



№4-2026

ISSN 1999-9429

ИЗВЕСТИЯ ЮФУ

ТЕХНИЧЕСКИЕ НАУКИ

- Алгоритмы обработки информации
- Анализ данных и моделирование

ИЗВЕСТИЯ ЮФУ. ТЕХНИЧЕСКИЕ НАУКИ IZVESTIYA SFedU. ENGINEERING SCIENCES

Свидетельство о регистрации средства массовой информации

ПИ № ФС77-28889 от 12.07.2007

*Федеральная служба по надзору в сфере массовых коммуникаций, связи
и охраны культурного наследия*

Научно-технический и прикладной журнал

Издается с 1995 года, до середины 2007 года под названием «Известия ТРТУ»

Подписной индекс ПС704

№ 4 (252). 2026 г.

Журнал включен в «Перечень рецензируемых научных изданий, в которых должны быть опубликованы основные научные результаты диссертаций на соискание ученой степени кандидата наук, на соискание ученой степени доктора наук».

Редакционный совет

Курейчик В.В. (гл. редактор); Кравченко Ю.А. (зам. гл. редактора); Бородинский И.М. (ученый секретарь); Абрамов С.М.; Агеев О.А.; Бабенко Л.К.; Боженюк А.В.; Борисов В.В.; Веселов Г.Е.; Гайдук А.Р.; Горбанёва О.И.; Еремеев А.П.; Зинченко Л.А.; Каляев И.А.; Касьянов А.О.; Коноплев Б.Г.; Коробейников А.Г.; Кравец А.Г.; Куповых Г.В.; Левин И.И.; Магид Е.А.; Массель Л.В.; Медведев М.Ю.; Мельник Э.В.; Никитов С.А.; Обуховец В.А.; Панич А.Е.; Петров В.В.; Пшихопов В.Х.; Редько В.Г.; Румянцев К.Е.; Сергеев Н.Е.; Середин Б.М.; Сидоркина И.Г.; Стемпковский А.Л.; Сухинов А.И.; Турулин И.И.; Тютиков В.В.; Угольницкий Г.А.; Целых А.Н.; Юханов Ю.В.

Учредитель Южный федеральный университет.

Издатель Южный федеральный университет.

Ответственный за выпуск Клевцов С.И.

Технический редактор Ярошевич Н.В.

Оригинал-макет выполнен Ярошевич Н.В.

Дата выхода в свет 31.08. 2026 г. Формат $70 \times 108 \frac{1}{16}$. Бумага офсетная.

Офсетная печать. Усл. печ. л. – 26,2. Уч.-изд. л. – 19,7.

Заказ № 10570. Тираж 250 экз.

Адрес издателя: 344090, г. Ростов-на-Дону, пр. Стачки, 200/1, тел. 8(863)243-41-66.

Адрес типографии: Отпечатано в отделе полиграфической, корпоративной и сувенирной продукции Издательско-полиграфического комплекса КИБИ МЕДИА ЦЕНТРА ЮФУ. 344090, г. Ростов-на-Дону, пр. Стачки, 200/1, тел. 8(863)243-41-66.

Адрес редакции: 347922, г. Таганрог, ул. Чехова, 22, ЮФУ, тел. +7 (928) 909-57-82, e-mail: iborodyanskiy@sfedu.ru, <http://izv-tn.tti.sfedu.ru/>.

16+

Цена свободная

ISSN 1999-9429 (Print)

ISSN 2311-3103 (Online)

© Южный федеральный университет, 2026

СОДЕРЖАНИЕ

РАЗДЕЛ I. АЛГОРИТМЫ ОБРАБОТКИ ИНФОРМАЦИИ

Е.С. Абрамов, Н.Е. Белов, Г.Е. Веселов АЛГОРИТМ КОНТЕКСТУАЛЬНОЙ ВАЛИДАЦИИ ПОВЕДЕНЧЕСКИХ ИНДИКАТОРОВ (ЮВ) ДЛЯ ДЕТЕКТИРОВАНИЯ СКРЫТЫХ АТАК В АСУ ТП	6
А.А. Александров, М.А. Бутакова ПОВЫШЕНИЕ КАЧЕСТВА РАЗНОМОДАЛЬНЫХ ДАННЫХ НА ОСНОВЕ ГИБРИДИЗАЦИИ МЕТОДОВ ПРЕДВАРИТЕЛЬНОЙ ОБРАБОТКИ И МУЛЬТИСЕНСОРНОГО СЛИЯНИЯ	21
Н.Н. Войт МОДИФИЦИРОВАННЫЙ АЛГОРИТМ ИНТЕГРАЦИИ СИСТЕМЫ УПРАВЛЕНИЯ ЖИЗНЕННЫМ ЦИКЛОМ ИЗДЕЛИЯ С ВИРТУАЛЬНОЙ ПЛОЩАДКОЙ ПРОЕКТИРОВАНИЯ	38
В.З. Гергов МУЛЬТИМОДАЛЬНОЕ РАСПОЗНАВАНИЕ ДЕЙСТВИЙ ЧЕЛОВЕКА НА ОСНОВЕ НЕЗАВИСИМОГО ОБУЧЕНИЯ ВИДЕОМОДЕЛИ И АУДИОМОДЕЛИ И ПОЗДНЕГО ОБЪЕДИНЕНИЯ ПРЕДСКАЗАНИЙ	51
К.С. Григорян, Е.С. Басан АЛГОРИТМ ПРОТИВОДЕЙСТВИЯ ПРЕДНАМЕРЕННЫМ ПОМЕХАМ НА ОСНОВЕ ТЕОРИИ ИГР И ОБУЧЕНИЯ С ПОДКРЕПЛЕНИЕМ	66
С.В. Ермоленко, И.Л. Каширина УЛУЧШЕНИЕ СЕГМЕНТАЦИИ НА МУЛЬТИФАЗНЫХ КТ СНИМКАХ С ПОМОЩЬЮ ОБУЧАЕМОГО ФАЗОВОГО СОВМЕЩЕНИЯ	77
И.Е. Лысенко МАЛОРЕСУРСНАЯ АБСТРАКТИВНАЯ СУММАРИЗАЦИЯ ТЕКСТОВ С ИСПОЛЬЗОВАНИЕМ КОНТРАСТИВНОГО ОБУЧЕНИЯ БЕЗ УЧИТЕЛЯ И ROUGE-ОРИЕНТИРОВАННОЙ ФУНКЦИИ ПОТЕРЬ	91
С.И. Носков, А.П. Медведев, И.Д. Кириллов ПРИМЕНЕНИЕ МЕТОДА СМЕШАННОГО ОЦЕНИВАНИЯ ПАРАМЕТРОВ ПРИ ПОСТРОЕНИИ ОДНОРОДНОЙ ВЛОЖЕННОЙ КУСОЧНО-ЛИНЕЙНОЙ РЕГРЕССИИ ПЕРВОГО ТИПА	102
Т.Р. Гамберов, Р.Н. Сафин, Э.В. Чеботарева, Т.Г. Цой, Е.А. Магид ОЦЕНКА АЛГОРИТМА СЛЕДОВАНИЯ МОБИЛЬНОГО РОБОТА TIAGO BASE ЗА ПОЛЬЗОВАТЕЛЕМ С ПРИМЕНЕНИЕМ ФУНКЦИОНАЛЬНЫХ ЦИФРОВЫХ МОДЕЛЕЙ ЧЕЛОВЕКА В СИМУЛЯТОРЕ GAZEBO	109
А.Д. Скорик АЛГОРИТМ ВЫБОРА ОПТИМАЛЬНОЙ РАБОЧЕЙ ЧАСТОТЫ И ПАРАМЕТРОВ ПЕРЕДАЧИ В ДКМ-КАНАЛЕ НА ОСНОВЕ ПРОГНОЗИРОВАНИЯ ПО ДАННЫМ ГНСС-МОНИТОРИНГА	120
М.М. Сорокин, Л.А. Гладков, Н.В. Гладкова СОВРЕМЕННЫЕ ПОДХОДЫ К РЕШЕНИЮ ЗАДАЧИ 3D УПАКОВКИ	131
В.С. Стародубцев, Л.К. Бабенко АЛГОРИТМЫ СНИЖЕНИЯ ВРЕМЕННЫХ ЗАТРАТ НА ВЫПОЛНЕНИЕ ОПЕРАЦИЙ КРИПТОСИСТЕМЫ ДОМИНГО-ФЕРРЕРА	150
А.Н. Целых, И.А. Валухов МЕТОД НЕЙРОСЕТЕВОГО ИЗВЛЕЧЕНИЯ ПРИЧИННО-СЛЕДСТВЕННЫХ СВЯЗЕЙ НА ОСНОВЕ СРАВНИТЕЛЬНОГО АНАЛИЗА СИНТЕТИЧЕСКИХ И ОТКРЫТЫХ КОРПУСОВ	160

РАЗДЕЛ II. АНАЛИЗ ДАННЫХ И МОДЕЛИРОВАНИЕ

М.Н. Бейтуганов, Х.Х. Чочаев РОЛЬ МОЛНИЙ В ПЕРЕХОДЕ ОБЛАКА ИЗ ГРОЗОВОГО СОСТОЯНИЯ В ГРАДОВОЕ И ИХ ИСКУССТВЕННОЕ ИНИЦИИРОВАНИЕ КАК МЕТОД ПРЕДОТВРАЩЕНИЯ ГРАДООБРАЗОВАНИЯ И ПЕРЕРАСПРЕДЕЛЕНИЯ ОСАДКОВ	170
О.В. Белоусова ГЛОБАЛЬНЫЕ И ЛОКАЛЬНЫЕ ВАРИАЦИИ ЭЛЕКТРИЧЕСКОГО ПОЛЯ В ПРИЗЕМНОМ ЭЛЕКТРОДНОМ СЛОЕ.....	180
И.С. Бурлаченко, Г.М. Чечин РЕЗОНАНСНОЕ ВОЗБУЖДЕНИЕ БУШЕЙ НЕЛИНЕЙНЫХ НОРМАЛЬНЫХ МОД В ДВУМЕРНЫХ РЕШЁТКАХ.....	191
Ж.-М.Н. Дакуо, С.В. Беззатеев, В.В. Давыдов, И.Д. Иогансон, А.Ф. Хуцаева МОДЕЛЬ СХЕМЫ ЭЛЕКТРОННОГО ГОЛОСОВАНИЯ С ВОЗМОЖНОСТЬЮ ИЗМЕНЕНИЯ РЕШЕНИЯ УЧАСТНИКОМ ДО МОМЕНТА ОКОНЧАНИЯ ГОЛОСОВАНИЯ.....	201
А.Н. Каркищенко КОЛИЧЕСТВЕННЫЕ ОЦЕНКИ ПОГРЕШНОСТИ ПОДДЕРЖАНИЯ СТРОЯ ПОДВОДНЫХ АППАРАТОВ ПРИ ЗАДЕРЖКАХ НАВИГАЦИОННЫХ ДАННЫХ	212
Д.А. Ковалёв, Д.Х. Имаев, С.Е. Душин МОДЕЛИРОВАНИЕ И ИДЕНТИФИКАЦИЯ СОДОРЕГЕНЕРАЦИОННОГО КОТЛОАГРЕГАТА КАК ОБЪЕКТА УПРАВЛЕНИЯ	226
А.В. Мартиросян, Д.В. Ромашин КОНЦЕПТУАЛЬНАЯ МОДЕЛЬ ФАКТОРОВ, ВЛИЯЮЩИХ НА ЭФФЕКТИВНОСТЬ ПРОЦЕССА ПОДГОТОВКИ И СЕПАРАЦИИ ГАЗА.....	241
В.Х. Пшихопов, А.Л. Веревкин, А.Р. Гайдук, Л.С. Веревкина МАТРИЧНЫЙ МЕТОД ОЦЕНКИ НАДЁЖНОСТИ СЛОЖНЫХ МУЛЬТИАГЕНТНЫХ РЕКОФИГУРИРУЕМЫХ СИСТЕМ	249
Э.М. Эндреев, В.А. Шаповалов, А.А. Аджиева, Г.В. Куповых, В.В. Гривцов СУХИЕ ВТОРЖЕНИЯ В СРЕДНЕЙ ТРОПОСФЕРЕ КАК ФАКТОР ФОРМИРОВАНИЯ ОПАСНЫХ НИСХОДЯЩИХ КОНВЕКТИВНЫХ ПОТОКОВ: АНАЛИЗ И ДИАГНОСТИКА НА ЮГЕ ЕВРОПЕЙСКОЙ ЧАСТИ РОССИИ.....	260
А.В. Боженюк, И.А. Дубчак, О.В. Косенко ПРИМЕНЕНИЕ НЕЧЕТКОГО МЕТОДА TOPSIS ДЛЯ ПРИНЯТИЯ РЕШЕНИЙ С ИСПОЛЬЗОВАНИЕМ СИНТАКСИЧЕСКИ НЕЗАВИСИМЫХ ЛИНГВИСТИЧЕСКИХ ПЕРЕМЕННЫХ.....	272
Е.А. Титенко, Т.М. Белова, Л.А. Лисицин, И.И. Пузанов, А.А. Положенцев МАТЕМАТИЧЕСКАЯ МОДЕЛЬ И СХЕМА УПРАВЛЕНИЯ ПРОДУКЦИЯМИ ДЛЯ РЕАЛИЗАЦИИ ПАРАЛЛЕЛЬНЫХ ВЫЧИСЛЕНИЙ	283

CONTENT

SECTION I. INFORMATION PROCESSING ALGORITHMS

E.S. Abramov, N.E. Belov, G.E. Veselov ALGORITHM FOR CONTEXTUAL VALIDATION OF INDICATORS OF BEHAVIOR (IOB) FOR DETECTING STEALTHY ATTACKS IN ICS	6
A.A. Alexandrov, M.A. Butakova IMPROVING THE QUALITY OF MULTI-MODAL DATA BASED ON THE HYBRIDIZATION OF PREPROCESSING AND MULTISENSORY FUSION METHODS	21
N.N. Voit MODIFY INTEGRATION ALGORITHM BETWEEN CONTROL SYSTEM OF PART LIFE CYCLE AND VIRTUAL DESIGN PLATFORM.....	38
V.Z. Gergov MULTIMODAL HUMAN ACTION RECOGNITION BASED ON INDEPENDENTLY TRAINED VIDEO AND AUDIO MODELS WITH LATE FUSION OF PREDICTIONS	51
K.S. Grigoryan, E.S. Basan GAME-THEORETIC AND REINFORCEMENT LEARNING-BASED ALGORITHM FOR INTENTIONAL JAMMING MITIGATION.....	66
S.V. Ermolenko, I.L. Kashirina IMPROVING SEGMENTATION IN MULTIPHASE CT IMAGES USING TRAINABLE PHASE SUPERIMAGING	77
I.E. Lysenko LOW-RESOURCE ABSTRACTIVE TEXT SUMMARIZATION VIA CONTRASTIVE UNSUPERVISED LEARNING WITH A ROUGE-ORIENTED LOSS FUNCTION	91
S.I. Noskov, A.P. Medvedev, I.D. Kirillov APPLICATION OF THE MIXED PARAMETER ESTIMATION METHOD IN THE CONSTRUCTION OF A HOMOGENEOUS NESTED PIECEWISE LINEAR REGRESSION OF THE FIRST TYPE	102
T.R. Gamberov, R.N. Safin, E.V. Chebotareva, T.G. Tsoy, E.A. Magid EVALUATING THE TIAGO BASE MOBILE ROBOT'S PERSON-FOLLOWING ALGORITHM USING REALISTIC DIGITAL HUMAN MODELS IN THE GAZEBO SIMULATOR.....	109
A.D. Skorik ALGORITHM FOR SELECTING THE OPTIMAL OPERATING FREQUENCY AND TRANSMISSION PARAMETERS IN THE HF CHANNEL BASED ON PREDICTION USING GNSS MONITORING DATA	120
M.M. Sorokin, L.A. Gladkov, N.V. Gladkova MODERN APPROACHES TO SOLVING THE 3D BIN PACKING PROBLEM.....	131
V.S. Starodubcev, L.K. Babenko ALGORITHMS FOR REDUCING THE TIME REQUIRED TO PERFORM OPERATIONS OF THE DOMINGO-FERRER CRYPTOSYSTEM.....	150
A.N. Tselykh, I.A. Valukhov A NEURAL NETWORK-BASED METHOD FOR EXTRACTING CAUSAL RELATIONS USING COMPARATIVE ANALYSIS OF SYNTHETIC AND OPEN CORPORA	160

SECTION II. DATA ANALYSIS AND MODELING

M.N. Beituganov, Kh.Kh. Chochaev ROLE OF LIGHTNING IN THE CLOUD TRANSITION FROM THUNDERSTORM TO HURRICANE AND THEIR ARTIFICIAL INITIATION AS A METHOD OF PREVENTING HURRICANE AND PRECIPITATIONS REDISTRIBUTION	170
O.V. Belousova GLOBAL AND LOCAL VARIATIONS OF THE ELECTRIC FIELD IN THE SURFACE ELECTRODE LAYER.....	181

I.S. Burlachenko, G.M. Chechin RESONANT EXCITATION OF BUSH NONLINEAR NORMAL MODES IN TWO-DIMENSIONAL LATTICES	191
Z.-M.N. Dakuo, S.V. Bezzateev, V.V. Davydov, I.D. Ioganson, A.F. Khutsaeva MODEL OF AN ELECTRONIC VOTING SCHEME WITH THE POSSIBILITY OF CHANGING THE PARTICIPANT'S DECISION BEFORE THE END OF VOTING	202
A.N. Karkishchenko QUANTITATIVE ESTIMATES OF THE ERROR IN MAINTAINING THE FORMATION OF UNDERWATER VEHICLES WITH DELAYS IN NAVIGATION DATA	213
D.A. Kovalev, D.H. Imaev, S.E. Dushin MODELING AND IDENTIFICATION OF A COGENERATION BOILER UNIT FOR KRAFT PULP PRODUCTION AS A CONTROL PLANT.....	226
A.V. Martirosyan, D.V. Romashin CONCEPTUAL MODEL OF FACTORS INFLUENCING THE EFFICIENCY OF GAS PREPARATION AND SEPARATION PROCESS	242
V.Kh. Pshikhopov, A.L. Verevkin, A.R. Gaiduk, L.S. Verevkina MATRIX METHOD FOR SIMULTANEOUS RELIABILITY ASSESSMENT OF COMPLEX MULTI-AGENT RECONFIGURABLE SYSTEMS	249
E.M. Andreyev, V.A. Shapovalov, A.A. Adzhieva, G.V. Kupovkh, V.V. Grivtsov DRY INTRUSIONS IN THE MIDDLE TROPOSPHERE AS A FACTOR IN THE FORMATION OF DANGEROUS DOWNWARD CONVECTIVE FLOWS: ANALYSIS AND DIAGNOSTICS IN THE SOUTH OF THE EUROPEAN PART OF RUSSIA	260
A.V. Bozhenyuk, I.A. Dubchak, O.V. Kosenko APPLICATION OF TOPSIS FUZZY METHOD FOR DECISION MAKING USING SYNTACTICALLY INDEPENDENT LINGUISTIC VARIABLES	273
E.A. Titenko, T.M. Belova, L.A. Lisitsin, I.I. Puzanov, A.A. Polozhenev A MATHEMATICAL MODEL AND PRODUCT MANAGEMENT SCHEME FOR IMPLEMENTING PARALLEL PRODUCT COMPUTING.....	283

Раздел I. Алгоритмы обработки информации

УДК 004.056.57

DOI 10.18522/2311-3103-2026-4-6-20

Е.С. Абрамов, Н.Е. Белов, Г.Е. Веселов

АЛГОРИТМ КОНТЕКСТУАЛЬНОЙ ВАЛИДАЦИИ ПОВЕДЕНЧЕСКИХ ИНДИКАТОРОВ (IOB) ДЛЯ ДЕТЕКТИРОВАНИЯ СКРЫТЫХ АТАК В АСУ ТП

Рассматривается актуальная проблема обеспечения информационной безопасности критической информационной инфраструктуры (КИИ) в условиях качественной эволюции киберугроз и массового перехода злоумышленников к скрытым атакам типа «Living-off-the-Land» (LotL). Ввиду того, что подобные атаки реализуются через легитимные инструменты администрирования, классические индикаторы компрометации (IOC) теряют свою эффективность, а традиционные системы мониторинга генерируют избыточное количество ложных тревог, провоцируя эффект «усталости от алертов». Для решения данной проблемы в работе предложена методика контекстуальной валидации поведенческих индикаторов (IOB) в сетях АСУ ТП. Основным научным результатом является разработанный алгоритм «обратного обогащения», который использует детерминированную природу технологического процесса в качестве строгого априорного фильтра. Алгоритм осуществляет верификацию каждого управляющего воздействия путём прогнозирования следующего состояния системы с помощью математической модели в пространстве состояний и проверки его принадлежности к формализованному множеству безопасных значений Ω_{safe} . Дополнительно учитывается организационный контекст S_{org} , включающий регламенты смен и окна техобслуживания. В отличие от вероятностных подходов машинного обучения, данный метод даёт строгий бинарный критерий допустимости команды. Эффективность предложенного подхода подтверждена результатами моделирования на базе верифицированного датасета академического стенда водоочистки SWaT. Внедрение физического и организационного фильтров позволило снизить коэффициент ложных срабатываний (FPRR) на 97,7 % и успешно выявить 1021 из 1035 внедрённых скрытых деструктивных воздействий. Вычислительная задержка алгоритма составила в среднем 1,4 мс, что полностью удовлетворяет строгим требованиям систем реального времени. Предложенный метод не требует замены устаревшего оборудования и обеспечивает точную атрибуцию киберинцидентов на основе законов физики производственного цикла.

Критическая информационная инфраструктура (КИИ); безопасность АСУ ТП; индикаторы атаки (IOA); Living-off-the-Land (LotL); предиктивное моделирование; снижение ложных срабатываний (FPRR); MITRE ATT&CK for ICS.

E.S. Abramov, N.E. Belov, G.E. Veselov

ALGORITHM FOR CONTEXTUAL VALIDATION OF INDICATORS OF BEHAVIOR (IOB) FOR DETECTING STEALTHY ATTACKS IN ICS

This article addresses the pressing issue of ensuring the information security of critical information infrastructure (CII) amidst the qualitative evolution of cyber threats and the massive shift of threat actors toward stealthy "Living-off-the-Land" (LotL) attacks. Because such attacks are executed using legitimate administration tools, classical indicators of compromise (IOCs) lose their effectiveness, and traditional monitoring systems generate an excessive number of false alarms, thereby provoking "alert fatigue". To solve this problem, the paper proposes a methodology for the contextual validation of indicators of behavior (IOB) in Industrial Control Systems (ICS) networks. The primary scientific result is the developed "reverse enrichment" algorithm, which utilizes the deterministic nature of the technological process as a strict a priori filter. The algorithm verifies every control action by predicting the next state of the system using a state-space mathematical model and checking its membership in a formalized set of safe values, Ω_{safe} . Additionally, the organizational context, S_{org} , including shift schedules and maintenance windows, is taken

into account. Unlike probabilistic machine learning approaches, this method provides a strict binary criterion for command admissibility. The effectiveness of the proposed approach is confirmed by simulation results based on the verified dataset of the SWaT (Secure Water Treatment) academic testbed. The implementation of physical and organizational filters achieved a 97.7% False Positive Reduction Rate (FPRR) and successfully detected 1,021 out of 1,035 injected stealthy destructive impacts. The average computational latency of the algorithm was 1.4 ms, which fully satisfies the stringent requirements of real-time systems. The proposed method does not require the instrumentation of legacy field equipment and ensures the precise attribution of cyber incidents based on the physical laws of the production cycle.

Critical information infrastructure (CII); ICS security; indicators of attack (IOA); Living-off-the-Land (LotL); predictive modeling; false positive reduction rate (FPRR); MITRE ATT&CK for ICS.

Введение. В условиях современной киберугрозы традиционные оборонительные стратегии для промышленных объектов стремительно теряют свою актуальность. Мы наблюдаем качественную эволюцию тактик атакующих – на смену авторскому, целевому вредоносному коду приходят методы «Living-off-the-Land» (LotL), эксплуатирующие штатный функционал системы для скрытного деструктивного воздействия [1, 2]. Использование легитимных утилит администрирования и инженерного ПО лишает смысла поиск классических индикаторов компрометации (IOC), таких как контрольные суммы файлов (хэши) или списки блокировки IP-адресов, которые просто не фиксируют работу доверенных инструментов [3].

Переход к мониторингу поведенческих индикаторов (IOB) стал ответом индустрии на возросшую скрытность атак, однако внедрение этого подхода в сетях АСУ ТП столкнулось с серьезными барьерами. Высокая волатильность данных на полевом уровне и фрагментарность журналов событий ПЛК делают действия легитимного оператора и атакующего практически идентичными для автоматизированных средств защиты [4]. В отсутствие глубокого понимания технологического контекста системы безопасности провоцируют эффект «усталости от алертов», перегружая персонал ложноположительными срабатываниями и, как следствие, повышая риск пропуска реальных инцидентов [5, 6].

Анализ существующих подходов показывает, что методы машинного обучения [7], [8] обладают высокой чувствительностью к отклонениям, но не содержат механизма различения вредоносного воздействия и штатной технологической операции, а также уязвимы к подмене показаний датчиков [9]. В свою очередь, киберфизические методы [10, 11, 13] и вероятностные подходы [14] либо требуют глубокой инструментовки устаревшего оборудования [15], либо не дают бинарного (детерминированного) критерия физической допустимости управляющего воздействия. Темпоральные [16] и IOC-ориентированные методы [17, 18] также фокусируются на поиске статистических отклонений, а не на проверке физической корректности команд.

Таким образом, существует критический пробел в методах детектирования: отсутствие инструмента, способного объединить абстрактное цифровое событие с его физическим смыслом без сложной модернизации контроллеров. В данной статье предлагается алгоритм «обратного обогащения» (reverse enrichment) поведенческих индикаторов. Его суть заключается в использовании детерминированных параметров техпроцесса (физических ограничений оборудования и организационных регламентов) в качестве априорного фильтра. Алгоритм осуществляет строгую проверку каждого управляющего воздействия на вхождение прогнозируемого состояния системы в заранее формализованное безопасное множество (Ω_{safe}) с учётом организационного контекста (S_{org}).

Целью настоящей работы является разработка методики и алгоритма контекстуальной валидации поведенческих индикаторов, основанного на проверке бинарной допустимости каждого управляющего воздействия через физические и организационные ограничения технологического процесса для надежного детектирования атак типа LotL.

