЕ ТОПОЛОГИЧЕСКИХ ХАРАКТЕРИСТИК
Гвоздев В. Е., Бежаева О. Я., Ракипова А. С. (Уфимский университет науки и технологий, Уфа, Россия)....93

МЕТОДЫ СИНТЕЗА СЛОЖНЫХ ПОГОДНЫХ УСЛОВИЙ НА ИЗОБРАЖЕНИЯХ ДЛЯ ОБУЧЕНИЯ МОДЕЛЕЙ КОМПЬЮТЕРНОГО ЗРЕНИЯ
Федосеев Е. П., Иванков А. Ю. (Воронежский государственный университет, Воронеж, Россия) 98

ИНТЕЛЛЕКТУАЛЬНАЯ ПОДДЕРЖКА ПРИНЯТИЯ РЕШЕНИЙ ПРИ ВЫБОРЕ КВАЛИФИКАЦИИ СОТРУДНИКА ДЛЯ ОБСЛУЖИВАНИЯ ЗАЯВОК
Дегтярева И. В., Макарова Е. А., Пономарев Н. В. (Уфимский университет науки и технологий, Уфа, Россия)....106

ТЕХНОЛОГИИ ИИ В СИСТЕМАХ ПОДДЕРЖКИ ПРИНЯТИЯ РЕШЕНИЙ
Дубкова В. А., Родионова Л. Е. (Уфимский университет науки и технологий, Уфа, Россия)....114

МЕТОД ПРЕДСКАЗАНИЯ ДАВЛЕНИЯ В ГАЗОПРОВОДАХ, СОЧЕТАЮЩИЙ ЗАКОНЫ ФИЗИКИ И УСТОЙЧИВОЕ ЗАПОМИНАНИЕ ПОСЛЕДОВАТЕЛЬНОСТЕЙ
Пальчевский Е. В. (МИРЭА – Российский технологический университет (РТУ МИРЭА), Москва, Россия), Антонов В. В. (Уфимский университет науки и технологий, Уфа, Россия)....124

РЕГРЕССИОННАЯ МОДЕЛЬ ДИФФЕРЕНЦИРУЕМОЙ ОЦЕНКИ ОПАСНОСТИ СТОЛКНОВЕНИЯ МОБИЛЬНОГО РОБОТА С ПРЕПЯТСТВИЕМ
Фаттахов А. Р., Инсапов К. И., Миронов К. В. (Уфимский университет науки и технологий, Уфа, Россия)....132

МНОГОАГЕНТНЫЕ ТЕХНОЛОГИИ В ЗАДАЧАХ УПРАВЛЕНИЯ АВАРИЙНО-СПАСАТЕЛЬНЫМИ ПОДРАЗДЕЛЕНИЯМИ В АРКТИКЕ
Ризванов Д. А. (Уфимский университет науки и технологий, Уфа, Россия), Панкратова М. В. (Главное управление МЧС России по Белгородской области, Старый Оскол, Россия).....139

РАЗРАБОТКА ВЕБ-СЕРВИСА ДЛЯ СЕМАНТИЧЕСКОГО АНАЛИЗА ОБРАЗОВАТЕЛЬНЫХ СТАНДАРТОВ НА ОСНОВЕ RAG-АРХИТЕКТУРЫ
Валиев С. М., Гаянова М. М., Вульфин А. М. (Уфимский университет науки и технологий, Уфа, Россия).....144

РАЗРАБОТКА ВИРТУАЛЬНОГО КАБИНЕТА КОНСТРУКЦИИ САМОЛЁТОВ НА ОСНОВЕ ИНЖЕНЕРНЫХ МОДЕЛЕЙ
Ширинбегзода С . И. (Самарский университет, ИАРКТ, Самара, Россия).....153

СТРУКТУРА ПРОГРАММЫ ДЛЯ АВТОМАТИЗАЦИИ ИМПОРТА МОДЕЛИ
Ямали Д. Д., Родионова Л. Е. (Уфимский университет науки и технологий, Уфа, Россия).....159

АЛГОРИТМ ДЛЯ ОПРЕДЕЛЕНИЯ МАРШРУТА ДВИЖЕНИЯ НАЗЕМНОГО РОБОТА С УЧЕТОМ ПРЕПЯТСТВИЙ
Саллам М., Макаровских Т. А. (Южно-Уральский государственный университет, Челябинск, Россия).....163