1. Описание и постановка решаемой задачи. Современные системы защиты автоматизированных систем управления технологическими процессами (АСУ ТП) сталкиваются с качественно новым классом угроз. Злоумышленники всё чаще отказываются от применения специализированного вредоносного программного обеспечения, переходя к тактикам класса LotL [1, 2]. Подобные атаки реализуются исключительно через легитим-

ные инструменты администрирования и штатное инженерное программное обеспечение, что делает классические индикаторы компрометации – хеши файлов, IP-адреса, сигнатуры эксплойтов – практически бесполезными [3]. По оценкам SANS, в 2023–2024 гг. доля LotL-атак в сетях АСУ ТП остаётся стабильно высокой, а сами инциденты характеризуются возросшей скрытностью и длительным латентным периодом¹.

Закономерной реакцией индустрии стал перенос акцента с поиска статических следов на мониторинг поведенческих индикаторов и индикаторов атаки. Однако в промышленных сетях этот подход порождает принципиальную трудность: данные полевого уровня обладают высокой летучестью, логи контроллеров фрагментарны, а действия оператора АСУ ТП зачастую неотличимы от действий злоумышленника [4]. Без учёта контекста технологического процесса средства обнаружения генерируют поток ложных тревог, вызывая у операторов центров мониторинга «усталость от алертов» и провоцируя пропуск реальных инцидентов [5, 6].

Обзор существующих подходов и их ограничения. В современной научной литературе можно выделить несколько направлений, нацеленных на решение обозначенной проблемы. Каждое из них обладает определёнными достоинствами, но демонстрирует принципиальные ограничения, не позволяющие полноценно решить задачу валидации поведенческих индикаторов в АСУ ТП.

Методы, основанные на машинном обучении без явного учёта физики процесса. Значительная часть работ опирается на глубокие нейронные сети и графовые модели для поиска аномалий в потоках данных. Так, в исследовании Атарской и соавторов [7] предложена гетерогенная модель детекторов на основе автоэнкодеров и LSTM-сетей для анализа технологических временных рядов. Работа демонстрирует высокую чувствительность к отклонениям, однако не содержит механизма различения вредоносного воздействия и штатной технологической операции – любое статистически редкое событие интерпретируется как аномалия. Аналогичный недостаток присущ графовому анализатору АНГА [Lyu], который, несмотря на глубокий проблемно-ориентированный анализ ассоциативных связей между гетерогенными устройствами, не учитывает физические ограничения самого технологического процесса. Как показано в работе [9], процессно-ориентированные детекторы аномалий, не опирающиеся на строгие физические инварианты, оказываются уязвимы для атак типа Synthetic Sensor Spoofing, когда злоумышленник подменяет показания датчиков синтезированными правдоподобными значениями.

Киберфизические методы, учитывающие физические параметры. Ряд исследователей предпринял попытки интеграции физических данных процесса в контур обнаружения вторжений. В работе [10] проведена систематическая оценка эффективности ML-классификаторов при совместном использовании сетевых и физических признаков и показано, что дополнение кибер-данных физическими переменными повышает точность детектирования. В работе [11] предложен метод анализа полезной нагрузки сетевых пакетов (Physics Reasoning) для выявления кода эксплойтов, а в [12] – вариационный автоэнкодер, устойчивый к скрытым киберфизическим атакам. Существенным шагом вперёд является система SCAPHY [13], которая использует граф физических зависимостей и динамический анализ путей исполнения SCADA-приложений для выявления поведений, не соответствующих легитимным фазам управления. Однако SCAPHY, как и большинство перечисленных методов, ориентирован на обнаружение аномалий на хостовом или сетевом уровне и не оперирует организационным контекстом – регламентами смен, нарядами-допусками, календарными графиками технологических операций. Кроме того, данные подходы требуют либо глубокой инструментовки SCADA-узлов, либо наличия полных логов с контроллеров, что трудно реализовать на устаревшем оборудовании, всё ещё доминирующем в промышленности [15].

Значимым шагом в направлении киберфизической валидации является метод доверенного функционирования CPS, предложенный Басан и соавторами [14]. Авторы вводят парадигму киберфизических параметров как признака, наиболее явно отражающего взаимодейст-

¹ SANS ICS/OT Cybersecurity Survey: 2023's Challenges and Tomorrow's Defenses.

вие системы с физическими процессами, и строят вероятностную функцию доверия к текущему управляющему процессу на основе двух метрик: вероятности попадания значений датчиков в доверительный интервал и контроля изменения параметров. В зависимости от значения функции доверия система классифицируется как находящаяся в доверенном, неопределённом или нестабильном состоянии, что позволяет выявлять как отказы, так и целенаправленные атаки. Предложенный подход обладает несомненным достоинством – минимальными требованиями к объёму начальных данных о системе и отсутствием необходимости в эталонных распределениях, а экспериментальная проверка на данных UAV-атаки подтвердила отсутствие ложноположительных оценок аномалии в нормальном режиме полёта.

Однако метод Басан и соавторов принципиально остаётся вероятностным: решение об аномалии принимается на основе количественной оценки доверия, а не бинарной физической допустимости управляющего воздействия. В условиях АСУ ТП, где множество безопасных состояний априори формализовано в технологических регламентах и уставах противоаварийной защиты, более естественным является критерий принадлежности прогнозируемого состояния системы множеству Ω_{safe} , что обеспечивает детерминированную, а не вероятностную классификацию. Кроме того, в отличие от работы [14], предлагаемый алгоритм учитывает не только показания сенсорной подсистемы в текущем режиме, но и организационные ограничения – регламенты смен, наряды-допуски, календарные графики операций, – что позволяет отличить легитимное управляющее воздействие от несанкционированного даже при статистически нормальных показаниях датчиков.

Темпоральные процессно-ориентированные методы. Отдельную группу составляют подходы, эксплуатирующие временные закономерности технологических процессов. В работе [16] предложена система обнаружения вторжений, обучающаяся темпоральному поведению переменных процесса и выявляющая отклонения от нормального ритма. Авторы показывают, что учёт временных свойств позволяет детектировать атаки, невидимые для классических процессно-ориентированных IDS. Однако данный метод опирается на статистические модели нормального поведения, что делает его чувствительным к редким, но легитимным событиям (например, пусконаладочные работы) и не даёт бинарного критерия допустимости конкретного управляющего воздействия.

Методы на основе индикаторов компрометации (IOC). В работе [17] представлен всесторонний анализ проблем использования IOC в промышленных системах. Авторы подчёркивают, что существующие стандарты представления IOC слабо учитывают отраслевую специфику, а процесс создания индикаторов для АСУ ТП требует нестандартных подходов. В развитие этого направления в [18] предложены методы ранжирования и классификации источников IOC по отраслям, а в [19] – скоринг источников индикаторов компрометации для различных индустрий. Однако все перечисленные работы фокусируются на статических индикаторах и не предоставляют механизма автоматической верификации поведенческих паттернов через призму физики технологического процесса.

Таким образом, анализ существующих работ выявляет критический пробел: отсутствие метода, который одновременно:

- ♦ использует детерминированные физические и организационные ограничения техпроцесса как строгий критерий валидации, а не как вероятностную оценку;
- ♦ не требует модификации или инструментовки полевых устройств и контроллеров;
- ♦ обеспечивает бинарную классификацию «легитимное действие / атака» на основе проверки принадлежности прогнозируемого состояния системы безопасному множеству;
- ♦ объединяет в единой процедуре валидации как физические (инерция, пределы безопасности), так и организационные (расписание смен, режимы техобслуживания) ограничения.

Постановка задачи и гипотеза исследования. Целью настоящей работы является разработка методики и алгоритма контекстуальной валидации поведенческих индикаторов, основанного на проверке допустимости каждого управляющего воздействия через физические и организационные ограничения технологического процесса.

Гипотеза исследования состоит в том, что использование детерминированных параметров техпроцесса (регламентов времени, физических ограничений оборудования) в качестве фильтра позволяет:

- ◆ снизить количество ложных срабатываний за счёт автоматического отсеивания легитимных технологических операций;
- ◆ подтвердить тактический замысел атаки даже при отсутствии полных журналов событий на полевых устройствах, компенсируя тем самым техническую отсталость устаревшего оборудования.

Предлагаемый алгоритм «обратного обогащения» (Reverse Enrichment) реализует иной по сравнению с существующими работами принцип: не обучение нормальному поведению процесса и поиск отклонений от него, а бинарную проверку каждого управляющего воздействия на вхождение прогнозируемого состояния системы в заранее формализованное множество безопасных состояний Ω_{safe} . Физические законы и организационные регламенты выступают как априорный, не требующий обучения фильтр, что принципиально отличает данный подход от статистических и ML-ориентированных методов.

2. Методика исследования

Стратегия и подходы к исследованию

Мы применяем гибридный подход к мониторингу. Он объединяет классический анализ сетевых пакетов (DPI) с проверкой физических ограничений технологического процесса. Главная идея – не доверять следам в IT-сети (логам, трафику), пока они не будут проверены через состояние физического оборудования.

Такой подход решает проблему «летучести» данных, описанную в [17]. Даже если злоумышленник стёр логи на сервере, он не может отменить физическую инерцию турбины или изменить расписание рабочей смены в бумажном журнале.

Архитектура данных

Система оперирует тремя типами структур данных, соответствующих уровням модели обнаружения:

1. Событие (Event, e): Сырые данные от сенсоров или зеркалированного трафика:

$$e = \{t, src, dst, proto, func, val\},$$

где t – время, $func$ – код функции (например, *Write Coil* в Modbus), val – значение полезной нагрузки.

2. Контекст (Context, C): Состояние системы в момент времени t .

$$C_t = \{\vec{S}_{phys}, \vec{S}_{org}\},$$

где $\vec{S}_{phys}$ – вектор физических тегов (давление, температура, обороты), а $\vec{S}_{org}$ – организационные метрики (ID текущей смены операторов, режим техобслуживания).

3. Правило валидации (Rule, R): Логическое условие, определяющее допустимость действия.

Для хранения и проверки правил R применяются графовые структуры. В этой модели вершины соответствуют физическим состояниям системы, а рёбра – разрешённым переходам между ними.

Связь с физикой здесь прямая: наличие ребра в графе определяется законами природы и регламентом работы оборудования. Если команда пытается инициировать переход по несуществующему ребру, алгоритм фиксирует нарушение логической целостности процесса [20]. Например, система заблокирует команду на открытие клапана, если основной насос выключен.

Использование графов позволяет обнаруживать сложные многоходовые атаки. В таких сценариях каждое отдельное действие кажется легитимным, но их совокупность ведёт к аварийной ситуации.

Алгоритм «обратного обогащения» (Reverse Enrichment)

Алгоритм назван «обратным», потому что обычно системы идут от сбора событий к поиску аномалий. Авторы идут от фиксации потенциально опасного действия (IOB) к проверке его физической невозможности или недопустимости. Если действие физически опасно или нелогично – мы «обогащаем» исходный сетевой пакет меткой «Инцидент». Выбор метода верификации поведенческих индикаторов (IOB) через проверку физиче-

ской допустимости обусловлен *детерминированной природой технологических процессов* АСУ ТП. В отличие от стохастического поведения пользователей в ИТ-системах, пространство состояний физического объекта управления является конечным и строго ограниченным инвариантными законами физики (инерция, термодинамика, гидравлическое сопротивление) и конструктивными характеристиками оборудования.

Множество недопустимых состояний системы *априори формализовано* в технологических регламентах, картах уставок и алгоритмах систем противоаварийной защиты (ПАЗ). Следовательно, задача обнаружения аномалии сводится к проверке математического условия принадлежности вектора состояния системы множеству безопасных значений ($\vec{S} \in \Omega_{safe}$), что обеспечивает существенно более высокую достоверность детектирования по сравнению с эвристическим поиском сигнатур неизвестных атак в условиях высокой энтропии входящего потока данных.

Формализация множества безопасных состояний Ω_{safe}

Для корректной работы алгоритма необходимо математически формализовать понятие «безопасность». Формализация множества допустимых состояний является развитием идей о форензик-готовности АСУ ТП [20]. Введём вектор физического состояния $S_{phys} = (s_1, s_2, \dots, s_n)$, где каждый элемент – это текущее значение конкретного датчика в системе. Множество Ω_{safe} – это область в n -мерном пространстве, в которой работа объекта считается допустимой.

Мы определим Ω_{safe} через систему строгих ограничений как условие принадлежности вектора состояния множеству безопасных значений:

$$\Omega_{safe} = \{S_{phys} \in \mathbb{R}^m \mid \forall i: g_i(S_{phys}) \leq 0\}, \quad (1)$$

где Ω_{safe} – множество безопасных состояний технологического процесса; S_{phys} – вектор текущего физического состояния системы (набор значений с датчиков); $\mathbb{R}^m$ – n -мерное пространство возможных состояний, где n – количество контролируемых параметров; g_i – набор функций, задающих физические, логические и регламентные ограничения (уставки, скорость изменения параметров, зависимости между узлами).

Функции g_i задают три типа границ, которые нарушитель не может обойти:

1. Статические уставки. Это базовые пределы оборудования. Например, давление в контуре не должно превышать 10 Бар. Если команда ведёт к $s_j > 10$, фильтр её блокирует.
2. Динамические фильтры. Они учитывают инерцию. Физика процесса не даёт температуре или оборотам расти мгновенно. Как отмечает [17], учёт такой волатильности данных – ключ к защите полевого уровня.
3. Логические связи. Это зависимости между узлами. Нельзя открыть клапан подачи хлора, если отключён основной насос. Подобные инварианты крайне сложно подделывать, так как они зашиты в саму конструкцию объекта.

Применение такого фильтра заменяет вероятностный поиск аномалий строгой проверкой. Мы не пытаемся выявить специфический почерк нарушителя. Вместо этого система контролирует соблюдение физических границ процесса. Это обеспечивает детерминированность мониторинга и подтверждает нашу гипотезу.

Предиктивная модель $Apply(B, S_{phys})$

Теперь определим функцию прогнозирования. Она вычисляет, в какое состояние S_{next} перейдёт система после выполнения команды.

Для этого мы используем модель в пространстве состояний в дискретном времени. Применение аппарата пространств состояний для верификации управляющих воздействий согласуется с принципами построения киберфизических детекторов, предложенными в [11].

Математически оператор перехода описывается системой разностных уравнений (2):

$$S_{next} = A \cdot S_{phys} + K \cdot U(B), \quad (2)$$

где S_{next} – прогнозируемый вектор физического состояния; S_{phys} – вектор текущего физического состояния системы; A – матрица собственной динамики системы (описывает инерцию оборудования без внешних команд); K – матрица управления, задающая влияние конкретных команд на объект; $U(B)$ – числовой вектор воздействия, извлечённый из семантического действия B .

Выбор систем разностных уравнений вместо алгоритмов машинного обучения обоснован двумя факторами. Первый – строгая детерминированность вычислений. Второй – высокая вычислительная эффективность. Матричные операции выполняются за доли миллисекунды, что критически важно для исключения задержек при мониторинге АСУ ТП в реальном времени. Если спрогнозированное состояние S_{next} не принадлежит безопасному множеству Ω_{safe} , система классифицирует событие как угрозу.

В ходе исследования авторы теоретически проверили работу алгоритма «обратного обогащения» на двух типовых сценариях атак. Для каждого события вычислялась функция допустимости $\Phi(B, C_t)$, определяющая, является ли действие легитимным (*Safe*) или угрозой (*Threat*).

Сценарий моделирования 1: Атака на физическом уровне (Водоочистка/Химпром)

Контекст: ПЛК управляет подачей реагента (хлора) в систему водоснабжения.

Вектор атаки: Инсайдер или вирус изменяет уставку (Setpoint) в памяти ПЛК, увеличивая дозировку в 10 раз.

Проблема: Сетевой трафик показывает штатную команду записи (Write Register). Полевые устройства не ведут логов «кто и зачем» изменил память.

Работа алгоритма:

1. Система фиксирует изменение значения регистра $val_{new} = 500$ (было 50).
2. Алгоритм применяет функцию прогнозирования $Apply(\vec{B}, \vec{S}_{phys})$.

$$S_{next} = CurrentDosage + \Delta(val_{new}). \quad (3)$$

3. Рассчитанное состояние S_{next} сравнивается с безопасным множеством Ω_{safe} (предельно допустимая концентрация, ПДК).

4. Результат: $S_{next} > Limit_{max}$. Действие физически опасно. Функция Φ возвращает 0. Генерируется инцидент с высоким приоритетом.

Далее описаны шаги алгоритма.

Шаг 1. Агрегация и нормализация

Поток сырых данных очищается от шума. Промышленные протоколы (Modbus, IEC-104, DNP3) разбираются до уровня функциональных кодов.

Проблема: Раскрытие алгоритмов проприетарных промышленных протоколов часто затруднено [21].

Решение: В алгоритме используются только открытые поля заголовков и значения регистров, игнорируя служебные данные вендора.

Шаг 2. Выделение поведенческого паттерна

Сырое событие e отображается на вектор поведения B .

$$B = M(e). \quad (4)$$

Функция M преобразует технический код (например, Modbus Function 5) в семантическое действие (например, Critical_Param_Change). На этом этапе отсеиваются операции чтения (Read), так как они (обычно) не меняют физику процесса.

Шаг 3. Запрос контекста

Для каждого действия B система делает мгновенный срез состояния из базы данных процесса (Historian) или «цифрового двойника».

Мы проверяем вектор C_t на момент времени $t - \Delta t$ и $t + \Delta t$. Это нужно, чтобы понять динамику – как система вела себя до команды и как отреагировала после.

Шаг 4. Доменная фильтрация

Это ядро методики. Мы вычисляем функцию допустимости $\Phi(B, C_t)$. Она возвращает False (аномалия), если действие противоречит физике или регламенту.

Фильтр проверяет, попадёт ли новое состояние в безопасное множество Ω_{safe} :

$$\Phi(B, C_t) = \begin{cases} 1(Safe), & \text{если } Apply(B, S_{phys}) \in \Omega_{safe} \wedge Check(S_{org}) \\ 0(Threat), & \text{иначе} \end{cases} \quad (5)$$

Здесь мы используем предиктивную модель *Apply* и рассчитываем прогнозный вектор состояния системы по модели (2) и сверяем его с границами безопасного множества (1), вычисляя функцию допустимости (5).

Примеры ограничений Ω_{safe} :

- ♦ *Градиентное ограничение*: Скорость изменения температуры не может превышать X градусов в секунду (физическая инерция).
- ♦ *Регламентное ограничение*: Команды управления запрещены, если *Shift_ID* (рабочая смена) не активна.

Сценарий моделирования 2: Атака на организационном уровне (Электроэнергетика)

Контекст: Региональная подстанция управляется удалённо по протоколу IEC-104. Доступ осуществляется через VPN.

Вектор атаки: Злоумышленник похитил валидные учётные данные инженера. Он подключается к сети и отправляет команду на размыкание выключателя (IOB).

Проблема: С точки зрения классических средств защиты (Firewall, IDS), всё чисто. IP-адрес верный, логин верный, протокол не нарушен. Это ложноотрицательное срабатывание (False Negative).

Работа алгоритма:

1. Система перехватывает пакет с функцией «Single Command» (IOB).
2. Алгоритм запрашивает контекст $C_t = \{\vec{S}_{org}\}$. В данном случае $\vec{S}_{org}$ содержит расписание смен дежурных диспетчеров.
3. Проверка условия $\Phi(B, C_t)$: $Check(S_{org}) = (CurrentTime \in ShiftSchedule)$
4. Результат: Время события (03:15) не попадает в плановое окно переключений. Функция возвращает 0 (*Threat*). Событие маркируется как IOA «Несанкционированное управление», несмотря на валидный пароль.

Шаг 5. Расчёт критичности и атрибуция

Если $\Phi = 0$, событие получает статус IOA (индикатор атаки). Дальнейший расчёт скоринга угрозы (*Score*) и прогнозирование развития атаки осуществляются с опорой на матрицы критичности активов и формальные модели компрометации объектов КИИ [22, 23].

$$Score = W_{asset} \times W_{impact}, \quad (6)$$

где W – весовые коэффициенты важности узла и потенциального ущерба.

Ниже представлен псевдокод алгоритма, демонстрирующий логику обработки потока данных.

3. Результаты и их обсуждение

Подготовка экспериментальных данных

Для проведения эксперимента использовался гибридный подход. В качестве базового легитимного трафика был взят верифицированный датасет академического стенда водоочистки SWaT [24]. Однако, поскольку публичные наборы данных редко содержат чистые примеры скрытых атак типа LoTL (где воздействие идёт через штатные команды [25]), исходный массив был дополнен. Мы сформировали итоговую выборку объёмом 100 000 пакетов управляющего трафика (Modbus TCP), где 99% событий составила реальная легитимная активность из датасета SWaT, а 1% – внедрённые нами синтетические аномалии LoTL. В реальных АСУ ТП трафик детерминирован. Команды отдаются по графику или в ответ на изменения среды. Поэтому 99% нормальных команд – это достоверное отражение рутины объекта.

Каждое событие в логе представлено в виде стандартизованного вектора:

[Timestamp, Src_IP, Dst_IP, Func_Code, Register, Value].

Листинг 1 – Псевдокод алгоритма обратного обогащения

```

BEGIN
SET Список_Подтвержденных_Инцидентов to пустой список
FOR each Событие in Поток_Данных:
### Шаг 1: Отсеиваем лишний шум, ищем только управляющие воздействия (IOB)
IF Функция_События is NOT "Команда_Управления":
    CONTINUE with next Событие

### Шаг 2: Запрашиваем состояние системы на момент события
SET Состояние_Системы to ЗАПРОСИТЬ_ИСТОРИЮ_ПРОЦЕССА(Временная_Метка_События)
SET Угроза_Обнаружена to FALSE

### Шаг 3: Доменный фильтр – Организационная проверка
IF Состояние_Системы.Режим_Техобслуживания is FALSE:
    SET Угроза_Обнаружена to TRUE
    SET Описание_Угрозы to "Несанкционированное управление вне расписания"

### Шаг 4: Доменный фильтр – Физическая проверка
CALCULATE Прогнозное_Значение = APPLY_PREDICTIVE_MODEL(Состояние_Системы, Собы-
тие)

IF Прогнозное_Значение exceeds Состояние_Системы.Предел_Безопасности:
    SET Угроза_Обнаружена to TRUE
    SET Описание_Угрозы to "Нарушение физических ограничений техпроцесса"

### Шаг 5: Атрибуция атаки (IOA)
IF Угроза_Обнаружена is TRUE:
    SET Новый_Инцидент to СОЗДАТЬ_ОПОВЕЩЕНИЕ(Событие.Источник, Описание_Угрозы)
    ADD Новый_Инцидент to Список_Подтвержденных_Инцидентов
END FOR
DISPLAY Список_Подтвержденных_Инцидентов
END

```

Значение каждого элемента:

- ◆ Timestamp – точное время фиксации события.
- ◆ Src_IP – IP-адрес источника команды (например, АРМ инженера).
- ◆ Dst_IP – IP-адрес получателя (контроллера ПЛК).
- ◆ Func_Code – код функции протокола (например, 06 – запись в один регистр).
- ◆ Register – адрес конкретного регистра памяти ПЛК, который отвечает за параметр.
- ◆ Value – передаваемое значение (уставка, команда включения).

Для полноценной проверки мы синхронизировали логи трафика с базой организационного контекста (S_{org}). Для каждого Timestamp в датасете определён статус технологического окна: активна ли дежурная смена (Shift_Active: True/False) и идёт ли плановый ремонт (Maintenance: True/False).

Штатный профиль поведения

В нормальном режиме инженерная станция плавно меняет уставку дозирования. Значения находятся в пределах нормы, время совпадает с активной сменой.

```

10:15:01, 192.168.1.10, 10.0.0.5, 06, 40001, 50
10:30:14, 192.168.1.10, 10.0.0.5, 06, 40001, 55
10:45:00, 192.168.1.10, 10.0.0.5, 06, 40001, 60

```

Эти команды классифицируются как легитимные. Прогнозное состояние не выходит за рамки допустимого ($Apply(B, S_{phys}) \in \Omega_{safe}$).

Физическая аномалия типа LotL. Нарушитель захватил АРМ и отправляет штатную команду в рабочее время, но с деструктивной уставкой (имитация Сценария 1).

```
14:20:05, 192.168.1.10, 10.0.0.5, 06, 40001, 60
14:25:11, 192.168.1.10, 10.0.0.5, 06, 40001, 60
14:28:33, 192.168.1.10, 10.0.0.5, 06, 40001, 500 <-- LotL Атака
(Shift_Active = False)
```

Здесь Value 500 означает критическое завышение дозировки реагента (хлора). При норме в 50-60 единиц значение 500 в десять раз превышает предел безопасности. Для обычной IDS это штатная операция записи. Но наш алгоритм фиксирует выход прогнозного вектора за пределы безопасного множества Ω_{safe} . Событие блокируется физическим фильтром.

Организационная аномалия LotL. В этом случае нарушитель отправляет физически безопасную команду (уставка 60), но делает это глубокой ночью, имитируя Сценарий 2.

```
03:15:22, 192.168.1.10, 10.0.0.5, 06, 40001, 60 <-- LotL Атака
(Shift_Active = False)
```

Значение 60 не нарушает физику процесса и находится в пределах Ω_{safe} . Базовая IDS полностью игнорирует пакет. Однако наш алгоритм запрашивает контекст S_{org} . Проверка условия $Check(S_{org})$ возвращает ноль, так как активной смены нет. Событие блокируется до физического исполнения.

Процедура эксперимента и оценка снижения шума (FPRR)

Для проведения эксперимента использовался верифицированный датасет со стенда SWaT (Secure Water Treatment). Этот набор данных содержит записи реальных технологических процессов и зафиксированных атак на систему очистки воды. Мы выбрали этот источник, так как он детально описывает работу датчиков уровней (LIT), расходомеров (FIT) и ПЛК Allen-Bradley в условиях киберфизических воздействий.

Техническое описание эксперимента

Вместо прямого управления моделью мы реализовали стенд для воспроизведения трафика. Данные из датасета подавались на вход системы мониторинга последовательно, имитируя поток событий в реальном времени.

Эксперимент проходил в два этапа:

1. Первичный анализ. Записи из датасета поступали в классическую IDS. В качестве базового метода использовался алгоритм изоляционного леса (Isolation Forest). Он помечал как «аномальные» 4 612 легитимных событий из-за естественной волатильности физических параметров.

2. Контекстуальная валидация. Те же 4 612 событий передавались на вход нашего алгоритма. Для каждой команды из лога система восстанавливала текущее состояние S_{phys} и рассчитывала прогноз S_{next} по матрицам динамики.

В ходе контекстуальной валидации 4 507 событий были классифицированы как физически допустимые. Это означает, что 97,7% статистических аномалий, выявленных базовой IDS, на самом деле являлись безопасными технологическими переходами. Оставшиеся 105 ложных срабатываний вызваны пограничными состояниями. В этих случаях расчётные параметры находились вблизи уставок систем противоаварийной защиты (ПАЗ), что объективно требует ручного анализа оператором SOC.

Коэффициент снижения шума составил:

$$FPRR = (4612 - 105) / 4612 \times 100\% \approx 97,7\%.$$

Технически результаты эксперимента доказывают, что большинство ложных тревог генерируется из-за несовершенства эвристических моделей. Классические IDS видят только математические отклонения, но не учитывают инерцию и физику оборудования.

Использование детерминированного фильтра Ω_{safe} позволяет отсеять этот информационный шум без потери чувствительности к целевым угрозам. Из 1 035 внедрённых в датасет скрытых LotL-атак алгоритм успешно выявил 1 021 воздействие. При этом сис-

тема отработала по двум независимым векторам. Порядка 60% угроз были заблокированы физическим фильтром (прогнозный вектор S_{next} явно нарушал пределы безопасности узлов водоочистки). Остальные 40% алгоритм отсекал через организационный фильтр S_{org} . Система зафиксировала легитимные по физике команды, которые отправлялись вне разрешённого окна расписания.

Оценка вычислительной задержки

В АСУ ТП время – это критический параметр. Если система защиты думает слишком долго, она задержит легитимную команду и сама спровоцирует аварию. Поэтому мы измерили задержку, которую вносит алгоритм.

Как отмечено в разделе 2, предиктивная модель *Apply* опирается на матричные вычисления. Мы замерили время расчёта функции для графа состояний, включающего 50 взаимосвязанных физических параметров (уровень типичного цеха).

Для оценки производительности мы замерили время работы алгоритма на графе логико-физических переходов. В качестве тестовой модели мы взяли сегмент сети, включающий 50 взаимосвязанных параметров:

- ◆ 24 датчика (уровни, давление, расход, температура);
- ◆ 16 исполнительных механизмов (состояние насосов, задвижек, частотных преобразователей);
- ◆ 10 критических уставок (границы ПАЗ и рабочие уставки).

Взаимосвязь параметров задаётся через разреженные матрицы динамики A и управления K . Например, изменение состояния насоса (параметр №30) через матрицу K мгновенно влияет на прогноз расхода (параметр №5) и давления (параметр №12). 50 переменных – это стандартный объём данных для одного технологического узла, например, линии фильтрации или насосной станции второго подъёма.

Технические условия замера:

Расчёты проводились на стандартном процессоре уровня Intel Core i5 (2.4 ГГц) с использованием библиотек линейной алгебры. Основное время тратится на два действия:

1. Перемножение матриц для получения вектора S_{next} .
2. Проверка 50 неравенств вида $g_i(S_{next}) \leq 0$ для поиска выхода за границы безопасности.

Среднее время обработки одного события составило 1.4 миллисекунды (мс). При пиковых скачках трафика максимальная задержка не превышала 2.1 мс.

Для сравнения: стандартный цикл опроса контроллеров по протоколу Modbus TCP занимает от 50 до 100 мс. Наш алгоритм отрабатывает на порядок быстрее. Он укладывается даже в жёсткие стандарты реального времени, где лимит передачи сообщения составляет 3-4 мс. Это доказывает, что даже при усложнении модели до уровня цеха (200–300 параметров), время отклика останется в пределах 5–10 мс, что полностью удовлетворяет требованиям реального времени.

Эти данные подтверждают, что отказ от тяжёлых моделей машинного обучения был обоснован. Алгоритм не создаёт узких мест в сети. Его можно ставить в разрыв трафика (in-line режим) для автоматической блокировки угроз, не опасаясь рассинхронизации технологического процесса.

Обоснование выбора критериев оценки

Для оценки эффективности методики были выбраны два ключевых критерия:

1. Коэффициент снижения шума (False Positive Reduction Rate, FPRR). Показывает, какую долю формальных аномалий алгоритм отсеял как легитимные технологические операции.
2. Полнота выявления MITRE-тактик. Способность обнаружить угрозу, не имея сигнатуры конкретного эксплойта.

Непротиворечивость критериев обеспечивается детерминированностью физических законов. В отличие от правил межсетевого экрана, которые могут конфликтовать друг с другом, физическое ограничение «объём бака не может быть больше 100%» абсолютно.

Полнота критериев достигается за счёт покрытия двух главных векторов: времени (организационный фильтр) и материи (физический фильтр). Если атака не нарушает регламент времени и не нарушает физику процесса, она, по определению, не наносит ущерба процессу в данный момент.

Обсуждение результатов

Полученные данные подтверждают гипотезу исследования о том, что физический контекст является эффективным фильтром для верификации киберугроз.

Подтверждение снижения ложных срабатываний

В ходе экспериментов классическая система обнаружения (IDS) реагировала на любые нестандартные команды управления, генерируя сотни предупреждений в сутки. Внедрение организационного фильтра (Сценарий 2) позволило автоматически классифицировать большинство этих событий как легитимные, так как они происходили в рамках открытых нарядов-допусков. Это решает проблему «усталости от тревог», описанную в [6], освобождая внимание оператора для реальных угроз.

Подтверждение выявления IOA без полных логов

Сценарий 2 показал, что для обнаружения атаки нам не потребовались логи аудита с самого ПЛК (которые часто отсутствуют или могут быть стёрты). Хватало факта передачи команды и знания физических ограничений объекта. Замысел (IOA) атаки – «Нарушение техпроцесса» – определялся через анализ её потенциального результата.

Это доказывает, что предложенная методика позволяет компенсировать техническую отсталость полевых устройств (отсутствие встроенных средств безопасности) за счёт внешней аналитики. Мы переносим доверие с уязвимого контроллера на неизменные законы физики.

Таким образом, использование цепочки «След → Физика → Замысел» позволяет с высокой точностью атрибутировать атаки типа LoTL, где злоумышленник использует штатные инструменты, но с деструктивными целями [1].

Заключение. В этой работе мы представили и протестировали алгоритм контекстуальной валидации поведенческих индикаторов (IOB), в основу которого легла идея «обратного обогащения» данных. Проведённое исследование привело к нескольким важным выводам:

1. Смена подхода к обнаружению. Мы показали, что для выявления скрытых атак недостаточно опираться на статические следы компрометации (IOC). Предложенный подход формализует переход от наблюдения за поведенческим паттерном (IOB) к подтверждению вредоносного замысла (IOA). В отличие от вероятностных моделей машинного обучения, алгоритм использует детерминированность физики процесса: каждое управляющее воздействие проверяется на вхождение в безопасное множество Ω_{safe} . Это даёт строгий бинарный критерий – действие либо физически допустимо, либо противоречит природе процесса и является атакой.

2. Экспериментальные результаты. Экспериментальная проверка на данных академического стенда SWaT подтвердила эффективность методики:

- ◆ Снижение «шума»: алгоритм позволил снизить коэффициент ложных срабатываний (FPRR) на 97,7%, автоматически отсеяв легитимные технологические операции.
- ◆ Выявление скрытых угроз: успешно детектирована большая часть инъекций типа LoTL (1021 из 1035 атак). При этом 60% угроз было заблокировано физическим фильтром, а 40% – организационным регламентом (S_{org}).
- ◆ Производительность: средняя вычислительная задержка составила 1,4 мс, что подтверждает применимость решения в системах реального времени.

3. Практическая значимость. Предложенный алгоритм решает проблему «усталости от тревог» у операторов промышленных SOC. Использование технологического контекста в качестве фильтра позволяет надёжно выявлять инциденты и компенсировать недостатки глубокого логирования на устаревшем полевом оборудовании.

Поставленная цель исследования достигнута: доказано, что связь IOC–IOB–IOA жёстко детерминирована спецификой отрасли, а контекст техпроцесса является надёжным критерием для верификации киберугроз.

В будущем планируется автоматизировать процесс создания правил валидации. Применение методов машинного обучения к историческим архивам данных АСУ ТП позволит системе самостоятельно профилировать нормальное физическое поведение объекта и динамически обновлять пороги безопасности без ручной настройки экспертов.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Ongun T., Stokes J., Bar Or J. Living-Off-The-Land Command Detection Using Active Learning // RAID. – 2021. – P. 442-455. – DOI: 10.48550/arXiv.2111.15039.
2. Ning R., Bu W., Yang J., Duan S. A Survey of Detection Methods Research on Living-Off-The-Land Techniques // 2023 IEEE International Conference on Sensors, Electronics and Computer Engineering (ICSECE). – 2023. – P. 159-164. – DOI: 10.1109/ICSECE58870.2023.10263445.
3. Kelley K. Indicators of Compromise (IoCs) and Their Role in Attack Defence (RFC 9424) // Internet Engineering Task Force. – 2023. – DOI: 10.17487/RFC9424.
4. Asiri M., Saxena N., Burnap P. Investigating Usable Indicators against Cyber-Attacks in Industrial Control Systems // USENIX Symposium on Usable Privacy and Security (SOUPS). – 2021.
5. Alahmadi B.A., Axon L., Martinovic I. 99% False Positives: A Qualitative Study of SOC Analysts' Perspectives on Security Alarms // 31st USENIX Security Symposium. – 2022. – P. 2783-2800.
6. Asiri M., Arunasalam A., Saxena N., Celik Z.B. Frontline responders: Rethinking indicators of compromise for industrial control system security // Computers & Security. – 2025. – Vol. 154. – P. 104421. – DOI: 10.1016/j.cose.2025.104421.
7. Атарская Е.А., Вульфин А.М., Узбекова Л.Я. Система обнаружения аномалий состояния киберфизических объектов в задаче обеспечения информационной безопасности // Вестник УГАТУ. – 2023. – Т. 27, № 2. – С. 98-107.
8. Lyu S., Wang K., Zhang L., Wang B. Process-Oriented heterogeneous graph learning in GNN-Based ICS anomalous pattern recognition // Pattern Recognition. – 2023. – Vol. 141. – Article 109661. – DOI: 10.1016/j.patcog.2023.109661.
9. Erba A., Taormina R., Galelli S. et al. No Need to Know Physics: Resilience of Process-based Model-free Anomaly Detection for Industrial Control Systems // ACM ACSAC. – 2020. – DOI: 10.1145/3427228.3427261.
10. Müller N., Ziras C., Heussen K. Assessment of Cyber-Physical Intrusion Detection and Classification for Industrial Control Systems // IEEE International Conference on Communications, Control, and Computing Technologies for Smart Grids (SmartGridComm). – 2022. – DOI: 10.1109/SmartGridComm52983.2022.9961010.
11. Yahya M., Sharaf N., Rrushi J. et al. Physics Reasoning for Intrusion Detection in Industrial Networks // 2nd IEEE International Conference on Trust, Privacy and Security in Intelligent Systems and Applications (TPS-ISA). – 2020. – DOI: 10.1109/TPS-ISA50397.2020.00043.
12. Aftabi N., Li D., Ramanan P. A Variational Autoencoder Framework for Robust, Physics-Informed Cyberattack Recognition in Industrial Cyber-Physical Systems // arXiv:2310.06948. – 2023. – DOI: 10.48550/arXiv.2310.06948.
13. Ike M., Phan K., Sadoski K. et al. SCAPHY: Detecting Modern ICS Attacks by Correlating Behaviors in SCADA and PHYsical // Proc. IEEE Symposium on Security and Privacy (SP). – 2023. – P. 362-379. – DOI: 10.1109/SP46215.2023.00067.
14. Basan E., Basan A., Nekrasov A., Fidge C., Ishchukova E., Basyuk A., Lesnikov A. Trusted Operation of Cyber-Physical Processes Based on Assessment of the System's State and Operating Mode // Sensors. – 2023. – Vol. 23, No. 4. – P. 1996. – DOI: 10.3390/s23041996.
15. Ahmed I., Obermeier S., Naedele M. et al. SCADA systems: Challenges for forensic investigators // Computer. – 2012. – Vol. 45, No. 12. – P. 44-51. – DOI: 10.1109/MC.2012.325.
16. Kabasele Ndonda G., Sadre R. Exploiting the Temporal Behavior of State Transitions for Intrusion Detection in ICS/SCADA // IEEE Access. – 2022. – Vol. 10. – P. 41535-41554. – DOI: 10.1109/ACCESS.2022.3168864.
17. Asiri M., Saxena N., Gjomemo R., Burnap P. Understanding Indicators of Compromise against Cyberattacks in Industrial Control Systems: A Security Perspective // ACM Transactions on Cyber-Physical Systems. – 2023. – Vol. 7, No. 2. – DOI: 10.1145/3565972.
18. Абрамов Е.С., Белов Н.Е. Классификация и оценка качества источников индикаторов компрометации // Системный синтез и прикладная синергетика: Сб. научных трудов XII Всерос. науч. конф. – Ростов-на-Дону; Таганрог: ЮФУ, 2024. – С. 392-395.

19. Туманов Д.А., Абрамов Е.С. Поиск и скоринг источников индикаторов компрометации по индустриям // Инженерный вестник Дона. – 2024. – № 8 (116).
20. Azzam M., Pasquale L., Provan G. et al. Forensic readiness of industrial control systems under stealthy attacks // *Computers & Security*. – 2023. – Vol. 125. – P. 103010. – DOI: 10.1016/j.cose.2022.103010.
21. Radvanovsky R., Brodsky J. SCADA/Control Systems Security. – Boca Raton: CRC Press, 2013. – 330 p. – DOI: 10.1201/b19545.
22. Котенко И.В., Хмыров С.С. Модель компрометации объектов критической информационной инфраструктуры // Актуальные проблемы инфотелекоммуникаций в науке и образовании (АПИНО 2022). – СПб., 2022. – Т. 1. – С. 614-619.
23. Knowles W., Prince D., Hutchison D. et al. A survey of cyber security management in industrial control systems // *International Journal of Critical Infrastructure Protection*. – 2015. – Vol. 9. – P. 52-80. – DOI: 10.1016/j.ijcip.2015.02.002.
24. Mathur A.P. SWaT: a water treatment testbed for research and training on ICS security // 2016 International Workshop on Cyber-physical Systems for Smart Water Networks (CySWater). – IEEE, 2016. – P. 31-36. – DOI: 10.1109/CySWater.2016.7469060.
25. Goh J. A dataset to support research in the design of secure water treatment systems // *Critical Information Infrastructures Security: 11th International Conference, CRITIS 2016*. – Springer, 2017. – P. 88-99. – DOI: 10.1007/978-3-319-71368-7_8.

REFERENCES

1. Ongun T., Stokes J., Bar Or J. Living-Off-The-Land Command Detection Using Active Learning, *RAID*, 2021, pp. 442-455. DOI: 10.48550/arXiv.2111.15039.
2. Ning R., Bu W., Yang J., Duan S. A Survey of Detection Methods Research on Living-Off-The-Land Techniques, 2023 *IEEE International Conference on Sensors, Electronics and Computer Engineering (ICSECE)*, 2023, pp. 159-164. DOI: 10.1109/ICSECE58870.2023.10263445.
3. Kelley K. Indicators of Compromise (IoCs) and Their Role in Attack Defence (RFC 9424), *Internet Engineering Task Force*, 2023. DOI: 10.17487/RFC9424.
4. Asiri M., Saxena N., Burnap P. Investigating Usable Indicators against Cyber-Attacks in Industrial Control Systems, *USENIX Symposium on Usable Privacy and Security (SOUPS)*, 2021.
5. Alahmadi B.A., Axon L., Martinovic I. 99% False Positives: A Qualitative Study of SOC Analysts' Perspectives on Security Alarms, 31st *USENIX Security Symposium*, 2022, pp. 2783-2800.
6. Asiri M., Arunasalam A., Saxena N., Celik Z.B. Frontline responders: Rethinking indicators of compromise for industrial control system security, *Computers & Security*, 2025, Vol. 154, pp. 104421. DOI: 10.1016/j.cose.2025.104421.
7. Atarskaya E.A., Vul'fin A.M., Uzbekova L.Ya. Sistema obnaruzheniya anomalii sostoyaniya kiberfizicheskikh ob'ektov v zadache obespecheniya informatsionnoy bezopasnosti [System for detecting anomalies in the state of cyber-physical objects in the task of ensuring information security], *Vestnik UGATU [Vestnik USATU]*, 2023, Vol. 27, No. 2, pp. 98-107.
8. Lyu S., Wang K., Zhang L., Wang B. Process-Oriented heterogeneous graph learning in GNN-Based ICS anomalous pattern recognition, *Pattern Recognition*, 2023, Vol. 141, Article 109661. DOI: 10.1016/j.patcog.2023.109661.
9. Erba A., Taormina R., Galelli S. et al. No Need to Know Physics: Resilience of Process-based Model-free Anomaly Detection for Industrial Control Systems, *ACM ACSAC*, 2020. DOI: 10.1145/3427228.3427261.
10. Müller N., Ziras C., Heussen K. Assessment of Cyber-Physical Intrusion Detection and Classification for Industrial Control Systems, *IEEE International Conference on Communications, Control, and Computing Technologies for Smart Grids (SmartGridComm)*, 2022. DOI: 10.1109/SmartGridComm52983.2022.9961010.
11. Yahya M., Sharaf N., Rrushi J. et al. Physics Reasoning for Intrusion Detection in Industrial Networks, 2nd *IEEE International Conference on Trust, Privacy and Security in Intelligent Systems and Applications (TPS-ISA)*, 2020. DOI: 10.1109/TPS-ISA50397.2020.00043.
12. Aftabi N., Li D., Ramanan P. A Variational Autoencoder Framework for Robust, Physics-Informed Cyberattack Recognition in Industrial Cyber-Physical Systems, *arXiv:2310.06948*, 2023. DOI: 10.48550/arXiv.2310.06948.
13. Ike M., Phan K., Sadoski K. et al. SCAPHY: Detecting Modern ICS Attacks by Correlating Behaviors in SCADA and PHYsical, *Proc. IEEE Symposium on Security and Privacy (SP)*, 2023, pp. 362-379. DOI: 10.1109/SP46215.2023.00067.
14. Basan E., Basan A., Nekrasov A., Fidge C., Ishchukova E., Basyuk A., Lesnikov A. Trusted Operation of Cyber-Physical Processes Based on Assessment of the System's State and Operating Mode, *Sensors*, 2023, Vol. 23, No. 4, pp. 1996. DOI: 10.3390/s23041996.

15. Ahmed I., Obermeier S., Naedele M. et al. SCADA systems: Challenges for forensic investigators, *Computer*, 2012, Vol. 45, No. 12, pp. 44-51. DOI: 10.1109/MC.2012.325.
16. Kabasele Ndonda G., Sadre R. Exploiting the Temporal Behavior of State Transitions for Intrusion Detection in ICS/SCADA, *IEEE Access*, 2022, Vol. 10, pp. 41535-41554. DOI: 10.1109/ACCESS.2022.3168864.
17. Asiri M., Saxena N., Gjomemo R., Burnap P. Understanding Indicators of Compromise against Cyberattacks in Industrial Control Systems: A Security Perspective, *ACM Transactions on Cyber-Physical Systems*, 2023, Vol. 7, No. 2. DOI: 10.1145/3565972.
18. Abramov E.S., Belov N.E. Klassifikatsiya i otsenka kachestva istochnikov indikatorov komprometatsii [Classification and quality assessment of sources of indicators of compromise], *Sistemnyy sintez i prikladnaya sinergetika: Sb. nauchnykh trudov XII Vseros. nauch. konf. [System Synthesis and Applied Synergetics: Proceedings of the XII All-Russian Scientific Conference]*. Rostov-on-Don; Taganrog: YuFU, 2024, pp. 392-395.
19. Tumanov D.A., Abramov E.S. Poisk i skoring istochnikov indikatorov komprometatsii po industriyam [Search and scoring of sources of indicators of compromise by industry], *Inzhenernyy vestnik Dona [Engineering Journal of Don]*, 2024, No. 8 (116).
20. Azzam M., Pasquale L., Provan G. et al. Forensic readiness of industrial control systems under stealthy attacks, *Computers & Security*, 2023, Vol. 125, pp. 103010. DOI: 10.1016/j.cose.2022.103010.
21. Radvanovsky R., Brodsky J. SCADA/Control Systems Security. Boca Raton: CRC Press, 2013, 330 p. – DOI: 10.1201/b19545.
22. Kotenko I.V., Khmyrov S.S. Model' komprometatsii ob"ektov kriticheskoy informatsionnoy infrastruktury [Model of compromise of critical information infrastructure objects], *Aktual'nye problemy infotelekkommunikatsiy v nauke i obrazovanii (APINO 2022) [Actual Problems of Infotelecommunications in Science and Education (APINO 2022)]*. Saint Petersburg., 2022, Vol. 1, pp. 614-619.
23. Knowles W., Prince D., Hutchison D. et al. A survey of cyber security management in industrial control systems, *International Journal of Critical Infrastructure Protection*, 2015, Vol. 9, pp. 52-80. DOI: 10.1016/j.ijcip.2015.02.002.
24. Mathur A.P. SWaT: a water treatment testbed for research and training on ICS security, *2016 International Workshop on Cyber-physical Systems for Smart Water Networks (CySWater)*. IEEE, 2016, pp. 31-36. DOI: 10.1109/CySWater.2016.7469060.
25. Goh J. A dataset to support research in the design of secure water treatment systems, *Critical Information Infrastructures Security: 11th International Conference, CRITIS 2016*. Springer, 2017, pp. 88-99. DOI: 10.1007/978-3-319-71368-7_8.

Абрамов Евгений Сергеевич – Южный федеральный университет; e-mail: abramoves@sfedu.ru; г. Таганрог, Россия; тел.: +78634371905; к.т.н.; доцент.

Белов Николай Евгеньевич – Южный федеральный университет; e-mail: info@ictis.sfedu.ru; г. Таганрог, Россия; тел.: +78634360450; аспирант.

Веселов Геннадий Евгеньевич – Южный федеральный университет; e-mail: gev@sfedu.ru; г. Таганрог, Россия; тел.: +78634360450; директор ИКТИБ; д.т.н.; профессор.

Abramov Eugene Sergeevich – Southern Federal University; e-mail: abramoves@sfedu.ru; Taganrog, Russia; phone: +78634371905; cand. of eng. sc.; associate professor.

Belov Nikolay Evgen'evich – Southern Federal University; e-mail: info@ictis.sfedu.ru; Taganrog, Russia; phone: +78634360450; graduate student.

Veselov Gennadiy Evgen'evich – Southern Federal University; e-mail: gev@sfedu.ru; Taganrog, Russia; phone: +78634360450; director of ICTIS; dr. of eng. sc.; professor.

А.А. Александров, М.А. Бутакова**ПОВЫШЕНИЕ КАЧЕСТВА РАЗНОМОДАЛЬНЫХ ДАННЫХ
НА ОСНОВЕ ГИБРИДИЗАЦИИ МЕТОДОВ ПРЕДВАРИТЕЛЬНОЙ ОБРАБОТКИ
И МУЛЬТИСЕНСОРНОГО СЛИЯНИЯ**

Исследуются методы предварительной обработки данных и слияния разномодальных данных. Собираемые с мультисенсорных устройств данные содержат шумы, аномалии и включают отказы датчиков. Использование неочищенных данных неизбежно приводит к получению ложных закономерностей при использовании моделей машинного обучения, в которых эти данные используются. Для улучшения качества данных необходимо применение специальных методов предварительной их обработки и слияния. Целью работы является создание гибридного конвейера предварительной обработки и мультисенсорного слияния данных, функционирующего на уровне периферийных вычислений. Для её достижения решены задачи удаления из данных последствий технических сбоев и шумов, временной синхронизации, масштабирования признаков и объединения дублирующих источников с учетом их динамического веса. Применены методы: межквартильного размаха для выявления выбросов, линейная интерполяция для их замены, фильтр высоких частот для подавления шумов и взвешенное мультисенсорное слияние с динамическим расчётом весов. Кроме того, представлены методы обработки визуальных, сенсорных и акустических данных в рамках гибридного конвейера, допускающие адаптацию для реализации на уровне периферийных вычислений. Экспериментальная проверка разработанного конвейера проведена на смоделированном наборе климатических данных с искусственно внесёнными зашумлениями и аномалиями. Предложенный подход продемонстрировал высокую эффективность восстановления данных. Для температуры воздуха коэффициент детерминации (R^2) увеличился с 0,9807 до 0,9964 при незначительном изменении средней абсолютной процентной ошибки (MAPE) с 10,11% до 10,48%. Для относительной влажности метрика R^2 выросла с 0,9248 до 0,9675, а ошибка MAPE снизилась с 2,87% до 2,57%. Наиболее значительное улучшение достигнуто для атмосферного давления. Метрика R^2 с отрицательного значения -0,124 выросла до 0,974, а ошибка MAPE сократилась с 0,12% до 0,06 %. Полученные результаты подтверждают, что предложенный конвейер снижает погрешность и повышает качество мультисенсорных данных, минимизирует объём передаваемых данных. Практическая ценность заключается в создании непрерывного цикла локальной фильтрации и слияния данных без накопления погрешностей.

Предварительная обработка данных; слияние данных; мультисенсорные системы; интернет вещей; периферийные вычисления; машинное обучение; качество данных; конвейер обработки данных.

A.A. Alexandrov, M.A. Butakova**IMPROVING THE QUALITY OF MULTI-MODAL DATA BASED
ON THE HYBRIDIZATION OF PREPROCESSING AND MULTISENSORY
FUSION METHODS**

The article explores the methods of data preprocessing and merging of multi-modal data. The data collected from multi-sensor devices contains noise, anomalies, and includes sensor failures. The use of raw data inevitably leads to false patterns when using machine learning models in which this data is used. To improve the quality of data, it is necessary to use special methods for preliminary processing and merging. The aim of the work is to create a hybrid preprocessing and multisensory data fusion controller operating at the peripheral computing level. To achieve this, the tasks of removing the consequences of technical failures and noise from the data, time synchronization, scaling of features and combining duplicate sources, considering their dynamic weight, have been solved. The following methods were applied: inter-quartile scale to detect outliers, linear interpolation to replace them, high-pass filter to suppress noise, and weighed multisensory fusion with dynamic weight calculation. In addition, methods for processing visual, sensory, and acoustic data within a hybrid pipeline are presented that can be adapted for implementation at the peripheral computing level. The experimental verification of the developed pipeline was carried out on a simulated set of climatic data with artificially introduced noises and anomalies. The proposed approach has demonstrated high efficiency of data recovery. For air temperature, the coefficient of determination (R^2) increased from 0.9807 to 0.9964 with a slight change in the mean absolute percentage error (MAPE) from 10.11% to 10.48%. For relative humidity, the R^2 metric increased from

0.9248 to 0.9675, and the MAPE error decreased from 2.87% to 2.57%. The most significant improvement has been achieved for atmospheric pressure. The R2 metric increased from a negative value of -0.124 to 0.974, and the MAPE error decreased from 0.12% to 0.06%. The results confirm that the proposed pipeline reduces the error and improves the quality of multisensory data, minimizes the amount of data being transferred. The practical value lies in creating a continuous cycle of local filtering and merging of data without accumulation of errors.