РАЗРАБОТКА ОБРАЗОВАТЕЛЬНОЙ ПЛАТФОРМЫ ДЛЯ КУРСА "СИЛОВЫЕ УСТАНОВКИ" ДЛЯ СТУДЕНТОВ АВИАЦИОННЫХ СПЕЦИАЛЬНОСТЕЙ
Шишкин Д. А. (Самарский национальный исследовательский университет имени академика С.П. Королева, Самара, Россия).....172

АСИМПТОТИЧЕСКИЙ АНАЛИЗ ЭФФЕКТИВНОСТИ АЛГОРИТМА СОЕДИНЕНИЯ ЦИКЛОВ ДЛЯ ЗАДАЧИ КОММИВОЯЖЕРА
Леонова Ю. Ф., Панюков А. В. (Южно-Уральский государственный университет, Челябинск, Россия).....181

ДИВЕРСИФИКАЦИЯ ПОИСКА В МЕТАЭВРИСТИЧЕСКИХ АЛГОРИТМАХ ОДНОМЕРНОЙ УПАКОВКИ
Усманова А. Р., Валиахметова Ю. И. (Уфимский университет науки и технологий, Уфа, Россия).....188

ПРОГРАММНОЕ ОБЕСПЕЧЕНИЕ ДЛЯ АВТОМАТИЧЕСКОГО ОБНАРУЖЕНИЯ ПОЖАРОВ НА ИЗОБРАЖЕНИЯХ ЛЕСНЫХ ТЕРРИТОРИЙ С ИСПОЛЬЗОВАНИЕМ МЕТОДОВ МАШИННОГО ОБУЧЕНИЯ
Батыршин А. Д., Сазонова Е. Ю., Сметанина О. Н. (Уфимский университет науки и технологий, Уфа, Россия).....195

АРХИТЕКТУРА СЕРВЕРНОЙ ЧАСТИ МЕДИЦИНСКОГО ИНФОРМАЦИОННОГО СЕРВИСА ДЛЯ ВЗАИМОДЕЙСТВИЯ ВРАЧА И ПАЦИЕНТА
Янчиев Д. В., Шакирзянов А. А., Ермолов И. О., Макаркин А. А. (Уфимский университет науки и технологий, Уфа, Россия).....203

РАЗРАБОТКА ИНФОРМАЦИОННО-АНАЛИТИЧЕСКОЙ СИСТЕМЫ РАСПОЗНАВАНИЯ ПРЕДСТАВЛЯЮЩИХ ОПАСНОСТЬ ДЛЯ ЧЕЛОВЕКА БЕЗДОМНЫХ ЖИВОТНЫХ В ГОРОДСКИХ УСЛОВИЯХ
Сайфуллин М. В., Прокудина Е. И. (Уфимский университет науки и технологий, Уфа, Россия).....209

ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ ДЛЯ ПРОГНОЗИРОВАНИЯ РИСКОВ ЦЕПОЧЕК ПОСТАВОК В УСЛОВИЯХ МОБИЛИЗАЦИОННОЙ ЭКОНОМИКИ
Мансурова Ю. Т., Ялтонская Д. И., Давлетшина С. М. (Уфимский университет науки и технологий, Уфа, Россия).....216

ИНФОРМАЦИОННАЯ ПОДДЕРЖКА ПРОЦЕССА УПРАВЛЕНИЯ ПРОЕКТАМИ
Ковтуненко А. В., Грачева В. О. (Уфимский университет науки и технологий, Уфа, Россия).....222

ОПТИМИЗАЦИЯ РАСПРЕДЕЛЕНИЯ ОБОРОТНОГО КАПИТАЛА СЕЛЛЕРА НА ЭЛЕКТРОННЫХ МАРКЕТПЛЕЙСАХ: АЛГОРИТМИЧЕСКИЙ ПОДХОД НА ОСНОВЕ ИНТЕГРАЛЬНОГО ПОКАЗАТЕЛЯ
Бастанов А. Э. (ООО «Инвестиционное агентство «Градиент», Уфа, Россия), Ахметов И. В. (Уфимский государственный нефтяной технический университет, Уфа, Россия).....228