Data preprocessing; data fusion; multisensory systems; internet of things; peripheral computing; machine learning; data quality; data processing pipeline.

Введение. Достоверность любого научного исследования зависит от правильно подобранных и подготовленных данных. Полученные результаты должны быть повторяемыми и воспроизводимыми в рамках конкретного класса объектов при одинаковых условиях эксперимента. Обоснованное научное исследование предполагает проведение экспериментов строго в рамках научных стандартов и правил, действующих в соответствующей области знаний [1]. Ключевым требованием при этом является использование актуальных данных из различных разномодальных источников, которые позволяют всесторонне рассмотреть изучаемый объект. В эпоху больших данных исследователи часто сталкиваются с заблуждением доступности и высокого качества данных. Большой объем данных не гарантирует их качество.

Особую остроту эта проблема приобретает в системах мониторинга [2], промышленного интернета вещей (IoT) [3], в таких областях как медицина и биология [4], робототехника [5], экология [6], где мультисенсорные устройства собирают гетерогенные данные (температура, влажность, давление, акустические и визуальные сигналы) непосредственно в полевых условиях. Низкое качество первичных измерений ведёт к ошибочным прогнозам и неверным управляющим решениям. Избежать этого помогает обработка информации прямо на самих устройствах.

Также, скрытые системные ошибки часто искажают результаты, а аналитик об этом даже не подозревает. Если обучить классические модели машинного обучения (МО) на таких «грязных» массивах, они найдут ложные закономерности. Проявляется принцип, известный в зарубежной литературе как Garbage In, Garbage Out (GIGO, «мусор на входе – мусор на выходе») [7, 8]. Чтобы этого избежать, данные необходимо тщательно проверять и обрабатывать еще до запуска моделей.

Предварительная обработка состоит из нескольких обязательных шагов, таких как восстановление отсутствующих значений, удаление аномалий, нормировка и кодирование признаков. Комплексная реализация сложных научных исследований зачастую не основывается только на одном источнике или методе проведения исследований. Нужны методы применения гибридного подхода, в том числе подхода слияния данных, известного в научной литературе как Data Fusion – это методы, позволяющие объединить различными образом формируемые массивы данных в едином информационном пространстве [9–11]. Слияние может быть проведено на уровнях данных, признаков и решений. Профессиональное объединение данных позволяет получать более качественные данные и повысить точность исследования. Поэтому предварительная очистка обязательна. Если её пропустить, ошибки из разных источников наложатся друг на друга. В итоге погрешность вырастет лавинообразно.

Текущие методы очистки имеют два подхода: они либо собирают и обрабатывают весь массив данных на центральном сервере, либо разбивают фильтрацию на отдельные несвязанные шаги. В работах [12–14] показана эффективность отдельных методов, но не предложено целостного конвейера, способного одновременно удалять аномалии, компенсировать отказы датчиков и выполнять мультисенсорное слияние с пересчётом весов в реальном времени на периферийном оборудовании. Разработчики все еще предпочитают передавать «сырые» данные на единый сервер и проводить предварительную обработку на нем, не используя адаптивный конвейер, объединяющий фильтрацию, восстановление пропусков и взвешенное объединение резервирующих датчиков без накопления ошибок на самих периферийных устройствах.

Микроконтроллеры и технологии интернета вещей автоматизируют научные эксперименты, снижают трудозатраты и позволяют собирать большие объемы данных в реальном времени при проведении экспериментов *in situ* [15, 16]. Однако, каждый датчик подвержен шумам и погрешностям, а нарушения в работе электрических схем могут приводить к появлению аномальных значений. Поэтому непрерывный поток данных необходимо фильтровать мгновенно. Первичную обработку сигналов переносят на уровень периферийных вычислений (Edge Computing), чтобы отсекал искажения до отправки на сервер [17].

В данной статье предлагается гибридный конвейер предварительной обработки и мультисенсорного слияния, ориентированный на выполнение в среде периферийных вычислений для использования в научных исследованиях. Его новизна состоит в совместном применении методов фильтрации, обнаружения и удаления аномалий, и взвешенного мультисенсорного слияния с динамическим расчётом весовых коэффициентов. В отличие от существующих решений, конвейер не требует передачи сырых данных на сервер, а реализует полный цикл локальной очистки, временной синхронизации, масштабирования признаков и объединения дублирующих источников в реальном времени. Кроме того, в его архитектуре предусмотрена адаптация к разнотипным модальностям – визуальным, сенсорным и акустическим данным. Далее в статье описаны этапы разработанного конвейера, приведены используемые методы обработки, представлены результаты экспериментов на смоделированном климатическом наборе данных и выполнено сравнение с исходными показателями.

Методика исследования. Для повышения качества мультисенсорных данных, собираемых в системах интернета вещей, используют три подхода: централизованная облачная обработка, гибридная архитектура и полностью локальная на периферии. Облачный сервис даёт очень большие вычислительные ресурсы. Однако объёмы передачи данных нагружают сеть, вырастают задержки и появляются уязвимости для доступа к исходным данным, что критично для ресурсоограниченных автономных устройств. Альтернативой служат методы локальной обработки данных.

В работе выбран экспериментальный подход с моделированием на синтетическом наборе данных. Это позволило строго контролировать характер и интенсивность вносимых искажений, что невозможно сделать на реальных «полевых» данных без эталонных значений. Стратегия заключается в создании гибридного конвейера предварительной обработки и мультисенсорного слияния прямо на Edge-устройствах, и его верификации на специально подготовленных данных с известным истинным сигналом. Преимуществом подхода является возможность количественно оценить вклад каждого этапа конвейера в итоговое повышение точности и сравнить полученные показатели с исходными «сырыми» данными.

Процесс разработки и валидации конвейера включает следующие шаги:

1. Подбор методов и алгоритмов гибридного конвейера, включающего стадии фильтрации, синхронизации, удаления аномалий, масштабирования признаков и слияния модальностей. Система должна обрабатывать разные типы данных – визуальные, сенсорные и акустические.
2. Генерация синтетического климатического набора данных (температура воздуха, относительная влажность, атмосферное давление). Для имитации условий реальных сенсорных сетей искусственно добавлены в данные типичные дефекты: шум, выбросы и отказы датчиков.
3. Проверка работы конвейера в условиях смоделированных сбоев и помех.
4. Оценка эффективности конвейера с помощью регрессионных метрик путём сравнения эталонных (истинных) значений и исходными «сырыми» показателями до обработки.

Постановка задачи. Обычно исследователи работают с готовыми архивными файлами, которые уже прошли очистку. Но если обучать или тестировать модели МО на «сырых» данных, система примет шум и случайные всплески за реальные закономерности.

Справиться с этой задачей помогает локальная обработка прямо на Edge-устройствах. Такой конвейер убирает искажения «на месте», разгружает сеть и отправляет на сервер чистый, компактный массив данных.

Конвейер обработки разномодальных данных построен строго последовательно. Сначала отсекаются явные технические сбои, затем убираются шумы на каждом датчике, и синхронизируются данные по времени. После этого масштабируются признаки, и устраняется избыточность с помощью объединения дублирующих источников. Такая последовательность действий не позволяет ошибкам накапливаться, так как каждый этап жестко зависит от предыдущего. Конвейер сразу же исключает аномалии и объединяет дублирующие потоки на уровне «сырых» данных. Такая интеграция работает только для однородных датчиков. Вместо нескольких параллельных сигналов резервных датчиков поступает единственный эталонный сигнал, прошедший полную очистку, нормализацию и подготовку к обработке с использованием МО.

Ключевой особенностью конвейера является метод взвешенного мультисенсорного слияния с динамическим расчётом весов. Обычное усреднение сохраняет ошибки всех каналов, поэтому показаниям присваивают весовые коэффициенты. Если датчик выдает аномальное отклонение, конвейер мгновенно обнуляет его вес. Опираясь на избыточность, алгоритм вычисляет наиболее вероятное истинное значение параметра. Это превращает несколько сомнительных сигналов в одно надежное измерение, что резко повышает точность и отказоустойчивость системы.

Таким образом, разработанный инструментариий закладывает основу для создания унифицированных систем интеллектуального анализа гетерогенных данных. Конвейер универсален, но логика фильтрации, нормализации и комплексирования зависит от физической природы сигналов. Поэтому следующей задачей является разработка методов, применяемых для разномодальных данных – для обработки визуальных, сенсорных и акустических сигналов. Также необходима экспериментальная проверка предлагаемых подходов, которая выполнена и представлена в работе на основе на примере обработки климатических данных.

Методы обработки разномодальных данных. Для количественной оценки степени размытия и зашумленности видеоданных или последовательностей изображений широко применяется фильтр Лапласа. Фильтр использует вторую производную для обнаружения краев на изображении:

$$F(s) = \int_0^{\infty} f(t) e^{-st} dt, \quad (1)$$

где s – комплексное число, t – вещественное число, t' – первая производная функции $f(t)$.

В задачах дискретной обработки визуальных данных преобразование Лапласа работает с двумерными массивами значений яркости пикселей и позволяет количественно оценить градиент яркости в локальной окрестности каждой точки. Технически этот расчет выполняется путем свертки исходного кадра с маской (ядром 3×3) фильтр, коэффициенты которого аппроксимируют вторую производную функции яркости:

$$\begin{bmatrix} 0 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 0 \end{bmatrix}.$$

Вычисление дисперсии, полученной после применения фильтра Лапласа, позволяет оценить уровень шума и степень размытости изображения. Для вычисления дисперсии $D[I_f]$ отфильтрованного изображения I_f размером $n \times m$ сначала вычислим математическое ожидание $E[I_f]$ по формуле:

$$E[I_f] = \frac{1}{n \cdot m} \sum_{i=1}^n \sum_{j=1}^m I_f(i, j). \quad (2)$$

Итоговое значение дисперсии вычисляется согласно выражению:

$$D[I_f] = E\left[\left(I_f - E[I_f]\right)^2\right] = \frac{1}{n \cdot m} \sum_{i=1}^n \sum_{j=1}^m \left(I_f(i, j) - E[I_f]\right)^2. \quad (3)$$

Расчёты по формулам (2) и (3) дают оценку дисперсии. Чем выше дисперсия, тем выше уровень шума в отфильтрованном изображении.

Для удаления шума и сохранения структуры изображения часто применяется медианный фильтр. Медианная фильтрация хорошо удаляет импульсный шум и работает с однородными областями. Цифровое изображение представляет собой двумерный массив чисел $\{x_{i,j}\}$, где (i, j) изменяются по некоторому подмножеству Z^2 . Двумерный медианный фильтр с окном A и размером $n \times m$ для изображения $\{x_{i,j}, (i, j) \in Z^2\}$ определяется как

$$y_{i,j} = \underset{A}{med} x_{i,j} \equiv \underset{A}{med} [x_{i+r, j+s}; (r, s) \in A], (i, j) \in Z^2,$$

где y – отфильтрованное изображение, med – функция, находящая медиану, (r, s) – индексы смещения координаты внутри скользящего окна.

Данный метод обработки изображений демонстрирует наибольшую эффективность в условиях, когда шум проявляется в виде импульсных помех и представлен ограниченным набором пиковых значений на фоне нулевых значений. В результате применения медианного фильтра показатели яркости в наклонных участках и резких переходах изображения остаются неизменными. Метод показывает ограниченную эффективность при фильтрации флуктуационного шума. Для минимизации ухудшения четкости визуального отображения возможно применение адаптивной медианной фильтрации, которая позволяет лучше справляться с различными типами шумов, сохраняя при этом четкость контуров. Размер скользящего окна (маски) изменяется динамически и зависит от яркости соседних пикселей. Примем пороговый коэффициент отклонения яркости $S_{порог} \in [0, 1]$, а величину отклонения яркости соседних пикселей – $A(r, n, m)$, попавших в окно размером $n \times m$. Тогда, относительно яркости центрального отсчета $A(r)$ запишется в виде:

$$S_{nm}(r) = \left| \frac{A(r, n, m) - A(r)}{A(r)} \right|,$$

где r – интенсивность центрального пикселя. Размер маски увеличиваем до тех пор, пока выполняется условие: $\max[S_{nm}(r)] < S_{порог}$.

Если медианный фильтр не дает достаточных результатов, переходят к фильтрации изображений на основе вейвлет-преобразования $\psi(t)$. Его используют для разложения изображения на составные части с разными частотами:

$$\psi_{a,b}(t) = \frac{1}{\sqrt{|a|}} \psi\left(\frac{t-b}{a}\right), \quad (4)$$

где t – время (или пространственная координата), a – масштабный параметр, b – параметр смещения.

В отличие от преобразования Фурье, преобразование на основе (4) обеспечивает возможность локального анализа (временного или пространственного) сигнала. С помощью преобразования (4) также очень эффективно удаляются мелкие детали, не содержащие в себе никакой полезной информации, которые проявляются в изображении в виде шума. Вейвлет-преобразование изображения $f(x, y)$ производится следующим образом:

$$W_f(a, b) = \iint f(x, y) \psi_{a,b}(x, y) dx dy, \quad (5)$$

где $W_f(a, b)$ – коэффициенты вейвлет-преобразования при масштабе a и смещении b .

Процесс фильтрации изображений осуществляется посредством дискретного вейвлет-преобразования. Алгоритм раскладывает изображение на низкочастотную и высокочастотную составляющие (аппроксимирующую и детализирующую матрицы). Анализ полученных коэффициентов вейвлет-преобразования (5) позволяет провести очистку изображения от шумов. Процесс уменьшения избыточных данных реализуется через пороговую фильтрацию, где решающим правилом выступает величина T :

$$W_f(a, b) = \begin{cases} W_f(a, b), & \text{если } |W_f(a, b)| > T; \\ 0, & \text{если } |W_f(a, b)| \leq T. \end{cases}$$

После фильтрации изображение восстанавливается из отфильтрованных коэффициентов с помощью обратного вейвлет-преобразования:

$$f(x, y) = \iint W_f(a, b) \psi_{a,b}(x, y) da db.$$

Альтернативный подход базируется на методах нелинейной фильтрации. Алгоритмы Калмана и Винера используют априорную информацию о сигнале. Это помогает минимизировать артефакты при обработке. В основе фильтра Калмана лежит рекурсивный подход. Также метод опирается на модель состояний. Благодаря этому удастся эффективно обрабатывать динамические последовательности. За счет экстраполяции оценок фильтр Калмана проявляет устойчивость к аддитивным шумам даже в условиях неопределенности. Процессы прогнозирования и коррекции ковариационной матрицы описываются классическими выражениями (6) и (7) соответственно:

$$\hat{x}_k^- = \mathbf{F}\hat{x}_{k-1} + \mathbf{B}u_{k-1}, \quad (6)$$

$$\mathbf{P}_k^- = \mathbf{F}\mathbf{P}_{k-1}\mathbf{F}^T + \mathbf{Q}. \quad (7)$$

Переменная в (6) $\hat{x}_k^-$ отвечает за прогноз состояния. Априорный характер этой оценки традиционно отмечают индексом «минус». Остальные параметры определяются следующим образом: $\mathbf{F}$ – матрица перехода между состояниями; $\hat{x}_{k-1}$ – состояние системы в предыдущий момент времени; $\mathbf{B}$ – матрица управляющего воздействия; u_{k-1} – управляющее воздействие в предыдущий момент времени; $\mathbf{P}$ – ковариационная матрица ошибки модели; $\mathbf{Q}$ – ковариационная матрица шума процесса.

Стандартный алгоритм фильтрации Калмана работает по двухэтапной схеме. Она последовательно совмещает шаги экстраполяции и апостериорной коррекции. Для точной проверки прогноза с учетом «свежих» данных рассчитывается оптимальный матричный коэффициент усиления Калмана $\mathbf{K}_k$. Его вычисляют в соответствии с выражением

$$\mathbf{K}_k = \mathbf{P}_k^- \mathbf{H}^T (\mathbf{H} \mathbf{P}_k^- \mathbf{H}^T + \mathbf{R})^{-1},$$

где $\mathbf{H}$ – матрица измерений, отображающая соотношение измерений и состояний, $\mathbf{R}$ – ковариационная матрица шума измерений.

Последующая оптимизация текущей оценки на основе векторов реальных измерений z_k осуществляется путем математического развертывания расчетной формулы

$$\hat{x}_k = \hat{x}_k^- + \mathbf{K}_k (z_k - \mathbf{H}\hat{x}_k^-),$$

где $\hat{x}_k$ – оценка состояния системы с учетом измерения, z_k – измерение в текущий момент времени.

Завершающий пересчет и обновление матрицы плотности ошибок модели выполняется в строгом соответствии со структурной формулой

$$\mathbf{P}_k = (\mathbf{I} - \mathbf{K}_k \mathbf{H}) \mathbf{P}_k^-,$$

где $\mathbf{I}$ – единичная матрица соответствующей размерности.

В случае работы с RGB-изображениями данные операции проводятся локально для каждой матрицы интенсивностей. Математический аппарат такой коррекции, направленный на устранение артефактов освещения и расфокусировки подробно описан в научной литературе [18].

Контрастность монохромного изображения можно определить согласно формуле

$$C = z_{\max} - z_{\min},$$

где $z_{\max}$ – максимальная яркость пикселей, $z_{\min}$ – минимальная яркость пикселей в изображении. Относительная контрастность монохромного изображения вычисляется по формуле

$$C_i = \frac{N}{z_m},$$

где N – яркость цвета, $z_m = 255$ – максимально возможная яркость.

Если динамический диапазон яркости исходного изображения не достигает максимально допустимого значения z_m , то усиление контрастности может быть достигнуто за счет линейного растяжения гистограммы яркостей по формуле

$$z' = \frac{z - z_{\min}}{C_i},$$

где z , $z_{\min}$ – яркость пикселя до и после преобразования.

Это преобразование заключается в перераспределении значений интенсивности пикселей таким образом, чтобы минимальное значение яркости отображалось в нулевое, а максимальное – в z_m , что приводит к расширению диапазона яркости и, как следствие, к улучшению визуального восприятия изображения. Если диапазон яркостных характеристик изображения приближен к максимальным значениям, то повышение различимости смежных фрагментов может быть достигнуто посредством применения нелинейного преобразования шкалы яркости с использованием методов гамма-коррекции или соляризации. Метод гамма-коррекции яркости и контраста реализуется в соответствии с формулой

$$z' = z_m \left(\frac{z}{z_m} \right)^n,$$

где $\frac{z}{z_m}$ – относительная яркость.

Для обработки изображения с высокой яркостью рекомендуется задавать n в пределах $\{2 \dots 10\}$. Если же яркость фрагментов недостаточна, параметр n следует выбирать в интервале $\{0,1 \dots 0,5\}$.

Применение метода соляризации позволяет выделить границы смежных участков за счёт повышения контрастности в соответствии с формулой:

$$z' = 4 \frac{z}{z_m} (z_m - z).$$

Изображения нормализуют, чтобы модели учились быстрее и работали лучше. Чаще всего используют минимаксную нормализацию. Этот метод переводит яркость пикселей в диапазон от 0 до 1. Границы данных фиксируются жестко: минимум становится нулем, а максимум единицей. Нормализация реализуется в соответствии формулой

$$z' = \frac{z - z_{\min}}{z_{\max} - z_{\min}}.$$

Выходные данные измерительных приборов строятся по принципу «параметр-значение». По сути, они представляют собой временные ряды. Данные с приборов часто подвержены воздействию шумов. Их вызывают как внешние факторы, так и внутренние

свойства самих датчиков [19, 20]. Под шумом понимаются случайные отклонения или искажения сигнала, которые маскируют полезный сигнал. Фильтрация – один из методов борьбы с шумом. Аналоговые или цифровые фильтры позволяют избавиться от шумов в определенных частотных диапазонах. Фильтр низких частот (ФНЧ) пропускает низкие частоты и подавляет высокие. Одним из применяемых ФНЧ является Фильтр Баттерворта, который имеет максимально плоскую амплитудную характеристику в полосе пропускания. Фильтр Баттерворта определяется своим порядком n и частотой среза f_c . Передаточная функция $H(s)$ в аналоговой области задается согласно формуле:

$$H(s) = \frac{1}{\sqrt{1 + \left(\frac{s}{\omega_c}\right)^{2N}}},$$

где s – комплексная частота, $\omega_c = 2\pi f_c$ – частота среза, N – порядок фильтра.

Для перехода в цифровую область используется билинейное преобразование по формуле $s = \frac{2}{T} \cdot \frac{1 - z^{-1}}{1 + z^{-1}}$, где T – период дискретизации, z – оператор сдвига.

Фильтр высоких частот (ФВЧ) пропускает высокие частоты и подавляет низкие. Фильтр Баттерворта может быть реализован как ФНЧ, так как и ФВЧ. Передаточной функцией $H(s)$ в аналоговой области является величина, описываемая формулой:

$$H(s) = \frac{\left(\frac{s}{\omega_c}\right)^N}{\sqrt{1 + \left(\frac{s}{\omega_c}\right)^{2N}}}.$$

Полосовой фильтр пропускает сигналы только в определенном диапазоне частот и подавляет сигналы за пределами этого диапазона. Фильтр может быть реализован каскадным соединением на основе ФНЧ и ФВЧ согласно формуле:

$$H(s) = \frac{1}{\sqrt{1 + \left(\frac{s}{\omega_0}\right)^2}} \cdot \frac{\frac{s}{\omega_1}}{\sqrt{1 + \left(\frac{s}{\omega_1}\right)^2}},$$

где ω_0 и ω_1 – частоты среза ФНЧ и ФВЧ соответственно.

Данные временных рядов могут содержать не только шумы, но также и искажения или выбросы, которые представляют собой аномалии, отклоняющиеся от ожидаемого паттерна данных. Аномалии могут быть вызваны ошибками измерений, сбоем в работе оборудования или внешними воздействиями, не имеющими ничего общего с исследуемым процессом. Аномалии во временных рядах могут искажать результаты анализа и прогнозирования, поэтому их обнаружение и устранение являются критически важным этапом.

Применение метода межквартильного размаха (Interquartile Range или IQR) оправдано в условиях, когда специфика формы распределения не является определяющим фактором. Сам показатель IQR охватывает центральную часть выборки. В этот диапазон попадают 50% всех значений. Расчет проводится по формуле $IQR = Q3 - Q1$, где $Q1$ – первый квартиль, ниже которого находится 25% данных, $Q3$ – третий квартиль, ниже которого находится 75% данных. Определение нижней B_l и верхней B_h границ для выбросов осуществляется на основе формул: $B_h = Q1 - 1.5 \times IQR$; $B_l = Q3 + 1.5 \times IQR$.

Найденные аномалии необходимо классифицировать. Это может быть сбой оборудования, случайный выброс или новый тренд. Далее выбираем метод очистки: сглаживаем значения, заменяем их логичными аналогами либо подключаем прогнозные модели.

Чтобы уменьшить влияние случайных выбросов, данные сглаживают. Обычно для этого используют скользящее среднее, медианный фильтр или фильтр Калмана. При работе со временными рядами $y = \{y_1, y_2, \dots, y_n\}$ значение $\tilde{y}_k$ после медианной фильтрации с заданным окном $2m + 1$ вычисляется по формуле

$$\tilde{y}_k = \text{median}(y_{k-m}, y_{k-m+1}, \dots, y_k, \dots, y_{k+m-1}, y_{k+m}),$$

где y_n – элементы вариационного ряда ранжированные в порядке возрастания значений.

Восстановить пропуски можно с помощью линейной или сплайновой интерполяции. Также подойдут прогностические модели, например, ARIMA или LSTM (Long Short-Term Memory). Они рассчитывают нужные значения по соседним точкам или прошлым данным.

Использование нескольких датчиков компенсирует недостатки одиночных приборов [21–24]. Сенсоры дублируют друг друга и при поломке одного из них мониторинг продолжится без потери данных. Комбинация разных устройств решает две задачи – повышает точность измерений и отсеивает информационный шум.

Для количественной оценки согласованности дублирующих потоков информации применяется коэффициент корреляции Пирсона. Он позволяет установить степень их линейной зависимости [25]. Пусть X и Y будут случайными величинами – наблюдаемыми двумерными точками данных. Тогда коэффициент корреляции Пирсона r вычисляется по формуле:

$$r = \frac{\sum_{i=1}^k (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^k (x_i - \bar{x})^2 \sum_{i=1}^k (y_i - \bar{y})^2}},$$

где $\bar{x} = \frac{\sum_{i=1}^k x_i}{k}$, $\bar{y} = \frac{\sum_{i=1}^k y_i}{k}$ – выборочные средние значения.

Поскольку для переменных X и Y стандартные отклонения являются константными. В подобной нотации данная математическая структура традиционно идентифицируется в научной литературе как классический коэффициент корреляции смешанных моментов Пирсона, либо фигурирует под более лаконичным термином «коэффициент корреляции»:

$$r = \frac{\text{cov}(X, Y)}{\sigma_x \sigma_y},$$

где σ_x , σ_y – среднеквадратические (стандартные) отклонения соответствующих выборочных совокупностей X и Y .

При сильной корреляции, когда значение коэффициента стремится к единице, исследуемые сигналы становятся практически идентичными. Так мы можем напрямую объединять временные ряды, сохраняя при этом практически весь объем полезной информации. В промежуточном диапазоне (от 0,3 до 0,7), то есть при наличии заметной, но не исчерпывающей связи, решение о слиянии данных принимается уже на основе детального сравнительного анализа измерительных характеристик. Исследователи смотрят на точность и разрешение приборов, а также на физическую специфику самих сигналов. В таком случае используют прецизионному сенсору для максимального сохранения детализации, либо прибору с наименьшим уровнем шума для четкого выделения общего тренда. Выбор зависит от конкретной цели текущего исследования. Естественно, любые собираемые данные требуют обязательной нормализации и глубокого изучения статистических свойств еще до их слияния. Только после приведения всех параметров к единому масштабу можно безопасно приступить к непосредственному слиянию временных рядов.

Выбор метода объединения временных рядов напрямую зависит от качества исходных данных. При разной степени их надежности применяется взвешенное среднее, где веса отражают уровень доверия к источнику. Для метрологически идентичных каналов – с одинаковой погрешностью и частотой записи – достаточно вычислить простое среднее арифметическое:

$$y_t = \frac{\sum_{i=1}^n \omega_i \cdot x_{i,t}}{\sum_{i=1}^n \omega_i},$$

где y_t – объединенное значение в момент времени t , ω_i – вес i -го источника, $x_{i,t}$ – значение из i -го источника в момент времени t .

Анализ акустических сигналов традиционно требует декомпозиции исходных данных и перехода от временной развертки к частотному спектру. Такая трансформация позволяет извлекать ключевые параметры процесса. К ним могут относиться спектральный состав, амплитудное распределение и фазовые соотношения. С математической точки зрения сигналы могут быть описаны в двух плоскостях: временной $x = f(t)$ и частотной $X = F(\omega)$. В первом случае физический процесс интерпретируется как функция $x(t)$, определяющая мгновенное значение амплитуды в произвольный момент времени t , выступающий в качестве независимой переменной.