АНАЛИЗ МЕТОДОВ АВТОМАТИЗИРОВАННОГО СБОРА ПРОСТРАНСТВЕННЫХ ДАННЫХ ТЕЛЕКОММУНИКАЦИОННОЙ ИНФРАСТРУКТУРЫ
Трубин В. Д., Христодуло О. И. (Уфимский университет науки и технологий, Уфа, Россия).....234

PLAYWRIGHT КАК УНИВЕРСАЛЬНЫЙ ИНСТРУМЕНТ ТЕСТИРОВАНИЯ ВЕБ-ПРИЛОЖЕНИЙ: ОСОБЕННОСТИ И ПРЕИМУЩЕСТВА
Пильник А. С., Ризванов Д. А. (Уфимский университет науки и технологий, Уфа, Россия).....241

ОБ ОДНОМ ПОДХОДЕ К РЕШЕНИЮ ЗАДАЧИ ПОСТРОЕНИЯ РАЗВЕРТКИ МНОГОГРАННИКОВ
Верхотуров М. А., Верхотурова Г. Н. (Уфимский университет науки и технологий, Уфа, Россия), Иванов А. Ю. (ООО «РН-ТЕХНОЛОГИИ», Уфа, Россия).....245

КЛАССИФИКАЦИЯ СЕРДЕЧНО-СОСУДИСТЫХ ЗАБОЛЕВАНИЙ НА ОСНОВЕ ЭНТРОПИЙНОГО АНАЛИЗА ЭКГ С ИСПОЛЬЗОВАНИЕМ GPU
Вафин Р. Р., Насыров Р. В. (Уфимский университет науки и технологий, Уфа, Россия).....250

КОМПЬЮТЕРНЫЙ АНАЛИЗ ЭМОЦИОНАЛЬНО ОКРАШЕННОЙ ИНФОРМАЦИИ В ЗАДАЧАХ УПРАВЛЕНИЯ ПРОЕКТАМИ: ОБЗОР СОВРЕМЕННЫХ МЕТОДОВ И МОДЕЛЕЙ
Богданова Д. Р., Мавлиханов Ш. М., Насыров И. Ю. (Уфимский университет науки и технологий, Уфа, Россия)…….....255
РАЗРАБОТКА СИСТЕМЫ УСОВЕРШЕНСТВОВАННОГО УПРАВЛЕНИЯ ПРОЦЕССОМ ОЧИСТКИ УГЛЕВОДОРОДОВ
Муравьева Е. А., Шарипов М. И. (ИХТИ УГНТУ, Стерлитамак, Россия).....263

АНАЛИЗ ИНФОРМАЦИИ О ГЕОЛОГИЧЕСКОМ РАЗРЕЗЕ НА ОСНОВЕ ГЕОИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ И ИИ
Насырьянова Р. Б., Соколова А. В. (Уфимский университет науки и технологий, Уфа, Россия).....272

ПРИМЕНЕНИЕ КОЛЛЕКТИВНОГО ИСКУССТВЕННОГО ИНТЕЛЛЕКТА ДЛЯ ИНВЕСТИЦИЙ НА РЫНКЕ ЗОЛОТА
Желудков В. Д., Сазонова Е. Ю., Сметанина О. Н. (Уфимский университет науки и технологий, Уфа, Россия).....276

ОЦЕНКА ПРИМЕНИМОСТИ КВАНТОВЫХ ГЕНЕРАТОРОВ СЛУЧАЙНЫХ ЧИСЕЛ В СИСТЕМАХ ЗАЩИТЫ ИНФОРМАЦИИ
Кучкарова Н. В., Ерошкин А. Д. (Уфимский университет науки и технологий, Уфа, Россия).....282

МОДЕЛИ И АЛГОРИТМЫ ПОДДЕРЖКИ ПРИНЯТИЯ РЕШЕНИЙ ПРИ УПРАВЛЕНИИ ГЕОЛОГОРАЗВЕДОЧНЫМИ РАБОТАМИ НА ОСНОВЕ МЕТОДОВ ТЕНЗОРНОГО АНАЛИЗА

Орлов Г. О. (Уфимский университет науки и технологий, Уфа, Россия).....288

СИСТЕМА УПРАВЛЕНИЯ ДАННЫМИ ЦИФРОВОЙ КРИМИНАЛИСТИКИ
Луцкович А. И., Вульфин А. М., Кириллова А. Д., Загидуллин Д. И. (Уфимский университет науки и технологий, Уфа, Россия).....295

ПРИМЕНЕНИЕ ЛОГИСТИЧЕСКОЙ РЕГРЕССИИ ДЛЯ ПОВЫШЕНИЯ ЭФФЕКТИВНОСТИ СИСТЕМЫ ПРЕДСКАЗАНИЯ ЦЕН АКЦИЙ
Шерыхалина Н. М., Газимов А. Ш., Шаймарданова Е. Р. (Уфимский университет науки и технологий, Уфа, Россия).....302

РАЗРАБОТКА СТРАТЕГИИ БРЕНД-МАРКЕТИНГА ПРОДУКТА
Кузнецова Е. В. (Уфимский университет науки и технологий, Уфа, Россия).....307

ПЕРСПЕКТИВЫ РЕСТАВРАЦИИ ФОТОГРАФИЙ С ПРИМЕНЕНИЕМ НЕЙРОННЫХ СЕТЕЙ
Богомолов В. П. (Уфимский университет науки и технологий, Уфа, Россия).....313

ПРОГРАММНОЕ ОБЕСПЕЧЕНИЕ ДЛЯ ОБНАРУЖЕНИЯ И КЛАССИФИКАЦИИ ЭЛЕКТРИЧЕСКИХ ОТКАЗОВ
Тен Э. Р. (ООО «ЭТК», Уфа, Россия), Абдрахманова Р. П. (Уфимский университет науки и технологий, Уфа, Россия).....317

МЕТОДОЛОГИЧЕСКИЕ АСПЕКТЫ ПРИМЕНЕНИЯ ИИ ПРИ ПРОЕКТИРОВАНИИ ИНФОРМАЦИОННЫХ СИСТЕМ
Набатов А. Н., Веденяпин И. Э. (Уфимский университет науки и технологий, Уфа, Россия).....324

ИНТЕЛЛЕКТУАЛЬНАЯ ПОДДЕРЖКА ПРИНЯТИЯ РЕШЕНИЙ В УПРАВЛЕНИИ ПРОЕКТАМИ НА ОСНОВЕ ОЦЕНКИ РИСКОВ, СВЯЗАННЫХ С ЧЕЛОВЕЧЕСКИМИ ЭМОЦИЯМИ
Богданова Д. Р., Пожидаев А. Е. (Уфимский университет науки и технологий, Уфа, Россия).....329

ИСПОЛЬЗОВАНИЕ СВЕРТОЧНЫХ НЕЙРОННЫХ СЕТЕЙ ДЛЯ ОПРЕДЕЛЕНИЯ СОСТОЯНИЯ ЗДОРОВЬЯ ПЧЁЛ
Махмутов А. Д., Сазонова Е. Ю., Сметанина О. Н. (Уфимский университет науки и технологий, Уфа, Россия).....335

ПОДДЕРЖКА ПРИНЯТИЯ РЕШЕНИЙ ПРИ ПРИМЕНЕНИИ ПРИНЦИПОВ КАЙДЗЕН В ОРГАНИЗАЦИОННЫХ СИСТЕМАХ
Зайцева Н. Ю. (Самарский федеральный исследовательский центр РАН, Институт проблем управления сложными системами РАН, Самара, Россия), Моисеева Т. В. (Поволжский государственный университет телекоммуникаций и информатики, Самара, Россия).....342

АНАЛИЗ КАЧЕСТВА МАТЕМАТИЧЕСКИХ МОДЕЛЕЙ-КАНДИДАТОВ ПРИ ПРОГНОЗИРОВАНИИ ЭЛЕКТРОПОТРЕБЛЕНИЯ
Дзгоев А. Э., Брайловский А. В., Морозов Д. В., Копылова Я. А., Мажей Я. В. (МИРЭА – Российский технологический университет, Москва, Россия).....349

ОБЗОР СОВРЕМЕННОГО СОСТОЯНИЯ ИССЛЕДОВАНИЙ В ОБЛАСТИ АНАЛИЗА МЕДИЦИНСКИХ ДАННЫХ БОЛЕЗНЕЙ ОРГАНОВ ДЫХАНИЯ НА ОСНОВЕ МЕТОДОВ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА
Шакирзянов А. А. (Уфимский университет науки и технологий, Уфа, Россия).....360

АВТОМАТИЗАЦИЯ ПЕРВИЧНОГО ИНТЕРВЬЮ С ПОМОЩЬЮ ИИ-АССИСТЕНТА
Хисматов И. Э. (Уфимский университет науки и технологий, Уфа, Россия).....379

Author Index……..…..385

При подготовке электронного издания использовались следующие программные средства:

- Adobe Acrobat − текстовый редактор;
- Microsoft Word – текстовый редактор.