В частотной области сигнал описывается как функция частоты $X(\omega)$. В данном контексте $X(\omega)$ выступает спектральной характеристикой сигнала, $\omega = 2\pi f$ – независимая переменная, представляющая угловую частоту, а f – частота в герцах. Одним из инструментов для получения спектральных характеристик является преобразование Фурье. Преобразование Фурье позволяет представлять временную функцию, например, звуковую волну, в частотной области. Непрерывное преобразование Фурье (НПФ) определяется по формуле

$$X(\omega) = \int_{-\infty}^{\infty} x(t) \cdot e^{-i\omega t} dt,$$

где $X(\omega)$ – спектральная характеристика сигнала, $x(t)$ – значение исходного сигнала в момент времени t , t – время, j – мнимая единица, ω – угловая частота.

На практике часто используется дискретное преобразование Фурье (ДПФ), которое применяется к дискретным сигналам, поскольку компьютеры работают с конечными выборками данных. ДПФ вычисляется по аналогичной формуле, но с использованием конечных сумм:

$$X(k) = \frac{1}{N} \sum_{n=0}^{N-1} x(nT) \cdot e^{-i2\pi nk/N} = \frac{1}{N} \sum_{n=0}^{N-1} x(nT) \cdot W^{nk},$$

где T – период дискретизации, $n = \{0, 1, \dots, N-1\}$ – номер отсчета дискретизированного сигнала, $k = \{0, 1, \dots, N-1\}$ – номер гармоники сигнала, частота гармоник равна $\frac{k}{T}$, T – период измерения, W – вспомогательная функция.

Если сигнал нестационарный, Фурье-преобразование не сможет показать, в какой момент времени произошли эти изменения. Вейвлет-преобразование, наоборот, позволяет анализировать сигнал одновременно в частотной и временной областях. Оно использует базисные функции (вейвлеты), которые локализованы как по частоте, так и по времени. Это позволяет увидеть, когда и на каких частотах происходят изменения в сигнале. Вейвлет-преобразования делятся на два основных типа – непрерывное Вейвлет-преобразование (НВП) и дискретное Вейвлет-преобразование (ДВП). Непрерывное Вейвлет-преобразование является основным инструментом вейвлет-анализа. Вычисление прямого и обратного НВП осуществляется в соответствии с выражениями (8) и (9):

$$W_s(a, b) = (S(t), \Psi_{ab}(t)) = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} S(t) \Psi\left(\frac{t-b}{a}\right) dt, \quad (8)$$

$$S(t) = \frac{1}{C_{\Psi}} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} W_s(a, b) \Psi_{ab}(t) \frac{dadb}{a^2}, \quad (9)$$

где $W_s(a, b)$ – результат вейвлет-преобразования, $S(t)$ – исходный сигнал, a – параметр масштаба (определяет частоту), b – параметр сдвига (определяет временное положение), $\Psi(t)$ – вейвлет-функция, масштабированная и сдвинутая, C_{Ψ} – нормирующий коэффициент, для случая ортонормированных вейвлетов его значение принимается равным единице.

В отличие от НВП, которое за счет анализа на всех масштабах и сдвигах требует значительных вычислительных мощностей, ДВП оперирует фиксированными значениями этих параметров. В основе ДВП лежит концепция многомасштабного анализа с использованием ортогональных вейвлетов. Это позволяет рекурсивно разлагать исследуемый сигнал $S(t)$ на детализирующие $y_{high}[n]$ и аппроксимирующие $y_{low}[n]$ компоненты, последовательно уточняя его представление на разных уровнях разрешения. Разложение сигнала $S(t)$ реализуется через систему фильтров, вычисляющих коэффициенты по формулам:

$$y_{low}[n] = \sum_{k=-\infty}^{\infty} S[k]g[2n-k] \quad (35), \text{ и } y_{high}[n] = \sum_{k=-\infty}^{\infty} S[k]h[2n-k] \quad (36), \text{ где } g[k] \text{ – низкочастотный (аппроксимирующий) фильтр, } h[k] \text{ – высокочастотный (детализирующий) фильтр.}$$

В данном разделе рассмотрены существующие методы очистки и объединения данных для создания на их основе единого гибкого конвейера. По отдельности каждый метод хорошо работает только в «своей» узкой задаче. Ни один из них не позволяет одновременно убирать шумы, подстраиваться под изменения и выдавать точный результат при работе с разными источниками. Поэтому рассмотренные методы объединяются в общую структуру для повышения качества данных. Такая гибридная система позволяет подстроиться под логику обработки разнородных данных и минимизировать вычислительные затраты. Оптимизация конвейера позволяет использовать рассмотренные методы так же на устройствах с ограниченными вычислительными ресурсами.

Результаты. Для численной оценки эффективности и валидации предложенных подходов был проведен имитационный эксперимент на основе верифицированных эталонных климатических данных (температура, относительная влажность, атмосферное давление) за месячный период. С целью моделирования реальных условий эксплуатации измерительных систем в исходные эталонные временные ряды были искусственно внедрены случайные высокочастотные шумы и единичные аномальные пики (выбросы). Параметры генерируемых искажений рассчитывались индивидуально для имитации выходных сигналов конкретных типов датчиков различного класса точности: температуры (DHT22, DS18B20, BMP280), влажности (DHT22, DHT11) и давления (BMP280, BMP180).

Далее к сформированным зашумленным массивам данных был применен конвейер предварительной обработки. Качество фильтрации оценивалось путем прямого сравнения обработанных временных рядов с исходными эталонными значениями. На первом этапе (рис. 1) было выполнено обнаружение аномальных значений с помощью метода IQR. Замена аномалий выполнялась с помощью линейной интерполяции.

Далее выполнена фильтрация высокочастотных компонентов сигнала с помощью ФВЧ. На основе метода взвешенного среднего реализовано слияние данных. Сопоставление обработанных данных с зашумленными наглядно представлено на рис. 2.

При генерации выборки для датчика влажности DHT11 и барометрического сенсора BMP180 была смоделирована повышенная дисперсия случайной составляющей погрешности. Это обусловлено их более низким технологическим классом по сравнению с другими сенсорами. В итоговой таблице весов (табл. 1) более стабильные и прецизионные сенсоры (DHT22, DS18B20, BMP280) получили больший математический приоритет. Веса сенсоров температуры распределились почти поровну.

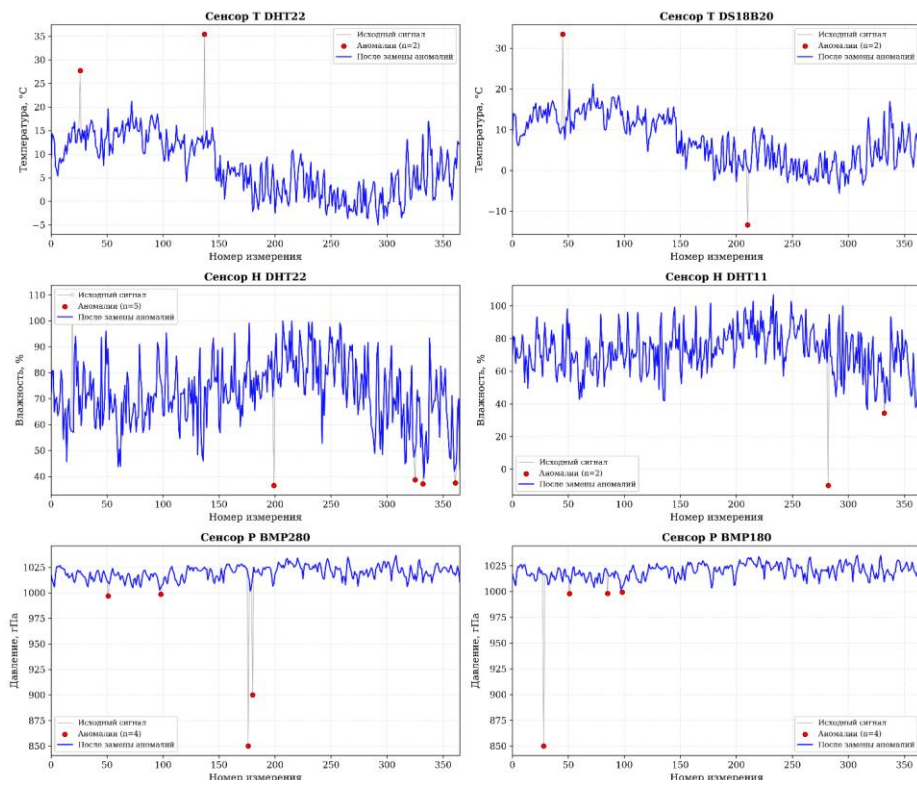


Рис. 1. Обнаружение и замена аномальных значений в показаниях сенсоров

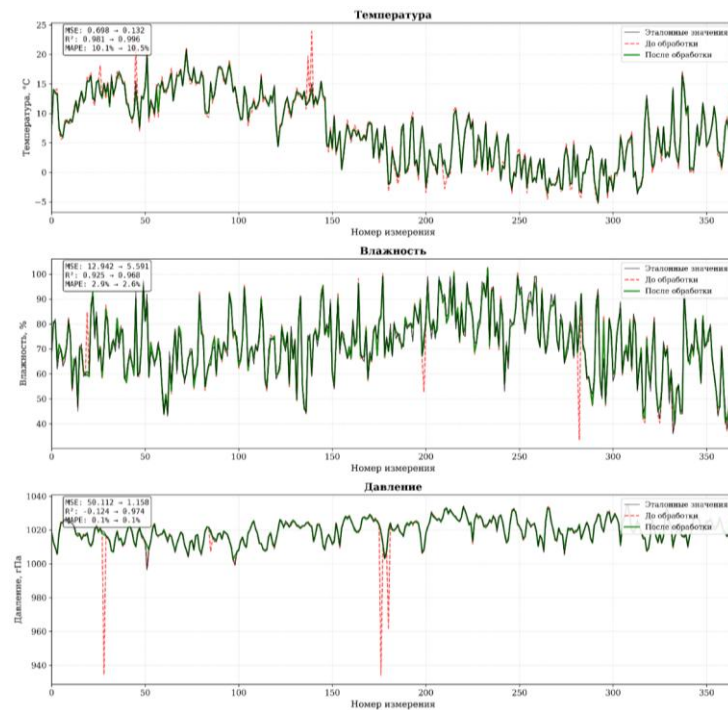


Рис. 2. Сравнение результатов слияния данных до и после обработки

Таблица 1

Весовые коэффициенты сенсоров

	Температура			Влажность		Давление	
Сенсор	T_DHT22	T_DS18B20	T_BMP280	H_DHT22	H_DHT11	P_BMP280	P_BMP180
Значение весов	0.336	0.336	0.329	0.522	0.478	0.504	0.496

Для глубокого анализа точности фильтрации было исследовано эмпирическое распределение плотности вероятности ошибок между эталоном и обработанным результатом. Из графиков (рис. 3) отчетливо видно, что распределения погрешностей для всех трех исследуемых параметров (температуры, влажности и давления) обладают свойствами симметричности и высокой концентрации значений вблизи нуля.

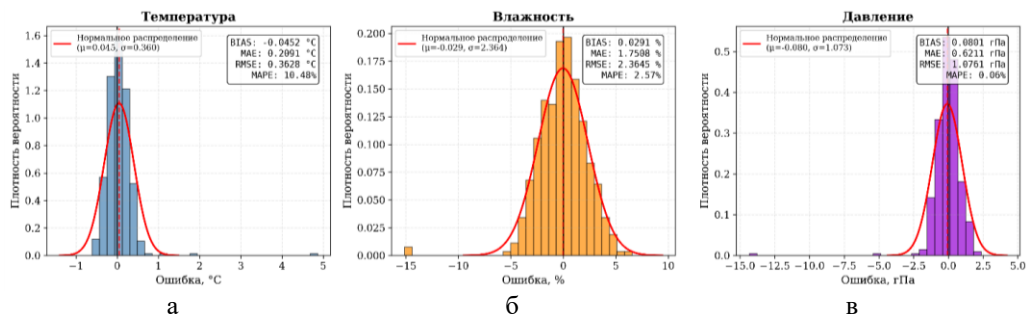


Рис. 3. Распределение ошибок обработанных данных

Эффективность предложенных решений была оценена количественно на основе расчета ключевых показателей качества. На рис. 4 представлен сравнительный анализ критериев точности до и после фильтрации.

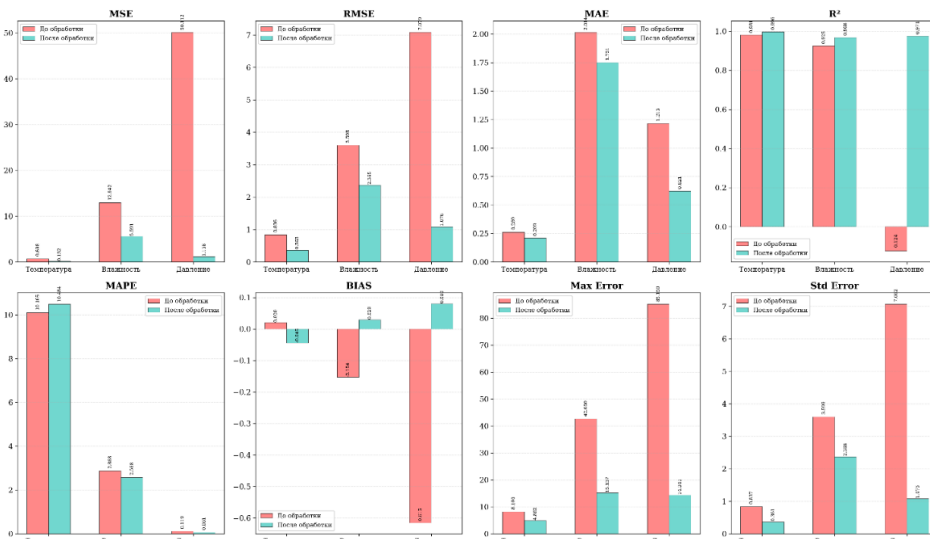


Рис. 4. Сравнительный анализ метрик качества до и после обработки

Гибридный конвейер позволил значительно повысить итоговое качество данных по нескольким ключевым метрикам. Средняя ошибка с учётом знака (BIAS) показывает систематическую погрешность данных. Среднеквадратическая ошибка (MSE) чувствительна к выбросам и аномалиям. Корень из среднеквадратической ошибки (RMSE) характеризует типичный разброс ошибок с усилением влияния крупных отклонений. Средняя абсо-

лютная ошибка (MAE) даёт линейную меру средней ошибки, менее чувствительна к выбросам, чем MSE/RMSE. Средняя абсолютная процентная ошибка (MAPE) позволяет сравнивать точность данных разного масштаба. Максимальная ошибка (Max Error) помогает обнаружить грубые промахи, выбросы, сбои датчиков или ошибки ввода. Стандартное отклонение ошибок (Std Error) отвечает за воспроизводимость данных. Коэффициент детерминации (R^2) показывает, какая доля разброса зависимой переменной объясняется построенной моделью зависимости.

Конвейер предварительной обработки и слияния данных демонстрирует универсальную высокую эффективность для всех рассматриваемых климатических параметров (табл. 2–4). Применение алгоритма привело к радикальному снижению величины ошибок (MSE, RMSE, MAE), минимизации отклонений Max Error и значительному росту коэффициента детерминации (R^2).

Таблица 2

Метрики качества обработки данных температуры

Метрика	До обработки	После обработки	Улучшение
MSE	0.6984	0.1317	+81.15%
RMSE	0.8357	0.3628	+56.58%
MAE	0.2604	0.2091	+19.71%
R^2	0.9807	0.9964	+0.0157
MAPE	10.1053	10.4844	-3.75%
BIAS	0.0201	-0.0452	—
Max Error	8.1000	4.8624	+39.97%
Std Error	0.8366	0.3605	+56.91%

Небольшой рост относительной ошибки (MAPE) для температуры на 3,75% обусловлен спецификой самой метрики. При этом абсолютные показатели (MSE, RMSE) резко улучшились. При приближении реальной температуры к 0 °С деление на близкие к нулю значения искусственно завышает MAPE. Этот скачок не указывает на реальное ухудшение качества обработки.

Таблица 3

Метрики качества обработки данных влажности

Метрика	До обработки	После обработки	Улучшение
MSE	12.9422	5.5911	+56.80%
RMSE	3.5975	2.3645	+34.27%
MAE	2.0145	1.7508	+13.09%
R^2	0.9248	0.9675	+0.0427
MAPE	2.8683	2.5681	+10.47%
BIAS	-0.1536	0.0291	—
Max Error	42.6500	15.1372	+64.51%
Std Error	3.5992	2.3676	+34.22%

Таблица 4

Метрики качества обработки данных давления

Метрика	До обработки	После обработки	Улучшение
MSE	50.1119	1.1580	+97.69%
RMSE	7.0790	1.0761	+84.80%
MAE	1.2148	0.6211	+48.87%
R^2	-0.1240	0.9740	+1.0980
MAPE	0.1191	0.0610	+48.77%
BIAS	-0.6153	0.0801	—
Max Error	85.1500	14.3415	+83.16%
Std Error	7.0618	1.0746	+84.78%

Регрессионный анализ подтверждает точность метода. На диаграммах рассеяния (рис. 5) видна сильная линейная связь между эталонными и обработанными значениями температуры, влажности и давления. Точки плотно ложатся на линию идеального соответствия, а коэффициент детерминации (R^2) более 0,96 доказывают, что конвейер сохраняет исходные физические закономерности.

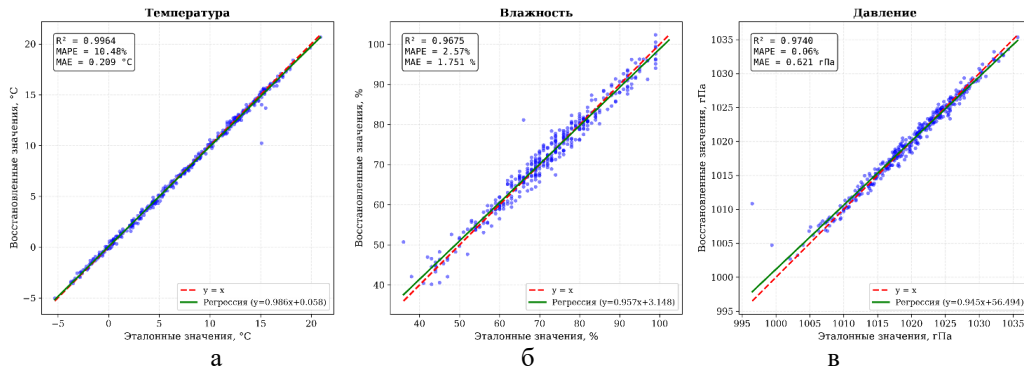


Рис. 5. Корреляция между эталонными и обработанными значениями

Эксперименты показали, что гибридный конвейер действительно справляется с шумами и аномалиями. Совместная работа IQR-фильтрации выбросов, высокочастотной очистки и динамического взвешенного слияния датчиков превращает сырые мультисенсорные данные в материал, пригодный для обучения моделей МО. Для всех трёх климатических параметров (табл. 2–4) R^2 вырос до 0,99, причём для сигнал атмосферного давления ($R^2 = -0,124$) удалось восстановить до 0,974. Конвейер не просто сглаживает шум, а возвращает реальную динамику измеряемой величины. Полученные цифры говорят о том, что подход снимает жёсткие требования к точности сенсоров: даже бюджетные датчики могут давать надёжный эталонный сигнал. В сравнении с обычным усреднением или фиксированными весами динамическое перераспределение приоритетов делает систему отказоустойчивой – временная деградация одного канала почти не влияет на итоговое значение. Локальная очистка и слияние данных без пересылки «сырых» потоков на сервер снижают нагрузку на сеть и энергопотребление, а в перспективе позволяют строить полностью децентрализованные самонастраивающиеся измерительные сети с высокой повторяемостью результатов.

Закключение. В ходе работы обоснован выбор методов предобработки разнородных данных и реализован конвейер, позволяющий повысить их качество. Для верификации конвейера сгенерирован синтетический набор данных.

Тесты на смоделированных климатических данных доказали высокую точность предложенного метода. Для значения параметра температуры R^2 увеличился на 0,0157 и составил 0,9964, а MAPE в процентах составляет 10,48%. Для относительной влажности R^2 вырос на 0,0427 до 0,9675, MAPE составляет 2,57%. Для атмосферного давления R^2 повысился на 1,0980, став положительным и составив 0,9740, MAPE составляет 0,06%.

Эксперименты подтвердили эффективность работы гибридного конвейера предобработки и мультисенсорного слияния данных. Алгоритм полностью устраняет шумы и аномалии, сводя к минимуму погрешности измерений. Объединение дублирующих каналов позволяет формировать единые информативные признаки и избегать ошибок, характерных для одиночных датчиков. Предложенные алгоритмы могут служить основой для систем мониторинга в режиме реального времени, комплексов проведения экспериментов in situ, а также для построения энергоэффективных сетей IoT с возможностью непосредственного развертывания на периферийных вычислительных устройствах.

Дальнейшие исследования будут направлены на оптимизирование вычислительной сложности алгоритмов для ещё более жёстких аппаратных ограничений. Также планируется разработать и интегрировать предложенное решение с протоколами передачи данных в беспроводных сенсорных сетях для создания сквозной самонастраивающейся системы сбора и обработки информации без участия центрального сервера.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Грязнов С.А., Кузнецов М.И. Факторы, влияющие на воспроизводимость научных исследований // Международный журнал гуманитарных и естественных наук. – 2024. – № 8-1 (95). – С. 69-71.
2. Jesus G. Casimiro A., Oliveira A. A Survey on Data Quality for Dependable Monitoring in Wireless Sensor Networks // Sensors. – 2017. – Vol. 17, No. 9. – 2010 p.
3. Zhang L. Jeong D., Lee S. Data Quality Management in the Internet of Things // Sensors. – 2021. – Vol. 21, No. 17. – 5834 p.
4. Боровская В.Г., Гомон Ю.М. Оценка качества данных реальной клинической практики // Реальная клиническая практика: данные и доказательства. – 2022. – Т. 2, № 4. – С. 10-16.
5. Papavasileiou A., Michalos G., Makris S. Quality control in manufacturing – review and challenges on robotic applications // International Journal of Computer Integrated Manufacturing. – 2025. – Vol. 38, No. 1. – P. 79-115.
6. Alexakis T., Adamopoulou E., Peppes N. [et al.]. Evaluating Data Quality: Comparative Insights on Standards, Methodologies, and Modern Software Tools // Electronics. – 2025. – Vol. 14, No. 15. – 3038 p.
7. Bogani R., Theodorou A., Arnaboldi L., Wortham R.H. Garbage in, toxic data out: a proposal for ethical artificial intelligence sustainability impact statements // AI and Ethics. – 2023. – Vol. 3, No. 4. – P. 1135-1142.
8. Kilkenny M.F., Robinson K.M. Data quality: “Garbage in – garbage out” // Health Information Management Journal. – 2018. – Vol. 47, No. 3. – P. 103-105.
9. Liu C., Ye F. A review of multimodal medical data fusion techniques for personalized medicine // Association for Computing Machinery. – New York: ACM, 2025. – P. 338-347.
10. Duan J., Xiong J., Li Y., Ding W. Deep learning based multimodal biomedical data fusion: An overview and comparative review // Information Fusion. – 2024. – Vol. 112. – 102536 p.
11. Azcarate S.M., Ríos-Reina R., Amigo J.M., Goicoechea H.C. Data handling in data fusion: Methodologies and applications // TrAC Trends in Analytical Chemistry. – 2021. – Vol. 143. – 116355 p.
12. Mourad A., Bertrand-Krajewski J.L., El-Youssef A. Edge Computing for Data Anomaly Detection of Multi-Sensors in IoT Environmental Monitoring // Electronics. – 2021. – Vol. 10, No. 3. – 302 p.
13. Zraunig M., Susto G.A., Cenedese A. Detecting Faults at the Edge via Sensor Data Fusion Echo State Networks // Sensors. – 2022. – Vol. 22, No. 8. – 2858 p.
14. Li L., Wang J., Zheng Y. Hybrid Machine Learning-Based Fault-Tolerant Sensor Data Fusion and Anomaly Detection for Fire Risk Mitigation in IIoT Environment // Sensors. – 2025. – Vol. 25, No. 7. – 2146 p.
15. Chauhan D., Shukla K., Desai D. [et al.]. A Comprehensive Review of IoT-Enabled Microgrids: Architectures, Challenges, and Future Prospects to achieve Sustainability // Next Research. – 2026. – 102156 p.
16. Bikakis N., Maroulis S., Papastefanatos G., Vassiliadis P. In-situ visual exploration over big raw data // Information Systems. – 2021. – Vol. 95. – 101616 p.
17. Maia A., Boutouchent A., Kardjadjia Y., Gherari M., Soyak E.G., Saqib M., Boussekar K., Cilbir I., Habibi S., Ali S.O., Ajib W. A survey on integrated computing, caching, and communication in the cloud-to-edge continuum // Computer Communications. – 2024. – Vol. 219. – P. 128-152.
18. Сергеев М.Б., Соловьев Н.В., Стадник А.И. Методы повышения контрастности растровых изображений для систем цифровой обработки видеoinформации // Информационно-управляющие системы. – 2007. – № 1. – С. 2-7.
19. Li D., Wang Y., Wang J., Wang C., Duan Y. Recent advances in sensor fault diagnosis: A review // Sensors and Actuators A: Physical. – 2020. – Vol. 309. – 111990 p.
20. Gasparyan L., Gasparyan F., Simonyan V. Internal Electrical Noises of BioFET Sensors Based on Various Architectures // OJBIPHY. – 2021. – Vol. 11, No. 02. – P. 177-204.
21. Jadhav S.P., Srinivas A., Dipak R.P., Ramkumar P.M., Suryawanshi J., Haldorai A. Deep learning approaches for multi-modal sensor data analysis and abnormality detection // Measurement: Sensors. – 2024. – Vol. 33. – 101157 p.
22. Tang Q., Liang J., Zhu F. A comparative review on multi-modal sensors fusion based on deep learning // Signal Processing. – 2023. – Vol. 213. – 109165 p.
23. Mitsis K., Zarkogianni K., Kalafatis E., Dalakleidi K., Jaafar A., Mourkousis G., Nikita K.S. A Multi-modal Approach for Real Time Recognition of Engagement towards Adaptive Serious Games for Health // Sensors. – 2022. – Vol. 22, No. 7. – 2472 p.
24. Chernov A.V., Savvas I.K., Alexandrov A.A., Kartashov O.O., Polyanchenko D.S., Butakova M.A., Soldatov A.V. Integrated Video and Acoustic Emission Data Fusion for Intelligent Decision Making in Material Surface Inspection System // Sensors. – 2022. – Vol. 22, No. 21. – 8554 p.
25. Wang Q.-G., Li X., Qin Q. Feature Selection for Time Series Modeling // JILSA. – 2013. – Vol. 05, No. 03. – P. 152-164.

REFERENCES

1. Gryaznov S.A., Kuznetsov M.I. Faktory, vliyayushchie na vosproizvodimost' nauchnykh issledovaniy [Factors affecting the reproducibility of scientific research], *Mezhdunarodnyy zhurnal gumanitarnykh i estestvennykh nauk* [International Journal of Humanities and Natural Sciences], 2024, No. 8-1 (95), pp. 69-71.
2. Jesus G. Casimiro A., Oliveira A. A Survey on Data Quality for Dependable Monitoring in Wireless Sensor Networks, *Sensors*, 2017, Vol. 17, No. 9, 2010 p.
3. Zhang L. Jeong D., Lee S. Data Quality Management in the Internet of Things, *Sensors*, 2021, Vol. 21, No. 17, 5834 p.
4. Borovskaya V.G., Gomon Yu.M. Otsenka kachestva dannykh real'noy klinicheskoy praktiki [Quality assessment of real-world data], *Real'naya klinicheskaya praktika: dannye i dokazatel'stva* [Real-World Data & Evidence], 2022, Vol. 2, No. 4, pp. 10-16.
5. Papavasileiou A., Michalos G., Makris S. Quality control in manufacturing – review and challenges on robotic applications, *International Journal of Computer Integrated Manufacturing*, 2025, Vol. 38, No. 1, pp. 79-115.
6. Alexakis T., Adamopoulou E., Peppes N. [et al.]. Evaluating Data Quality: Comparative Insights on Standards, Methodologies, and Modern Software Tools, *Electronics*, 2025, Vol. 14, No. 15, 3038 p.
7. Bogani R., Theodorou A., Arnaboldi L., Wortham R.H. Garbage in, toxic data out: a proposal for ethical artificial intelligence sustainability impact statements, *AI and Ethics*, 2023, Vol. 3, No. 4, pp. 1135-1142.
8. Killenny M.F., Robinson K.M. Data quality: “Garbage in – garbage out”, *Health Information Management Journal*, 2018, Vol. 47, No. 3, pp. 103-105.
9. Liu C., Ye F. A review of multimodal medical data fusion techniques for personalized medicine, *Association for Computing Machinery*. New York: ACM, 2025, pp. 338-347.
10. Duan J., Xiong J., Li Y., Ding W. Deep learning based multimodal biomedical data fusion: An overview and comparative review, *Information Fusion*, 2024, Vol. 112, 102536 p.
11. Azcarate S.M., Ríos-Reina R., Amigo J.M., Goicoechea H.C. Data handling in data fusion: Methodologies and applications, *TrAC Trends in Analytical Chemistry*, 2021, Vol. 143, 116355 p.
12. Mourad A., Bertrand-Krajewski J.L., El-Youssef A. Edge Computing for Data Anomaly Detection of Multi-Sensors in IoT Environmental Monitoring, *Electronics*, 2021, Vol. 10, No. 3, 302 p.
13. Zraunig M., Susto G.A., Cenedese A. Detecting Faults at the Edge via Sensor Data Fusion Echo State Networks, *Sensors*, 2022, Vol. 22, No. 8, 2858 p.
14. Li L., Wang J., Zheng Y. Hybrid Machine Learning-Based Fault-Tolerant Sensor Data Fusion and Anomaly Detection for Fire Risk Mitigation in IIoT Environment, *Sensors*, 2025, Vol. 25, No. 7, 2146 p.
15. Chauhan D., Shukla K., Desai D. [et al.]. A Comprehensive Review of IoT-Enabled Microgrids: Architectures, Challenges, and Future Prospects to achieve Sustainability, *Next Research*, 2026, 102156 p.
16. Bikakis N., Maroulis S., Papastefanatos G., Vassiliadis P. In-situ visual exploration over big raw data, *Information Systems*, 2021, Vol. 95, 101616 p.
17. Maia A., Boutouchent A., Kardjadjia Y., Gherari M., Soyak E.G., Saqib M., Boussekar K., Cilbir I., Habibi S., Ali S.O., Ajib W. A survey on integrated computing, caching, and communication in the cloud-to-edge continuum, *Computer Communications*, 2024, Vol. 219, pp. 128-152.
18. Sergeev M.B., Solov'ev N.V., Stadnik A.I. Metody povysheniya kontrastnosti rastrovyykh izobrazheniy dlya sistem tsifrovoy obrabotki videoinformatsii [Methods for increasing the contrast of raster images for digital video processing systems], *Informatsionno-upravlyayushchie sistemy* [Information and Control Systems], 2007, No. 1, pp. 2-7.
19. Li D., Wang Y., Wang J., Wang C., Duan Y. Recent advances in sensor fault diagnosis: A review, *Sensors and Actuators A: Physical*, 2020, Vol. 309, 111990 p.
20. Gasparyan L., Gasparyan F., Simonyan V. Internal Electrical Noises of BioFET Sensors Based on Various Architectures, *OJBIPHY*, 2021, Vol. 11, No. 02, pp. 177-204.
21. Jadhav S.P., Srinivas A., Dipak R.P., Ramkumar P.M., Suryawanshi J., Haldorai A. Deep learning approaches for multi-modal sensor data analysis and abnormality detection, *Measurement: Sensors*, 2024, Vol. 33, 101157 p.
22. Tang Q., Liang J., Zhu F. A comparative review on multi-modal sensors fusion based on deep learning, *Signal Processing*, 2023, Vol. 213, 109165 p.
23. Mitsis K., Zarkogianni K., Kalafatis E., Dalakleidi K., Jaafar A., Mourkousis G., Nikita K.S. A Multi-modal Approach for Real Time Recognition of Engagement towards Adaptive Serious Games for Health, *Sensors*, 2022, Vol. 22, No. 7, 2472 p.
24. Chernov A.V., Savvas I.K., Alexandrov A.A., Kartashov O.O., Polyanichenko D.S., Butakova M.A., Soldatov A.V. Integrated Video and Acoustic Emission Data Fusion for Intelligent Decision Making in Material Surface Inspection System, *Sensors*, 2022, Vol. 22, No. 21, 8554 p.
25. Wang Q.-G., Li X., Qin Q. Feature Selection for Time Series Modeling, *JILSA*, 2013, Vol. 05, No. 03, pp. 152-164.

Александров Александр Алексеевич – Научно-исследовательский и проектно-конструкторский институт информатизации, автоматизации и связи на железнодорожном транспорте АО «НИИАС», Ростовский филиал; e-mail: aa.alexandrov@vniias.ru; г. Ростов-на-Дону, Россия; тел.: + 79064278038.

Бутакова Мария Александровна – Ростовский государственный университет путей сообщения; e-mail: butakova@rgups.ru; г. Ростов-на-Дону Россия; тел.: +791859766 71; д.т.н.; профессор.

Aleksandrov Aleksandr Alekseevich – JSC NIIAS; e-mail: a.alexandrov@vniias.ru; Rostov-on-Don, Russia; phone: + 79064278038; Master's Degree in Computer Science.

Butakova Maria Aleksandrovna – Rostov State Transport University; e-mail: butakova@rgups.ru; Rostov-on-Don, Russia; phone: +79185976671; dr. of eng. sc., professor.

УДК 621.372

DOI 10.18522/2311-3103-2026-4-38-50

Н.Н. Войт

МОДИФИЦИРОВАННЫЙ АЛГОРИТМ ИНТЕГРАЦИИ СИСТЕМЫ УПРАВЛЕНИЯ ЖИЗНЕННЫМ ЦИКЛОМ ИЗДЕЛИЯ С ВИРТУАЛЬНОЙ ПЛОЩАДКОЙ ПРОЕКТИРОВАНИЯ

Низкая успешность разработки аппаратно-программных средств сложных изделий является проблемой в САПР, которая влечет траты материальных, технических и людских ресурсов, поэтому с целью повышения успешности разработки изделий работа вносит вклад в решение широко известной и важной задачи твердотельного моделирования – это задача по интеграции виртуальных прототипов. Задача включает цифровое представление изделия, отражающее пространственное взаимодействие компонентов и позволяющее оценивать работоспособность конструкций, и относится к PSPACE-классу сложных задач. В работе описана интеграция виртуальных прототипов на пример системы управления жизненным циклом изделия ЛОЦМАН:PLM с виртуальной площадкой проектирования Unity 3D. Работа содержит разработку модифицированного алгоритма поиска и извлечения в репозиторий 3D-сборок. Алгоритм отличается блоком поиска окружения сборки api.getinfo, позволяющим найти и поместить в графический репозиторий конструкторскую, технологическую и эксплуатационную документацию (данные, метаданные и прочее). По помещенным данным в виртуальную площадку специалистом выполняется оценка эргономики 3D-сборок. В рамках крупного проектно-промышленного предприятия проведены вычислительные эксперименты с модифицированным алгоритмом, реализованном как интегратор 3D-сборок, и получены примеры применения, а также выполнено внедрение интегратора. В виртуальной реальности специалистом было оценено расположение рабочих органов и блоков по примерам. Модифицированный алгоритм вносит вклад в решение проблемы успешности, сокращая сроки построения, оценки и правок 3D-сборок на 10 % (на 1,25 рабочих дня) по результатам экспериментов. Интегратор используется в конструкторском проектировании 3D-сборок и в процессе обучения (повышении квалификации, переподготовке) специалистов без отрыва от производства. Благодаря интерфейсной реализации интегратор взаимодействует с продуктами САПР компаний России и других стран. В работе используется теоретико-множественное обоснование для математического описания выражений, параметров и значений. Дальнейшие исследования связаны с выявлением ошибок по поведенческой модели специалиста.

Виртуальная реальность; ЛОЦМАН; Unity 3D; проектирование; интеграция; изделие; прототип.

N.N. Voit

MODIFY INTEGRATION ALGORITHM BETWEEN CONTROL SYSTEM OF PART LIFE CYCLE AND VIRTUAL DESIGN PLATFORM

The low hardware and software development success is a problem in CAD. The problem entails large expenditures of material, technical and human resources. In order to increase the product development success need to solve the issue of integrating virtual prototypes in the design stage. The issue includes a digital representation that allows evaluate the design's operability. The paper describes the integration between LOTSMAN:PLM and Unity 3D, that includes the modified algorithm development for searching and extracting 3D assemblies in the repository. This algorithm has a name api.getinfo, that

allows to aggregate design, technical and operational information. These information puts into a query to extract a 3D assembly for a graphical repository. A specialist can estimate the assembly ergonomics in virtual reality. At industrial enterprise computational experiments were carried out with a modified algorithm implemented like an integrator of 3D assemblies into virtual reality. 3D assembly's examples were obtained in a graphical repository. A specialist can estimate the location of working bodies and blocks in virtual reality. The integrator interacts with products from ASCON, SolidWorks, AutoCAD, Altium Designer and Delta Design. The modified algorithm reduces the time required to build, evaluate, and edit a 3D assembly by 10% (1.25 business days). Results are based on the experiments. The integrator is embedded in a design and industrial enterprise and in training (advanced training, retraining) of specialists. The work uses a set-theoretic justification for mathematical description of expressions, parameters, etc. Further research is related to the identification of errors in the behavioral specialist model.

Virtual reality; LOODSMAN; Unity 3D; design; integration; part; prototype.

Введение. В настоящее время в проектной деятельности предприятий используются как системы автоматизированного проектирования (PLM/CAD/CAM/CAE-системы), так и автоматизированные системы управления и накопления знаний. Результатом проектной деятельности являются проектные решения, систематизируемые и накапливаемые для повторного использования в производстве. При этом типовым методом моделирования в CAD -системах является создание трехмерного объекта из двумерного эскиза, над которым совершаются операции выдавливания, вращения и др. Основным недостатком такого метода моделирования является то, что при создании трехмерной модели проектировщику приходится мыслить не в категориях человеко-компьютерного диалога (эргономика), а в терминах проектируемых деталей (основание, отверстие, ребро жесткости и т.д.). Эту проблему возможно решить, переместив моделируемые объекты в виртуальное окружение и «подменив» им реальное, тем самым создается *виртуальная реальность* (Virtual Reality, VR), новый тип человеко-машинного диалога, при котором зрительные, слуховые, осязательные и моторные ощущения заменяются их имитацией [1].

За последние годы на многих предприятиях в России [2–4] и за рубежом внедрена технология виртуальной реальности, ее успешно применяют на таких предприятиях, как NASA, Lockheed Martin Corporation, Fiat Chrysler Automobiles, Ford, General Motors, «Норникель», «АвтоВАЗ», «Росатом», «Роснефть», «Лукойл» [1].

Технология виртуальной реальности обеспечивает высокореалистичное моделирование трехмерного пространства, поддерживает динамическое интерактивное взаимодействие с пользователем, создает визуальный, аудиальный, кинестетический эффект погружения в моделируемое киберпространство [5–8].

Проектирование с помощью технологии виртуальной реальности (виртуальное проектирование), согласно лабораторным исследованиям, позволяет снизить срок разработки изделия до 30 %, сократить количество проектных ошибок, повысить качество изделия за счет улучшений в проработке эргономики, ремонтных и эксплуатационных характеристик, повысить удовлетворенность клиента изделием за счет его вовлечения в стадию испытания изделия, снизить ремонтные и эксплуатационные издержки за счет прототипирования процессов производства и эксплуатации изделий [5–8]. Эта технология применяется на всех этапах жизненного цикла, от концептуального проектирования до обслуживания изделия. В отличие от физического макета, создаваемого только после выработки требований к изделию и организации производства, виртуальный макет может быть создан сразу после того, как заданы основные требования к изделию. Благодаря этому становится возможным проводить процессы верификации и разработки параллельно, что ведет к сокращению числа ошибок. Проектирование в виртуальной реальности позволяет дополнительно обратить особое внимание на такие стороны продукта, как эргономика, расположение рабочих органов и блоков, проектирование инструментария, удобство обслуживания.

Среди преимуществ использования технологий виртуальной реальности выделяют следующее:

- ◆ задействован спектр рецепторных систем человека: визуальной, аудиальной, кинестетической. За счет этого осуществляется согласованный процесс передачи информации по нескольким каналам одновременно;

- ◆ полностью погрузиться в созданную среду;
- ◆ интерактивно закрепляются полученные знания и усвоенные навыки.

Проектирование в виртуальной реальности обладает следующими преимуществами перед бумажным и «мониторным»:

а) больше площадь для проектирования по сравнению с мониторами, даже если их будет больше одного. К тому же, в последнем случае существенно увеличиваются затраты на оборудование: мониторы, кабели, контроллеры и прочее;

б) окна на VR-сцене обладают свойством интерактивности, с ними можно взаимодействовать – перемещать, сворачивать, разворачивать, добавлять, убирать, вводить и выводить данные – и настроить так, чтобы они обновлялись в режиме реального времени;

в) проектировщик работает в трехмерном пространстве, в то время как бумага и монитор предоставляют только двумерное.

Проблемами интеграции промышленных прикладных программ САПР (ПП САПР [9–13]) с технологиями виртуальной реальности более чем два десятилетия занимаются многие ведущие зарубежные и российские университеты и компании. Далее по тексту представлен обзор известных работ.

В статье [14] автор предлагает фреймворк для виртуальной сборки, связывающий элементарные логические сборочные шаги и изменения геометрического состояния сцены. Используемая в нем модель, основанная на фреймворке знаний, состоит из соединяемых частей (портов, *ports*) и видов соединений (*connections*), а также их таксономий из более чем 20 понятий. Алгоритмы сборки на системном этапе, излагаемые на естественном языке, обеспечивают более абстрактный уровень моделирования. При этом структура одного и того же соединения применяется как в процессе сборки, так и в разборке. Однако авторы работы не учитывают операции, для которых могут предназначаться два инструмента одновременно или настраиваемые устройства (например, паяльная станция).

В статье [15] авторы разработали модель данных для обмена между VR и ПП САПР. Модель представляет собой промежуточное звено, включающее элементы визуализации и порты ввода, вывода. Модель имеет дерево построения и граничное представление (Boundary Representation, B-Rep). В модель заключен принцип жесткого именования (*persistent naming*) элементов граничного представления, исключающий неоднозначности при обновлении дерева построения. Мультимаркировка (*Multi-labelling*) с помощью меток имеет следующий формат:

NAME (ID_skel, ID_op) или *NAME (ID_op)*,

где *NAME* – название метки, *ID_op* – идентификатор операции, *ID_skel* – идентификатор родительского B-Rep-элемента.

Такой формат позволяет более полно представить дерево построения. Благодаря такой модели нет необходимости переводить 3D-сборку из одного формата в другой для ее визуализации в VR. Однако данная модель не учитывает физические параметры 3D-объектов (масса, материал и т.д.).

В статье [16] авторы предложили интеграцию ПП САПР и VR на основе языка моделирования VRML. Модель интеграции основана на кроссплатформенном обеспечении (VRML, XML, Java). VRML-модели, получаемые преобразованием 3D-сборок в ПП САПР, хранятся в собственной базе данных, свойства моделей хранятся в отдельной базе данных, конфигурацииборок ПП САПР описаны на XML-языке. Специалист имеет возможность взаимодействовать с моделью в виртуальной реальности, меняя атрибуты элементов сборки (материал, внешний вид и др.) и ее конфигурацию. При этом отсутствие математического описания модели в работе затрудняет оценку степени масштабируемости, а также авторы не рассмотрели случаи изменений 3D-сборок.

Таким образом, нет известных алгоритмов по интеграции виртуальных прототипов, учитывающих эргономические параметры 3D-сборок и совместимых с отечественным ПП САПР, а также алгоритмов позволяющих собирать конструкторскую, технологическую и эксплуатационную информацию по 3D-сборке для использования в виртуальной реальности в решении задачи по оценке эргономики; поэтому целью работы является

разработка и реализация (как программного интегратора) модифицированного алгоритма интеграции прикладного программного обеспечения САПР (на примере системы управления жизненным циклом изделия ЛОЦМАН:PLM) с виртуальной площадкой проектирования (на примере Unity 3D). Алгоритм отличается блоком поиска окружения сборки *api.getinfo*, позволяющим агрегировать конструкторскую, технологическую и эксплуатационную информацию и поместить ее в запрос на извлечение в графический репозиторий 3D-сборки. Интегратор используется специалистом в виртуальной реальности в оценке эргономических параметров изделия.

Постановка задачи. Широко известной задачей твердотельного моделирования является задача по интеграции виртуальных прототипов (3D-сборок), включающей цифровое представление изделия и имеющей -класс сложности. Цифровое представление изделия отражает пространственное взаимодействие компонентов и позволяет оценивать работоспособность конструкции в автоматизированном проектировании аппаратно-программных средств сложных изделий.

Следует отметить, что последующая формализация задачи в работе необходима для понимания того, как можно сократить затраченное время специалистом на оценку эргономических характеристик изделия, поэтому начальные данные, необходимые для решения задачи, описаны в формате Хоара [21] на теоретико-множественном языке, а именно:

$$3D\_assembly = (DesignBegin \bowtie DesignBody \bowtie DesignEnd) \vee ErgonomicEstimate,$$

где *DesignBegin* – старт проекта;

DesignBody – тело проекта;

DesignEnd – завершение проекта;

ErgonomicEstimate – интегральная величина затраченного времени специалистом на оценку параметров эргономики по 3D-сборке,

ErgonomicEstimate $\rightarrow min$, где *min* – минимальное значение;

$\vee$ – логический оператор «И»;

$\bowtie$ – композиция членов выражения.

Суть постановки задачи сводится к следующему для того, чтобы на этапе конструкторского проектирования до создания физического макета параллельно проводить процессы выявления ошибок и создания структуры (конструкции), что ведет к сокращению сроков, надо обратить внимание на эргономику 3D-сборок в виртуальной реальности, т.е. на расположение рабочих органов и блоков, проектирование инструментария, удобство обслуживания и прочее; поэтому необходимо обнаружить ошибки конструкторского, технологического и эксплуатационного характера в 3D-сборке по полученным через интерфейсы данным и метаданным от прикладным программным средствам САПР, и для этого надо модифицировать известный алгоритм интерфейсного взаимодействия прикладных средств САПР. Такая модификация будет расширять функционал и выполнять поиск и извлечение конструкторской, технологической и эксплуатационной информации (метаданных, данных и др.).

Указанное расширение функционала интеграции можно представить функцией *Find_Extraction*, а именно:

$$[Value_n] = Find\_Extraction[3D\_assembly, Interface],$$

где $\{Value_n\}$ – множество из данных по конструкторской, технологической и эксплуатационной документации по 3D-сборке, где $n \in \mathbb{N}$,

$$BodyTree \cup BuildProcedure \cup Relation \cup Technology \cup Use \subseteq \{Value_n\},$$

где *BodyTree* – дерево построения 3D-сборки в прикладных программах САПР,

$$BodyTree \in \{\text{эскиз, деталь, сборка,}$$

фронтальная проекция, горизонтальная проекция, профильная проекция, ... \};

$$BuildProcedure = Logic \cup 3d,$$

где *Logic* = {И, ИЛИ, НЕ, ...} – базис из логических операций «И», «ИЛИ» и «НЕ»,

3d = {выдавливание, вращение, построение по траектории, ...} – базис из проектных процедур;

Relation = {соосность, параллельность, перпендикулярность, касание, ...} – набор отношений расположения элементов конструкции изделия;

Technology = {маршрут, нормы времени, требования по ресурсам, ...} – характеристики технологического процесса изготовления изделия;

Use = {технические условия, руководство по эксплуатации, ...} – описание эксплуатационной документации по изделию;

Interface $\in \{PLM, CAD, CAM, CAE, \dots\}$ – множество из программных интерфейсов по работе с соответствующими средствами САПР: PLM-, CAD-, CAM, CAE-системами.

«...» обозначает масштабируемость множества.

Таким образом, для решения поставленной задачи необходимо при соблюдении условия *ErgonomicEstimate* = *min* разработать функцию *Find_Extraction*, являющуюся модификацией алгоритма интеграции, реализовать модифицированный алгоритм интеграции, провести вычислительные эксперименты и оценить результаты.

Модифицированный алгоритм интеграции. ЛОЦМАН:PLM [17, 18] представляет собой отечественную систему управления жизненным циклом изделий, позволяющую решать задачи с учетом специфики любого проектно-промышленного предприятия. Базовыми функциями являются управление данными об изделии, документами, структурой, конфигурацией изделия. Прикладные модули ЛОЦМАН:PLM позволяют решать задачи в процессах технологической подготовки производства, управления изменениями конструкторской и технологической документации. Инструменты разработки API и SDK, предоставляемые с системой, позволяют настраивать и масштабировать ее функционал в соответствии с требованиями предприятий. Электронное описание изделия в ЛОЦМАН:PLM организовано в виде совокупности взаимосвязанных информационных объектов, состоящих из множества атрибутов и файлов, описывающих характеристики изделия, его конструкцию, технологию изготовления и другие сведения на всех стадиях его жизненного цикла. ЛОЦМАН:PLM позволяет взаимодействовать не только с продуктами компании АСКОН, но и с различными сторонними прикладными программами, такими как SolidWorks, AutoCAD, Altium Designer, Delta Design.

Далее на рис. 1 приведена структурная схема обмена сведениями между акторами процесса интеграции. Интегратор подключается к серверу ЛОЦМАН:PLM и получает доступ к программному интерфейсу (API) и базе данных (БД) системы. С помощью API он производит поиск объектов в БД, имеющих тип «3D-модель сборки». Для каждого из найденных объектов интегратор находит привязанные к ним файлы, считывает их и сохраняет в графическом репозитории. Сам процесс интеграции представлен в виде UML-диаграммы последовательности на рис. 2.

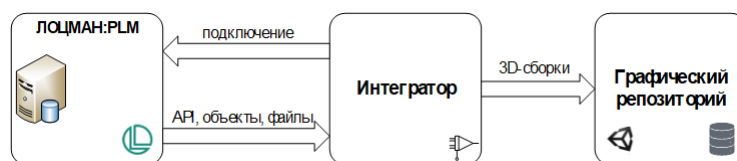


Рис. 1. Обмен

Модифицированный алгоритм извлечения 3D-сборок приведен на рис. 3., в котором вставлен новый блок *api.getinfo*, являющийся реализацией *Find_Extraction* и расширением функционала алгоритма.

Далее по работе описан процесс интеграции в теоретико-множественном описании с листингами исходного кода.

Интегратор подключается к базе данных ЛОЦМАН:PLM следующим образом:

$$CONNECT: DATA \rightarrow LOODSMAN \cup API,$$

где $DATA = \{DBName, LoodsmanConnection\}$ – параметры подключения;

DBName – имя базы данных;

LOODSMAN – база данных ЛОЦМАН:PLM;

API – программный интерфейс (API) ЛОЦМАН:PLM.



Рис. 2. Взаимодействие акторов (Интегратор, ЛОЦМАН:PLM, Графический репозиторий) в процессе интеграции

Подключение интегратора реализовано в листинге 1.

```

var temp = Type.GetTypeFromCLSID(Guid.Parse(
    "044259A0-FEAE-4B55-B6FE-E9B024FCCD4B"));
ILoodsmanConnection connection =
    Activator.CreateInstance(temp) as ILoodsmanConnection;
int code = 0; string msg = "";
string stDBName = "demo_22";
connection.API8.UniConnect(stDBName, "");
Console.WriteLine("Подключение успешно...");
var api = connection.API8;
api.RunMethod("SetFormat", new object[] { "", code, msg });
    
```

Листинг 1. Подключение интегратора к ЛОЦМАН:PLM

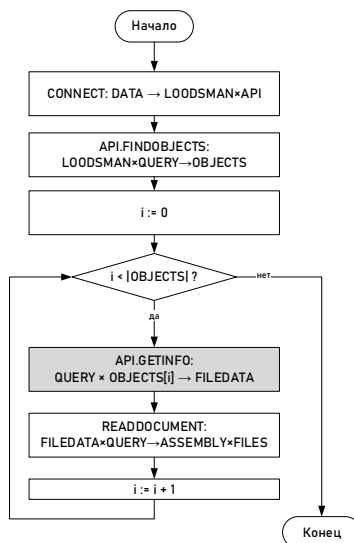


Рис. 3. Модифицированный алгоритм поиска и извлечения 3D-сборок в ЛОЦМАН:PLM

В базе производится поиск 3D-сборок [18]:

$$API.FINDOBJECTS: LOODSMAN \times QUERY \rightarrow OBJECTS,$$

где $QUERY = (Type, Product, Version, State, AttrCondition, Projects, Link)$ – параметры запроса, $ObjectType = \{ "3D - модель сборки" \}$, $Product = ""$, $Version = ""$, $State = \emptyset$, $AttrCondition = \emptyset$, $Projects = ""$, $Link = ""$;

$OBJECTS = \{ OBJECT_i \}$ – объекты, соответствующие запросу,

$OBJECT_i = (id, type, product, version, state, document, access, locked | id \in \mathbb{N}, document \in \{0,1\}, access \in \{1,2,3\}, locked \in \{0,1,2\})$.