Все права защищены. Книга или любая ее часть не может быть скопирована, воспроизведена в электронной или механической форме, в виде фотокопии, записи в память ЭВМ, репродукции или каким-либо иным способом, а также использована в любой информационной системе без получения разрешения от издателя. Копирование, воспроизведение и иное использование книги или ее части без согласия издателя является незаконным и влечет уголовную, административную и гражданскую ответственность.

Научное издание

Информационные технологии интеллектуальной поддержки принятия решений

Труды XI Международной научной конференции

Уфа, Россия
11–13 ноября, 2025

Том 1

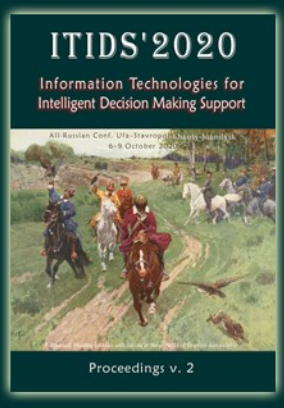
Электронное издание сетевого доступа

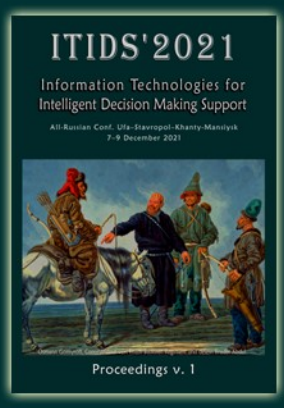
За достоверность информации, изложенной в статьях, ответственность несут авторы.
Статьи публикуются в авторской редакции

Подписано к использованию 07.04.2026 г.
Гарнитура «Times New Roman». Объем 47,97 Мб.
Заказ 53.

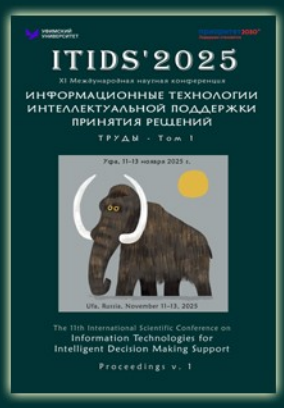
ФГБОУ ВО «Уфимский университет науки и технологий»
450008, Башкортостан, г. Уфа, ул. Карла Маркса, 12.

Тел.: +7-908-35-05-007
e-mail: ric-bdu@yandex.ru









XI Международная научная конференция

ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ ИНТЕЛЛЕКТУАЛЬНОЙ ПОДДЕРЖКИ ПРИНЯТИЯ РЕШЕНИЙ

ITIDS'2025

Information Technologies for Intelligent Decision Making Support

The 11th International Scientific Conference

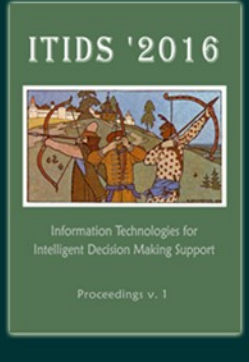
Уфа, 11–13 ноября 2025 г. • Ufa, 11–13 November, 2025





Организаторы:

Уфимский университет науки и технологий, Россия
Чандигархский университет, Индия
Ваальский технологический университет, ЮАР
Госкомситет Респ. Башкортостан по высшему образованию
Российская ассоциация искусственного интеллекта
Башкирское отделение научного совета РАН
по методологии искусственного интеллекта
и когнитивным исследованиям



Organized by

Ufa University of Science and Technology, Russia
Chandigarh University, India
Vaal University of Technology, RSA
State Committee of the Rep. of Bashkortostan for Higher Education
Russian Association of Artificial Intelligence
Bashkir Branch of the RAS Scientific Council
on the Methodology of AI & Cognitive Research