Поиск 3D-сборок реализован в листинге 2.

```
var link = api.RunMethod("FindObject", new object[] {
    "3D-модель сборки", "", "", "", "", "", "", code, msg });
DataProvider.DataSet ds = DataParse(link);
```

Листинг 2. Поиск 3D-сборок в ЛОЦМАН:PLM

Далее в блоке для каждой сборки по идентификатору версии осуществляется поиск и извлечение конструкторской, технологической и эксплуатационной информации (прикрепленных файлов, метаданных и др.):

$$API.GETINFO: QUERY \times OBJECTS[i] \rightarrow FILEDATA,$$

где $OBJECTS[i] \in OBJECTS$;

$QUERY = \{ DocumentID \}$ – параметры запроса,

$DocumentID$ – уникальный номер 3D-сборки, или версии документа;

$FILEDATA = \{ FILE_i \}$ – данные о прикрепленных файлах,

$FILE_i = (id, file, name, localname, creationdate, size, readonly, crc, modified | id = DocumentID, file \in \mathbb{N}, size \in \mathbb{N}, readonly \in \mathbb{N}, crc \in \mathbb{N})$.

Получение информации о 3D-сборке реализовано в листинге 3.

```
var fileversion = api.RunMethod(
    "GetInfoAboutVersionsFiles", new object[] {
    Int32.Parse("" + ds.FieldValue[ds.FieldName[i]]), code, msg });
DataProvider.DataSet filedata = DataParse(fileversion);
```

Листинг 3. Получение данных о 3D-сборке

По полученным параметрам выгружается файл 3D-сборки [18]:

$$READDOCUMENT: FILEDATA \times QUERY \rightarrow ASSEMBLY \cup FILES,$$

где $QUERY = \{ VersionID, DocumentID, Linktype, Name, LocalName, Mode, Type \}$ – параметры запроса,

$VersionID$ – номер версии документа,

$DocumentID$ – номер документа в объекте,

$Linktype$ – тип связи,

$Linktype = "Документы"$,

$Name$ – имя документа,

$LocalName$ – локальное имя документа,

$Mode$ – режим получения файла,

$Mode = 0$,

$Type$ – тип документа, $Type = 43$;

$ASSEMBLY$ – исходный файл 3D-сборки;

$FILES$ – дополнительные файлы, загружаемые с 3D-сборкой (детали, вложенные сборки, чертежи, спецификации).

Чтение и выгрузка файла 3D-сборки реализовано в листинге 4.

```
int res = loadsmanIntegrator.Connect(
"SRV-ASCON-APP1", "demo_22", 0, "", "");
byte[] _doc = loadsmanIntegrator.ReadDocument(
    ds.FieldName[i].ToString(),
    filedata.FieldValue["_ID_DOCUMENT"],
    "Документы",
    filedata.FieldValue["_NAME"] as string,
    filedata.FieldValue["_LOCALNAME"] as string,
    0,
    43);
File.WriteAllText(filedata.FieldValue["_NAME"] +
"_log.txt", ObjectToString(_doc));
string destination = @"3D-сборки/" +
filedata.FieldValue["_NAME"] + "/";
if (!Directory.Exists(destination))
Directory.CreateDirectory(destination);
// предварительно создаем папки для вспомогательных файлов...
List<string> dirs = new List<string>
(Directory.GetDirectories(buffer));
while (dirs.Count != 0)
{
    List<string> newDirs = new List<string>();
    foreach (var _dir in dirs)
    {
        Directory.CreateDirectory(destination +
_dir.Substring(buffer.Length));
        newDirs.AddRange(Directory.GetDirectories(_dir));
    }
    dirs = newDirs;
}
// ...и перемещаем их...
List<string> files = new List<string>
(Directory.GetFiles(buffer, "*", SearchOption.AllDirectories));
foreach (var _file in files)
{
    string newPath = destination +
_file.Substring(buffer.Length);
    if (!File.Exists(newPath))
        File.Copy(_file, newPath);
    System.Security.AccessControl.FileSecurity fs =
File.GetAccessControl(_file);
    File.Delete(_file);
}
// ...после чего очищаем буфер
dirs = new List<string>(Directory.GetDirectories(buffer));
foreach (var _dir in dirs)
{
    Directory.Delete(_dir);
}
Console.WriteLine("Файл загружен и доступен по адресу:
{0}{1}/{2}", destination,
filedata.FieldValue["_LOCALNAME"], filedata.FieldValue["_NAME"]);
```

Листинг 4. Чтение и сохранение файла 3D-сборки

Полный алгоритм извлечения 3D-сборок в ЛОЦМАН:PLM приведен в листинге 5.

```
// 1) Подключение к ЛОЦМАН:PLM
var temp = Type.GetTypeFromCLSID(Guid.Parse(
"044259A0-FEAE-4B55-B6FE-E9B024FCCD4B"));
ILoodsmanConnection connection =
Activator.CreateInstance(temp) as ILoodsmanConnection;
int code = 0; string msg = "";
string stDBName = "demo_22";
connection.API8.UniConnect(stDBName, "");
Console.WriteLine("Подключение успешно...");
var api = connection.API8;
api.RunMethod("SetFormat", new object[] { "", code, msg });
// 2) Поиск объектов по типу "Сборочная единица"
var link = api.RunMethod("FindObjects", new object[] {
"3D-модель сборки", "", "", "", "", "", "", code, msg });
DataProvider.DataSet ds = DataParse(link);
Integrator loadsmanIntegrator = new Integrator();
while (!ds.EOF)
{
    for (int i = 0; i < ds.FieldCount; i++)
    {
        if (ds.FieldName[i] == "_ID_VERSION") {
// 3) Поиск данных о 3D-сборке
link = api.RunMethod("GetLinkedFast", new object[] {
Int32.Parse("" + ds.FieldValue[ds.FieldName[i]]),
"Документы", true, code, msg });
DataProvider.DataSet linkdata = DataParse(link);
var fileversion = api.RunMethod("GetInfoAboutVersionsFiles",
new object[] { Int32.Parse("" +
ds.FieldValue[ds.FieldName[i]]), code, msg });
DataProvider.DataSet filedata = DataParse(fileversion);
// 4) Чтение файла 3D-сборки
Console.WriteLine("Открывается файл {0}...",
filedata.FieldValue["_NAME"]);
int res = loadsmanIntegrator.Connect("SRV-ASCON-APP1", "demo_22",
0, "", "");
byte[] _doc = loadsmanIntegrator.ReadDocument(
linkdata.FieldValue["_ID_VERSION"].ToString(),
filedata.FieldValue["_ID_DOCUMENT"],
"Документы",
filedata.FieldValue["_NAME"] as string,
filedata.FieldValue["_LOCALNAME"] as string, 0, 43);
Console.WriteLine("Перемещаем файл...");
string destination = @"3D-сборки/" +
filedata.FieldValue["_NAME"] + "/";
if (!Directory.Exists(destination))
Directory.CreateDirectory(destination);
// предварительно создаем папки для вспомогательных файлов...
List<string> dirs =
new List<string>(Directory.GetDirectories(buffer));
while (dirs.Count != 0)
{
```

```

        List<string> newDirs = new List<string>();
        foreach (var _dir in dirs)
        {
            Directory.CreateDirectory(destination
+ _dir.Substring(buffer.Length));
            newDirs.AddRange(Directory.GetDirectories(_dir));
        }
        dirs = newDirs;
    }
    // ...и перемещаем их...
    List<string> files = new List<string>
(Directory.GetFiles(buffer, "*", SearchOption.AllDirectories));
    foreach (var _file in files)
    {
        string newPath = destination + _file.Substring(buffer.Length);
        if (!File.Exists(newPath))
            File.Copy(_file, newPath);
        System.Security.AccessControl.FileSecurity fs =
        File.GetAccessControl(_file);
        File.Delete(_file);
    }
    // ...после чего очищаем буфер
    dirs = new List<string>(Directory.GetDirectories(buffer));
    foreach (var _dir in dirs)
    {
        Directory.Delete(_dir);
    }
    Console.WriteLine("Файл загружен и доступен по адресу:
{0}/{1}/{2}", destination, filedata.FieldValue["_LOCALNAME"],
filedata.FieldValue["_NAME"]);
    }
}
ds.Next();
}

```

Листинг 5. Извлечение 3D-объекта в ЛОЦМАН:PLM

Вычислительные эксперименты. Далее описаны проведенные вычислительные эксперименты с модифицированным алгоритмом, оценивающие сокращение затраченного времени на построение, оценку и правку 3D-сборок. Замеры затраченного времени без/с применение модифицированного алгоритма приведены в табл. 1 (единицы измерения времени – это часы), а примеры 3D-сборок, интегрированные в виртуальную реальность – на рис. 4 и 5.

Таблица 1

Замеры затраченного времени по вычислительным экспериментам

	Построение, ч.	Оценка, ч.	Правка, ч.	Сумма, ч.
Без модифицированного алгоритма	72	19	7	98
С модифицированным алгоритмом	72	9	7	88

Из таблицы взяты суммы по затраченному времени без применения модифицированного алгоритма 98 (часов) и с применением модифицированного алгоритма 88 (часов) соответственно для расчета сокращения сроков η :

$$\eta = \frac{88 \text{ (часов)}}{98 \text{ (часов)}} \approx 0,9.$$

Таким образом, вычислительные эксперименты показали уменьшение затраченного времени на 10 % (на 1,25 рабочих дня) по сравнению с применением алгоритма без нового блока поиска окружения сборки *api.getinfo*.

Также следует отметить то, что новый блок в модифицированном алгоритме дополняет 3D-сборку описанием по конструированию, технологии изготовления и рекомендациям по эксплуатации, которое важно при оценке специалистом эргономических характеристик (расположения рабочих органов и блоков, проектирования инструментария, удобства обслуживания и др.) в виртуальной реальности.

В ходе выполнения экспериментов в работе для визуализации 3D-сборок использовалась среда разработки сцен Unity 3D [19]. По причине того, что она загружает ограниченное число форматов трехмерных объектов, то требуется преобразовать сборку в один из читаемых форматов. Процесс преобразования 3D-сборок выглядит следующим образом: 1) открыть 3D-сборку в КОМПАС-3D; 2) выбрать команду меню «Файл» → «Экспорт» и выбрать формат ACIS (*.sat); 3) экспортированную сборку открыть в программе CAD Exchanger Lab [20]; 4) выбрать команду меню «Export» → FBX; 5) если остались непреобразованные сборки, перейти к шагу 1.

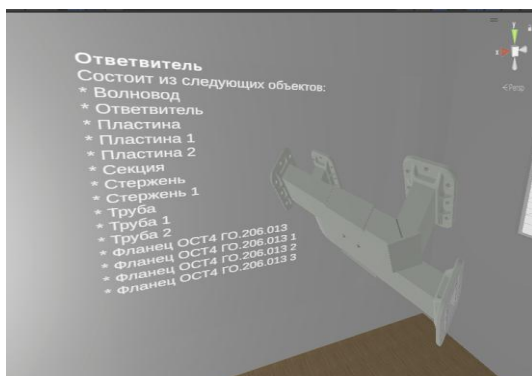


Рис. 4. Пример 3D-сборки «Ответвитель» в виртуальной реальности

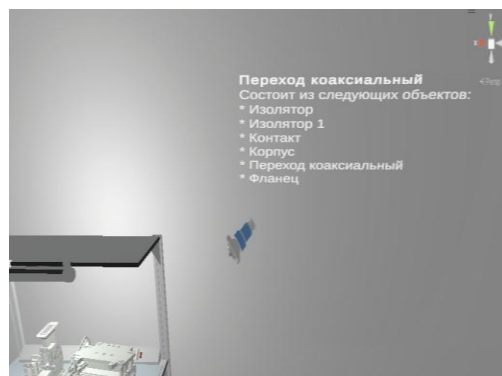


Рис. 5. Пример 3D-сборки «Переход коаксиальный» в виртуальной реальности

Заключение. Работа вносит вклад в решение одной из важнейших задач проектирования виртуальных прототипов – цифровое представление изделия, отражающее пространственное взаимодействие компонентов и позволяющее оценивать работоспособность конструкции.

В статье решена поставленная задача и представлена реализация функции *Find_Extraction* в виде нового блока *api.getinfo* в структурной схеме модифицированного алгоритма поиска и извлечения 3D-сборок в ЛОЦМАН:PLM.

За счет применения модифицированного алгоритма в конструкторском проектировании аппаратно-программных средств сложных изделий сокращение затраченного времени на построение, оценку и правку 3D-сборок было уменьшено на 10% (на 1,25 рабочих дня). Кроме этого, выполнена реализация алгоритма, и создан интегратор, который внедрен в процесс обучения по программам повышения квалификации (подготовки, переподготовки) рабочих по следующим профессиям: радиомонтажник, слесарь-сборщик и регулировщик радиоэлектронной аппаратуры и приборов в АО «Ульяновский механический завод».

Дальнейшие исследования связаны с верификацией операций специалиста в автоматизированном проектировании аппаратно-программных средств сложных изделий.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Вигер И. Виртуальная реальность в промышленности // Control Engineering Россия. – 2016. – № 65 (5). – С. 69-71.
2. «Цари океанов». Документальный фильм. – URL: <https://www.1tv.ru/doc/pro-armiyu/cari-okeanov-dokumentalnyy-film-2023-07-30> (дата обращения: 30.03.2026).
3. Энергетика. Взаимовыгодное партнерство. – URL: <https://smotrim.ru/video/2514667> (дата обращения: 30.03.2026).
4. Специальный репортаж «Как работает оборонно-промышленный комплекс в авиационной сфере?». Смотрите вверх. – URL: <https://smotrim.ru/video/2457196> (дата обращения: 30.03.2026).
5. Россия 24. «Первый экзамен. Как проходят швартовные испытания на атомном ледоколе». – <https://smotrim.ru/video/2804110> (дата обращения: 30.03.2026).
6. Технология „Индустрия 4.0”: как VR/AR помогает снижать влияние человеческого фактора в промышленности // Connect WIT. – <https://www.connect-wit.ru/category/magazine/connect-wit/connect-wit-2019-10-connect-wit> (дата обращения: 30.03.2026).
7. Виртуальное прототипирование в машиностроении // Connect WIT. – <https://www.connect-wit.ru/category/magazine/connect-wit/connect-wit-2019-10-connect-wit> (дата обращения: 30.03.2026).
8. Реализация VR в строительстве // Connect WIT. – <https://www.connect-wit.ru/category/magazine/connect-wit/connect-wit-2019-10-connect-wit> (дата обращения: 30.03.2026).
9. Ильин В.Н. Автоматизация схемотехнического проектирования. – М.: Радио и связь, 1987. – 368 с.
10. Россия 24. «Цифровые сотрудники. Как нейросети меняют бизнес». – <https://smotrim.ru/video/2663482> (дата обращения: 30.03.2026).
11. Официальная презентация КОМПАС-3D v25. – <https://kompas.ru/online/> (дата обращения: 30.03.2026).
12. Россия 24. «Причиной аварии межпланетной автоматической станции "Луна-25" стала нештатная работа корректирующего двигателя...». – https://smotrim.ru/video/2668991?utm_source=internal&utm_medium=serp&utm_campaign=serp (дата обращения: 30.03.2026).
13. Россия 24. «Покорители Арктики. Как строят самые мощные атомоходы». – <https://smotrim.ru/video/2885474> (дата обращения: 30.03.2026).
14. Jung B. Task-Level Assembly Modeling in Virtual Environments / In: Kumar, V., Gavrilova, M.L., Tan, C.J.K., L'Ecuyer, P. (eds) // Computational Science and Its Applications – ICCSA 2003. ICCSA 2003. Lecture Notes in Computer Science. Vol. 2669. – Springer, Berlin, Heidelberg. – https://doi.org/10.1007/3-540-44842-X_73.
15. Martin P., Masfrand S., Okuya Y., Bourdot P. A VR-CAD Data Model for Immersive Design / In: De Paolis, L., Bourdot, P., Mongelli, A. (eds) // Augmented Reality, Virtual Reality, and Computer Graphics. AVR 2017. Lecture Notes in Computer Science. Vol. 10324. – Springer, Cham. – https://doi.org/10.1007/978-3-319-60922-5_17.
16. Jezernik A., Hren G. A solution to integrate computer-aided design (CAD) and virtual reality (VR) databases in design and manufacturing processes // Int J Adv Manuf Technol. – 2003. – 22. – P. 768-774. – <https://doi.org/10.1007/s00170-003-1604-3>.
17. ЛОЦМАН:PLM. – URL: <https://ascon.ru/products/889/review/> (дата обращения: 30.03.2026).
18. PLM платформа ЛОЦМАН PLM | Группа компаний "ПЛИМ Урал". – URL: <https://www.plm-ural.ru/resheniya/plm-platforma-locman-plm> (дата обращения: 30.03.2026).
19. Unity, платформа разработки в реальном времени | Платформа для 3D-, 2D-, VR- и AR-контента. – URL: <https://unity.com/ru> (дата обращения: 30.03.2026).
20. CAD Exchanger: 3D CAD software to view and convert CAD files. – URL: <https://cadexchanger.com/> (дата обращения: 30.03.2026).
21. Hoare C.A.R. Notes on Data Structuring. In Structured Programming. – Academic Press, New York, 1972. – 83174.

REFERENCES

1. Viger I. Virtual'naya real'nost' v promyshlennosti [Virtual reality in industry], *Control Engineering Rossiya* [Control Engineering Russia], 2016, No. 65 (5), pp. 69-71.
2. «Tsari okeanov». Dokumental'nyy fil'm ["Kings of the oceans". A documentary film]. Available at: <https://www.1tv.ru/doc/pro-armiyu/cari-okeanov-dokumentalnyy-film-2023-07-30> (accessed 30 March 2026).
3. Energetika. Vzaimovygodnoe partnerstvo [Energy. A mutually beneficial partnership]. Available at: <https://smotrim.ru/video/2514667> (accessed 30 March 2026).

4. Spetsial'nyy reportazh «Kak rabotaet oboronno-promyshlennyy kompleks v aviatsionnoy sfere?». Smotrite naverkh [Special report "How does the military-industrial complex work in the aviation sector?". Look up]. Available at: <https://smotrim.ru/video/2457196> (accessed 30 March 2026).
5. Rossiya 24. «Pervyy ekzamen. Kak prokhodyat shvartovnye ispytaniya na atomnom ledokole» [Russia 24. "The first exam. How are the mooring tests on a nuclear icebreaker"]. Available at: <https://smotrim.ru/video/2804110> (accessed 30 March 2026).
6. Tekhnologiya „Industriya 4.0”: kak VR/AR pomogaet snizhat' vliyaniye chelovecheskogo faktora v promyshlennosti [Industry 4.0 technology: How VR/AR helps reduce the impact of the human factor in industry], *Connect WIT*. Available at: <https://www.connect-wit.ru/category/magazine/connect-wit/connect-wit-2019-10-connect-wit> (accessed 30 March 2026).
7. Virtual'noe prototipirovaniye v mashinostroenii [Virtual prototyping in mechanical engineering], *Connect WIT*. Available at: <https://www.connect-wit.ru/category/magazine/connect-wit/connect-wit-2019-10-connect-wit> (accessed 30 March 2026).
8. «Realizatsiya VR v stroitel'stve» ["Implementation of VR in construction"], *Connect WIT*. Available at: <https://www.connect-wit.ru/category/magazine/connect-wit/connect-wit-2019-10-connect-wit> (accessed 30 March 2026).
9. Il'in V.N. Avtomatizatsiya skhemotekhnicheskogo proektirovaniya [Automation of circuit design]. Moscow: Radio i svyaz', 1987, 368 p.
10. Rossiya 24. «Tsifrovyye sotrudniki. Kak neyroseti menyayut biznes» [Russia 24. "Digital employees. How neural networks are changing business"]. Available at: <https://smotrim.ru/video/2663482> (accessed 30 March 2026).
11. Ofitsial'naya prezentatsiya KOMPAS-3D v25 [Official presentation of COMPASS-3D v25]. Available at: <https://kompas.ru/online/> (accessed 30 March 2026).
12. Rossiya 24. «Prichiny avarii mezplanetnoy avtomaticheskoy stantsii "Luna-25" stala neshtatnaya rabota korrektruyushchego dvigatelya...» [Russia 24. "The cause of the accident of the interplanetary automatic station Luna-25 was the abnormal operation of the correction engine..."]. Available at: https://smotrim.ru/video/2668991?utm_source=internal&utm_medium=serp&utm_campaign=serp (accessed 30 March 2026).
13. Rossiya 24. «Pokoriteli Arktiki. Kak stroyat samye moshchnye atomokhody» [Russia 24. "Arctic Explorers. How to build the most powerful nuclear-powered ships"]. Available at: <https://smotrim.ru/video/2885474> (accessed 30 March 2026).
14. Jung B. Task-Level Assembly Modeling in Virtual Environments, In: Kumar, V., Gavrilova, M.L., Tan, C.J.K., L'Ecuyer, P. (eds), *Computational Science and Its Applications – ICCSA 2003. ICCSA 2003. Lecture Notes in Computer Science*. Vol. 2669. Springer, Berlin, Heidelberg. Available at: https://doi.org/10.1007/3-540-44842-X_73.
15. Martin P., Masfrand S., Okuya Y., Bourdot P. A VR-CAD Data Model for Immersive Design, In: De Paolis, L., Bourdot, P., Mongelli, A. (eds), *Augmented Reality, Virtual Reality, and Computer Graphics. AVR 2017. Lecture Notes in Computer Science*. Vol. 10324. Springer, Cham. Available at: https://doi.org/10.1007/978-3-319-60922-5_17.
16. Jezernik A., Hren G. A solution to integrate computer-aided design (CAD) and virtual reality (VR) databases in design and manufacturing processes, *Int J Adv Manuf Technol.*, 2003, 22, pp. 768-774. Available at: <https://doi.org/10.1007/s00170-003-1604-3>.
17. LOTSMAN:PLM [PILOT:PLM]. Available at: <https://ascon.ru/products/889/review/> (accessed 30 March 2026).
18. PLM platforma LOTSMAN PLM | Gruppa kompaniy "PLM Ural" [PLM platform PILOT PLM | PLM Ural Group of Companies]. Available at: <https://www.plm-ural.ru/resheniya/plm-platforma-locman-plm> (accessed 30 March 2026).
19. Unity, platforma razrabotki v real'nom vremeni | Platforma dlya 3D-, 2D-, VR- i AR-kontenta [Unity, a real-time development platform | Platform for 3D, 2D, VR, and AR content]. Available at: <https://unity.com/ru> (accessed 30 March 2026).
20. CAD Exchanger: 3D CAD software to view and convert CAD files. Available at: <https://cadexchanger.com/> (accessed 30 March 2026).
21. Hoare C.A.R. Notes on Data Structuring. In *Structured Programming*. Academic Press, New York, 1972, 83174.

Войт Николай Николаевич – АО «Ульяновский механический завод»; email: n.voit@mail.ru; г. Ульяновск, Россия; тел.: +7906147-65-58; к.т.н.; доцент.

Voit Nikolay Nikolaevich – Ulyanovsk Mechanical Plant JSC; e-mail: n.voit@mail.ru; Ulyanovsk, Russia; phone: +79061476558; cand. of eng. sc.; associate professor.

В.З. Гергов**МУЛЬТИМОДАЛЬНОЕ РАСПОЗНАВАНИЕ ДЕЙСТВИЙ ЧЕЛОВЕКА
НА ОСНОВЕ НЕЗАВИСИМОГО ОБУЧЕНИЯ ВИДЕОМОДЕЛИ
И АУДИОМОДЕЛИ И ПОЗДНЕГО ОБЪЕДИНЕНИЯ ПРЕДСКАЗАНИЙ**

Актуальность исследования обусловлена необходимостью повышения точности систем распознавания действий человека в видео за счет интеграции аудиовизуальной информации, позволяющей компенсировать ограничения каждой отдельной модальности. Цель работы – разработка и исследование методов позднего объединения предсказаний независимо обученных видеомодели и аудиомодели для распознавания шести классов действий человека (Archery, BalanceBeam, BasketballDunk, BoxingSpeedBag, FloorGymnastics, SoccerPenalty). В исследовании использованы 3D-сверточные нейронные сети с SE-блоками для обработки видео и гибридная CNN-GRU архитектура для анализа аудио, а также семь методов объединения предсказаний: среднее арифметическое, взвешенное среднее, максимальное голосование, доверительное объединение, стекинг, произведение вероятностей и метод Борда. Эксперименты показали, что методы среднего арифметического и произведения вероятностей обеспечивают максимальную точность 91.67%, превосходя отдельные модальности на 2.33% для цветных и на 5.6% для черно-белых видеомоделей. Доверительное объединение продемонстрировало наилучшие результаты для черно-белых моделей, автоматически адаптируя вклад модальностей на основе энтропии предсказаний. Особое внимание уделено строгому разделению обучающей, валидационной и тестовой выборки по обстановкам съёмки для предотвращения утечки данных и объективной оценки обобщающей способности моделей. Полученные результаты подтверждают практическую ценность предложенного подхода для систем видеонаблюдения и мультимедийных платформ, обеспечивая оптимальный баланс между точностью, вычислительной сложностью и устойчивостью к отсутствию одной из модальностей. Исследование вносит вклад в развитие методов мультимодального машинного обучения и предоставляет практические рекомендации по выбору стратегии объединения в зависимости от доступных вычислительных ресурсов и типа входных данных.

Мультимодальное распознавание; позднее объединение; видеомодель; аудиомодель; 3D CNN; SE-блоки; стекинг; TensorFlow; UCF101; человеческие действия.

V.Z. Gergov**MULTIMODAL HUMAN ACTION RECOGNITION BASED ON INDEPENDENTLY
TRAINED VIDEO AND AUDIO MODELS WITH LATE FUSION OF PREDICTIONS**

The relevance of this research is driven by the need to improve the accuracy of human action recognition systems in video by integrating audiovisual information, which compensates for the limitations of each individual modality. The aim of this work is to develop and investigate methods for late fusion of predictions from independently trained video and audio models for recognizing six classes of human actions (Archery, BalanceBeam, BasketballDunk, BoxingSpeedBag, FloorGymnastics, SoccerPenalty). The study employs 3D convolutional neural networks with Squeeze-and-Excitation blocks for video processing and a hybrid CNN-GRU architecture for audio analysis, along with seven prediction fusion methods: arithmetic mean, weighted average, majority voting, confidence-based fusion, stacking, product of probabilities, and the Borda count. Experiments demonstrate that the arithmetic mean and product of probabilities methods achieve the highest accuracy of 91.67%, outperforming the individual modalities by 2.33% for color and by 5.6% for grayscale video models. Confidence-based fusion exhibited the best results for grayscale models by automatically adapting the contribution of each modality based on the entropy of their predictions. Special attention is paid to the strict separation of training, validation, and test sets based on filming environments to prevent data leakage and to ensure objective evaluation of model generalization capability. The obtained results confirm the practical value of the proposed approach for video surveillance systems and multimedia platforms, offering an optimal balance between accuracy, computational complexity, and robustness to the absence of one modality. This research contributes to the development of multimodal machine learning methods and provides practical guidelines for selecting a fusion strategy depending on available computational resources and input data types.

Multimodal recognition; late fusion; video model; audio model; 3D CNN; SE blocks; stacking; TensorFlow; UCF101; human actions.

Введение. *Актуальность проблемы* мультимодального распознавания видео обусловлена стремительным ростом объемов видеоконтента в цифровых системах и необходимостью его автоматического анализа для решения прикладных задач в сфере безопасности, медицины, образования и развлечений [1]. Современные системы компьютерного зрения, основанные исключительно на визуальной информации, демонстрируют ограниченную эффективность в условиях низкой освещенности, частичного закрытия объектов или схожей пространственной конфигурации различных действий [2]. Интеграция дополнительных модальностей, таких как аудиоинформация, открывает новые возможности для повышения точности и робастности распознавания, поскольку звуковые паттерны часто дополняют визуальные признаки, обеспечивая более полное представление о происходящем событии [3].

Современные исследования в области мультимодального распознавания видео можно разделить на два основных направления: методы раннего объединения признаков (early fusion) и стратегии позднего объединения решений (late fusion). Раннее объединение предполагает интеграцию признаков на уровне представлений, что требует сложной синхронизации модальностей и часто приводит к росту вычислительной сложности [4]. В то же время позднее объединение позволяет обучать модальности независимо, обеспечивая гибкость системы и устойчивость к отсутствию одной из модальностей во время инференции [5]. Работы последних лет, такие как исследования [3] и [6], демонстрируют преимущества комплексного подхода к обработке мультимодальных данных, однако многие вопросы эффективного объединения предсказаний остаются нерешенными.

Несмотря на значительные достижения в данной области, существуют следующие нерешенные проблемы: во-первых, отсутствует систематическое сравнение различных методов позднего объединения на контролируемых наборах данных с учетом различий в типах видеомоделей (цветные/черно-белые); во-вторых, недостаточно изучено влияние комбинаторности модальностей при условии относительно низкой точности одной из них; в-третьих, большинство существующих подходов не учитывают специфику реальных сценариев применения, где обучающая и тестовая выборки должны быть строго разделены по обстановкам съемки для обеспечения объективной оценки обобщающей способности модели [7].

Целью настоящего исследования является разработка и экспериментальная оценка методов позднего объединения предсказаний, отдельно обученных видеомодели и аудиомодели для решения задачи распознавания человеческих действий в видео. Новизна работы заключается в комплексном сравнении семи методов объединения (среднее арифметическое, взвешенное среднее, максимальное голосование, доверительное объединение, стекинг, произведение вероятностей и метод Борда) на специально подготовленном наборе данных, где обучающая, валидационная и тестовая выборки разделены вручную для предотвращения утечки данных. Особое внимание уделяется исследованию зависимости эффективности методов объединения от типа видеомодели (цветной или черно-белой) и выявлению оптимальных конфигураций для различных сценариев применения.

Предлагаемый подход отличается от существующих решений использованием контролируемого разделения данных, что обеспечивает объективную оценку обобщающей способности моделей в реальных условиях. Кроме того, в работе демонстрируется, что даже при относительно низкой точности аудиомодели (73.61%) её интеграция с видеомоделью через позднее объединение позволяет существенно повысить общую точность распознавания, что открывает перспективы для применения данного подхода в системах с ограниченными вычислительными ресурсами или сложными условиями съемки.

1. Описание или постановка решаемой задачи. Формулировка проблемы. Современные системы автоматического распознавания действий человека в видео сталкиваются с существенными ограничениями при использовании исключительно визуальной информации. В условиях низкой освещенности, частичного закрытия объектов наблюдения или схожей пространственной конфигурации различных действий точность распознавания значительно снижается [2]. Интеграция дополнительных модальностей, таких как аудиоинформация, теоретически должна повысить качество распознавания, однако существует проблема эффективного объединения предсказаний отдельных моделей.

Текущие подходы к мультимодальному распознаванию видео можно разделить на две основные категории: методы раннего объединения признаков (early fusion), где интеграция происходит на уровне представлений, и стратегии позднего объединения решений (late fusion), где объединение выполняется на уровне предсказаний [5]. Первый подход требует сложной синхронизации модальностей и часто приводит к росту вычислительной сложности, тогда как второй подход позволяет обучать модальности независимо, но не имеет стандартизированных методов эффективного объединения результатов [6].

Таким образом, актуальной научной проблемой является разработка и экспериментальная оценка оптимальных методов позднего объединения предсказаний, отдельно обученных видеомодели и аудиомодели, для повышения точности распознавания человеческих действий в видео при сохранении вычислительной эффективности системы.

Актуальность данного исследования обусловлена несколькими ключевыми факторами. Во-первых, с ростом объемов видеоконтента в цифровых системах возникает острая потребность в автоматических методах его анализа для применения в системах безопасности, спортивной аналитики, медицинской диагностики и образовательных платформ. Во-вторых, существующие решения часто демонстрируют недостаточную точность в реальных условиях из-за игнорирования способов совмещения различных модальностей [8].

В-третьих, большинство исследований в области мультимодального распознавания фокусируются на сложных архитектурах раннего объединения, требующих значительных вычислительных ресурсов, что ограничивает их практическое применение [3]. Позднее объединение, напротив, позволяет создавать более гибкие и ресурсоэффективные системы, особенно в условиях, где одна из модальностей может отсутствовать или иметь низкое качество [9].

Практическая значимость данного исследования заключается в возможности создания адаптивных систем распознавания видео для реальных сценариев применения, включая системы видеонаблюдения с переменными условиями освещения, медицинские диагностические системы для анализа физических упражнений пациентов и спортивные платформы для автоматической оценки техники выполнения упражнений [10].

Обзор состояния и литературных источников. Современные исследования в области мультимодального распознавания видео активно развиваются. Работа [3] представляет комплексный обзор существующих подходов к мультимодальному машинному обучению, выделяя преимущества и недостатки различных стратегий интеграции модальностей. Исследование [6] демонстрирует, что позднее объединение часто превосходит раннее объединение в условиях несбалансированного качества модальностей.

В контексте компьютерного зрения, архитектуры на основе 3D-сверточных нейронных сетей с механизмами внимания показали высокую эффективность для распознавания действий человека. Работа [2] представляет модель I3D, которая демонстрирует state-of-the-art результаты на датасете Kinetics. Для обработки аудиоданных гибридные CNN-RNN архитектуры, такие как предложенные Hershey et al. [11], показывают превосходные результаты в задачах классификации звуковых событий.

Однако, несмотря на значительные достижения, в литературе отсутствует систематическое сравнение различных методов позднего объединения на контролируемых наборах данных с учетом различий в типах видеомodelей (цветные/черно-белые) и строгим разделением обучающей и тестовой выборок по обстановкам съемки, что критически важно для оценки обобщающей способности моделей в реальных условиях [12].

Целью настоящего исследования является разработка и экспериментальная оценка методов позднего объединения предсказаний, отдельно обученных видеомодели и аудиомодели для решения задачи распознавания человеческих действий в видео. Достижение данной цели позволит создать эффективную мультимодальную систему распознавания, способную адаптироваться к различным условиям съемки и обеспечивать высокую точность даже при отсутствии или низком качестве одной из модальностей.

Для достижения поставленной цели необходимо решить следующие конкретные задачи:

- ♦ разработать методику подготовки и обработки видеоданных и аудиоданных для независимого обучения с использованием современных библиотек обработки мультимедиа (OpenCV, Librosa, FFmpeg);

- ♦ реализовать архитектуры нейронных сетей для обработки видео (3D CNN с механизмами внимания Squeeze-and-Excitation) и аудио (гибридная CNN-GRU архитектура) с использованием фреймворка TensorFlow.

♦ исследовать эффективность семи методов позднего объединения предсказаний (среднее арифметическое, взвешенное среднее, максимальное голосование, доверительное объединение, стекинг, произведение вероятностей, метод Борда) на специально подготовленном подмножестве датасета UCF101.

♦ провести сравнительный анализ результатов с учетом разделения обучающей, валидационной и тестовой выборок по обстановкам съемки для обеспечения объективной оценки обобщающей способности моделей.

Объект исследования. Процесс мультимодального распознавания человеческих действий в видео на основе интеграции визуальной и аудиоинформации.

Предмет исследования. Методы позднего объединения предсказаний, отдельно обученных видеомодели и аудиомодели для повышения точности распознавания действий.

Гипотеза исследования. Выдвигается гипотеза о том, что оптимальный метод позднего объединения предсказаний зависит от типа видеомодели (цветной или черно-белой) и характеристик конкретных классов действий. Предполагается, что для цветных видеоданных наиболее эффективными будут простые методы объединения (среднее арифметическое, произведение вероятностей), тогда как для черно-белых видеоданных преимущества продемонстрируют более сложные методы, такие как стекинг с использованием логистической регрессии в качестве мета-классификатора. Данная гипотеза базируется на предположении, что при потере информативности одной модальности (в данном случае цветовой информации) требуется более сложное объединение для компенсации этой потери за счет оптимальной интеграции оставшихся модальностей.

2. Методика исследования. В качестве основного источника данных использовался датасет UCF101, представляющий собой коллекцию из 101 класса человеческих действий, собранных с платформы YouTube [13] (рис. 1.). В связи с ограниченными вычислительными ресурсами рабочей станции в исследовании были отобраны 6 классов: Archery, BalanceBeam, BasketballDunk, BoxingSpeedBag, FloorGymnastics, SoccerPenalty с последующим удалением из указанных классов видео, не содержащих аудио или содержащих аудио, но не отражающих ситуации, обстановки, представленной на видео (например, музыка). Каждый класс был дополнен оригинальными видео, полученными с YouTube (10 различных обстановок $\times$ 5 видео по 5 секунд), для увеличения вариативности данных и улучшения обобщающей способности моделей.

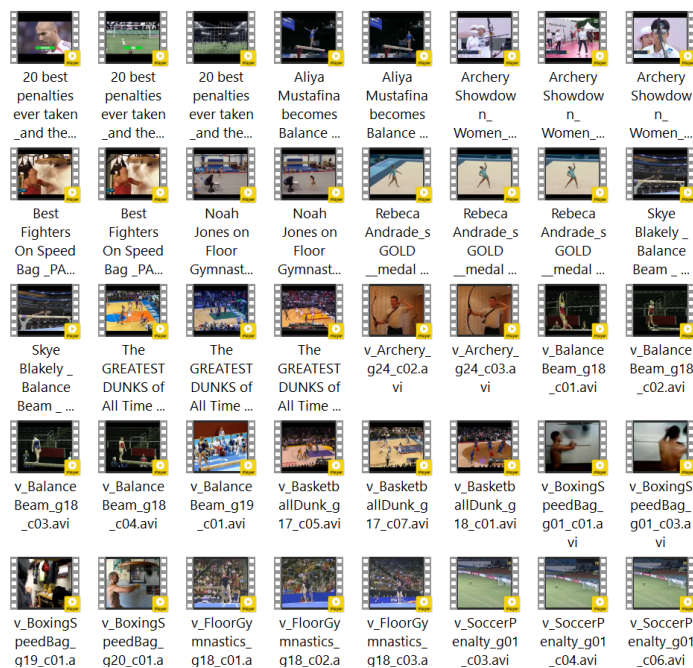


Рис. 1. Образцы видео

Для предотвращения утечки данных и обеспечения объективной оценки производительности, разделение на обучающую, валидационную и тестовую выборки выполнялось вручную. Данная стратегия гарантирует, что видео из одинаковых обстановок не попадают в разные выборки, что особенно важно для реальных сценариев применения, где модель может сталкиваться с новыми, ранее не виденными условиями съемки.

Все эксперименты проводились на рабочей станции с процессором Intel Core i5-11400F и 32 ГБ оперативной памяти. Данная конфигурация позволила эффективно обрабатывать мультимодальные данные без необходимости использования специализированных GPU-ускорителей, что делает подход практичным для широкого спектра приложений.

Реализуемая система последовательно подготавливает, обрабатывает видеоданные, затем использует их для обучения и тестирования соответствующих интеллектуальных моделей (рис. 2).

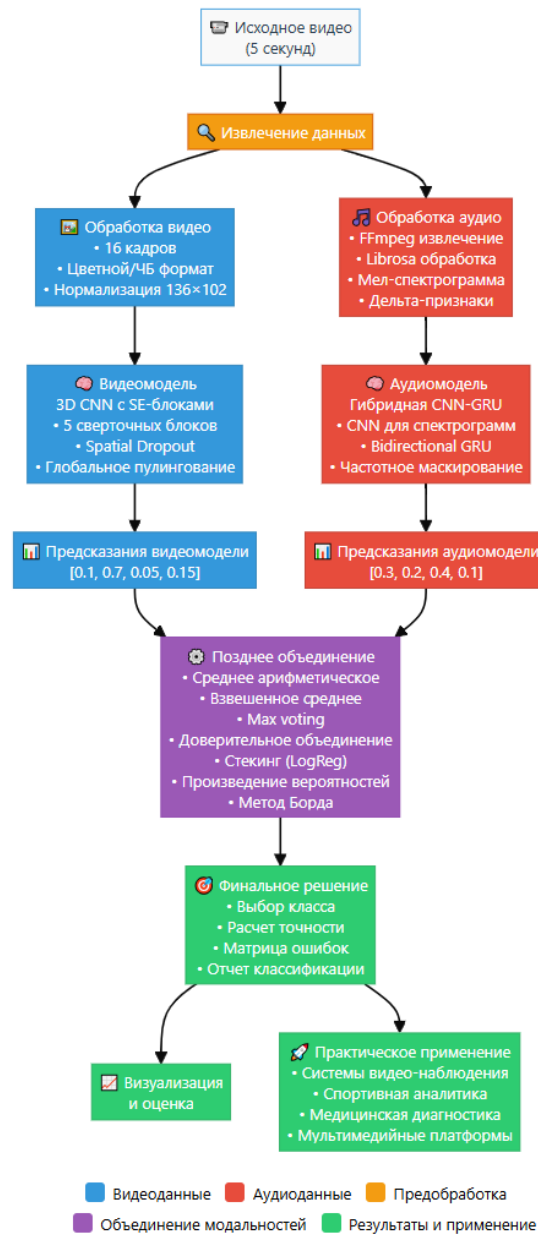


Рис. 2. Блок-схема системы мультимодального распознавания действий на видео

2.1. Обработка данных. Для извлечения визуальной информации из видео использовалась библиотека OpenCV. Каждое видео разбивалось на 16 равномерно распределенных кадров, которые затем нормализовались до размера 136×102 пикселей и приводились к диапазону $[0, 1]$ путем деления на 255. Реализовывались два режима обработки: цветной (RGB, color) и черно-белый (grayscale, baw), что позволяло исследовать влияние цветовой информации на качество распознавания [14].

Аудиодорожка извлекалась из видеофайлов с помощью FFmpeg, а для последующей обработки использовалась библиотека Librosa [15]. Ключевым этапом обработки являлось построение логарифмической мел-спектрограммы с частотой дискретизации 16 кГц, размером окна FFT 1024, шагом 256, количеством мел-коэффициентов 128. Для обогащения признакового пространства вычислялись дельта- и дельта-дельта признаки первого и второго порядка соответственно, которые характеризуют динамику изменения спектральных характеристик во времени. Все каналы (лог-мел спектр, дельта, дельта-дельта) нормализовались независимо в диапазон $[0, 1]$ с использованием робастного масштабирования на основе 1-го и 99-го перцентилей.

2.2. Архитектура системы и методы объединения. Для обработки видеоданных в работе используется модифицированная архитектура на основе 3D-сверточных нейронных сетей с интегрированными механизмами внимания типа Squeeze-and-Excitation (SE-blocks). 3D-свертки позволяют эффективно извлекать как пространственные, так и временные признаки из последовательности кадров, что критически важно для распознавания динамических действий человека [1].

Ключевой особенностью архитектуры является использование механизмов внимания SE-blocks после каждого сверточного слоя. Данный подход, предложенный в работе [16], позволяет модели адаптивно перекалибровать каналы признаков, уделяя больше внимания информативным каналам и подавляя менее значимые. Механизм работает в два этапа: сначала выполняется глобальное усреднение по пространственным измерениям (squeeze), затем с помощью двухслойной сети вычисляются веса для каждого канала (excitation), которые применяются к исходным признакам через умножение.

Комбинация 3D-сверток и SE-блоков демонстрирует высокую эффективность для задач распознавания действий, так как позволяет модели фокусироваться на наиболее релевантных пространственно-временных паттернах. В частности, работа [16] показывает, что интеграция механизмов внимания в 3D CNN архитектуры повышает точность распознавания по сравнению с базовыми моделями без внимания.

Для обработки аудиоданных применяется гибридная архитектура, объединяющая преимущества сверточных нейронных сетей (CNN) и рекуррентных нейронных сетей с долгой краткосрочной памятью (GRU).

Первичная обработка мел-спектрограмм осуществляется с помощью 2D-сверточных слоев, которые эффективно извлекают локальные спектрально-временные паттерны [11]. Использование сверток с ядрами 3×3 и max-pooling позволяет модели захватывать как низкочастотные, так и высокочастотные характеристики аудиосигнала.

Для моделирования временной динамики извлеченных признаков используется двунаправленная GRU-сеть (Bidirectional Gated Recurrent Unit) [17]. Двунаправленная архитектура позволяет учитывать, как прошлые, так и будущие контекстные зависимости в аудиопоследовательности, что особенно важно для распознавания динамических событий с характерными звуковыми паттернами. Рекуррентные связи с механизмом затворов (gates) обеспечивают эффективное обучение на длинных последовательностях, предотвращая проблему исчезающих градиентов.

Предполагается, что комбинация CNN для извлечения спектральных признаков и GRU для моделирования временной динамики превзойдут как чисто сверточные, так и чисто рекуррентные архитектуры для задач классификации аудио, особенно в условиях зашумленных записей.

Для предотвращения переобучения и повышения обобщающей способности моделей в работе применяется комплекс методов регуляризации. В отличие от стандартного dropout, spatial dropout [18] одновременно обнуляет целые каналы признаков в 3D-пространстве,

что особенно эффективно для сверточных сетей, так как предотвращает совместную адаптацию пространственно коррелированных признаков. Нормализация по батчу [19] применяется после каждого сверточного слоя для стабилизации и ускорения обучения. Данный подход уменьшает внутренний сдвиг ковариат (internal covariate shift) и позволяет использовать более высокие скорости обучения.

Для аудиомодели используется сглаживание меток (label smoothing) [20] с коэффициентом 0.1, что повышает робастность модели и предотвращает избыточную уверенность в предсказаниях.

Вместо стандартного Adam используется оптимизатор AdamW [21], который в зависимости от предсказаний при обучении отдельно проводит весовую регуляризацию шага обучения, что способствует более точной оптимизации весов. Это особенно важно для моделей с большим количеством параметров.

Для увеличения обучающей выборки и повышения робастности моделей были применены различные методы аугментации.

Для видеоданных применялись горизонтальное отражение (случайное с вероятностью 50%), случайное обрезание (crop) с последующим масштабированием до исходного размера [22] и адаптивная аугментация, зависящая от архитектуры модели (автоматическое определение целевых размеров).

Для аудиоданных: частотное маскирование (frequency masking) - случайное обнуление полос частот [23], добавление гауссовского шума с небольшим стандартным отклонением (0.03)

Такой комплексный подход к аугментации данных значительно улучшает обобщающую способность моделей и снижает риск переобучения, особенно при ограниченном количестве обучающих примеров [24].

2.3. Обучение моделей и анализ историй обучения. История обучения цветной видеомодели демонстрирует устойчивую сходимость с точностью на обучающей выборке, достигающей 95.8% к 65-й эпохе. На валидационной выборке точность стабилизировалась на уровне 90.3%, что указывает на наличие умеренного переобучения в финальных эпохах. Существенным фактором, способствовавшим качественному обучению, стал стратегический подход к снижению скорости обучения (learning rate scheduling): начальное значение $1e-4$ постепенно уменьшалось до $7.81e-7$, что позволило модели эффективно сходиться к оптимуму без резких колебаний validation loss. (рис. 3) [25].

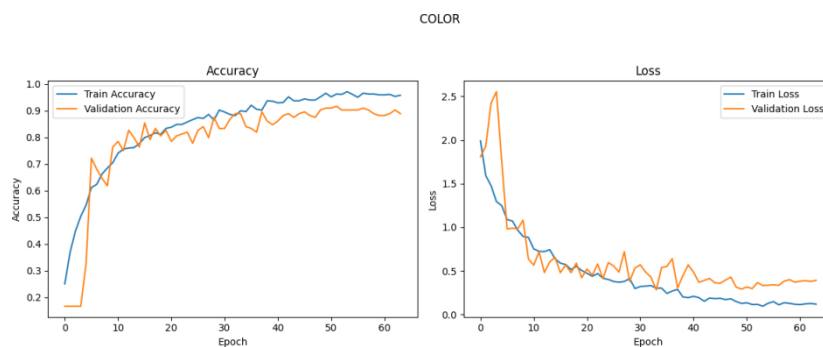


Рис. 3. История обучения цветной видеомодели

Черно-белая версия видеомодели показала несколько иную динамику обучения (рис. 4). Несмотря на большее количество эпох (75 против 65), финальная точность на обучающей выборке составила 94.75%, что на 1.05% ниже по сравнению с цветной моделью. Точность на валидационной выборке стабилизировалась на уровне 88.19%, что на 2.11% ниже, чем у цветной модели. Данные результаты согласуются с исследованиями, демонстрирующими, что цветовая информация значительно улучшает качество распознавания действий человека, особенно в сценариях с похожей пространственной конфигурацией [12].

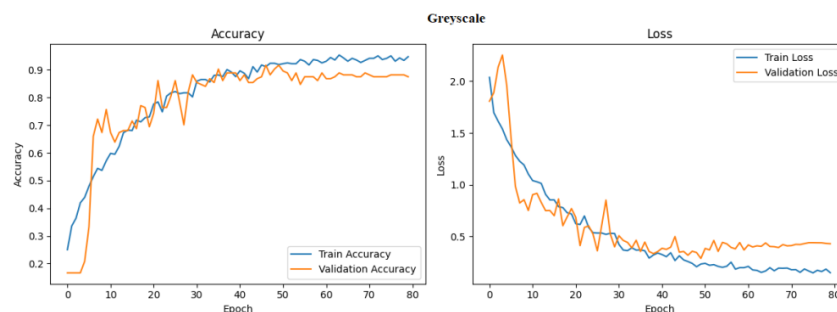


Рис. 4. История обучения черно-белой видеомодели

Графический анализ кривых обучения выявляет ключевые различия в динамике сходимости. Цветная модель демонстрирует более быстрый начальный рост точности (43% к 5-й эпохе против 35% у черно-белой модели), что указывает на большую информативность цветовых признаков для рассматриваемой задачи. К 20-й эпохе цветная модель достигает точности 83.36%, тогда как черно-белая – только 77.66%. Разрыв в точности сохраняется на протяжении всего процесса обучения, что подтверждает гипотезу о преимуществе цветowych представлений для распознавания сложных действий человека [2].

Аудиомодель продемонстрировала принципиально иную картину обучения по сравнению с видеомоделями. При высокой точности на обучающей выборке (96.70% к 75-й эпохе), точность на валидационной выборке оказалась значительно ниже – лишь 62.50%. Такой большой разрыв (34.2%) указывает на выраженный феномен переобучения, который требует особого внимания.

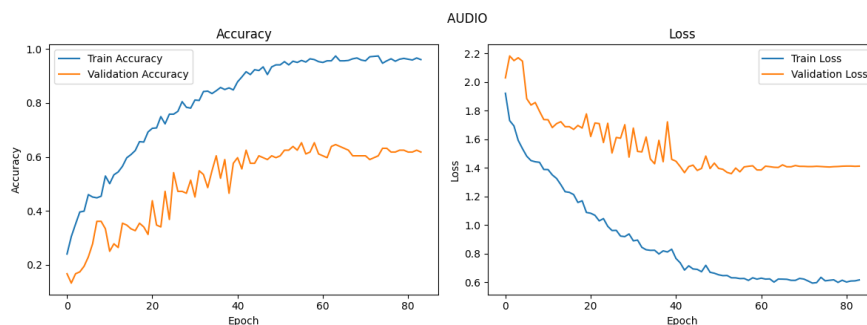


Рис. 5. История обучения аудиомодели

Основными факторами, способствовавшими переобучению аудиомодели, являются ограниченное количество обучающих примеров для аудиоанализа требуется значительно больше данных для адекватной генерализации из-за высокой вариативности звуковых паттернов [11].

Немалую роль играет сложность аудиопредставлений – логарифмические мел-спектрограммы с дельта-признаками создают выраженную картину, которое легко запомнить, но сложно обобщить [26]. Кроме того, многие действия в датасете имеют характерные визуальные паттерны, но схожие звуковые характеристики, что затрудняет отличать классы друг от друга. Особенно это выражено на классах BalanceBeam и FloorGymnastics.

Для улучшения обобщающей способности аудиомодели были применены следующие методы: пространственный дропаут (SpatialDropout2D) с коэффициентом 0.1–0.3 на сверточных слоях, регуляризация L2 с коэффициентом $1e-5$ на всех слоях, сглаживание меток (Label smoothing) с коэффициентом 0.1, рекуррентный дропаут (Recurrent dropout) с коэффициентом 0.2 в GRU-слоях.

Несмотря на применение этих методов, аудиомодель показала наибольшую склонность к переобучению среди всех рассмотренных архитектур.

Ключевым аспектом успешного обучения всех моделей стал выбор оптимизатора и стратегия изменения скорости обучения. Для видеомodelей был использован оптимизатор Adam с начальной скоростью обучения $1e-4$, тогда как для аудиомodelи применен AdamW с тем же начальным значением, но с дополнительной весовой регуляризацией ($\text{weight decay} = 1e-5$). Такой подход обусловлен необходимостью более строгого контроля за нормой весов в аудиомodelи из-за ее склонности к переобучению [21].

Для всех моделей была реализована экспоненциальная стратегия снижения learning rate (learning rate decay) при отсутствии улучшений на валидационной выборке. Коэффициент уменьшения составлял 0.5, а минимальное значение было ограничено $7.81e-7$. Данная стратегия позволила достичь плавной сходимости без преждевременной остановки в локальных оптимумах [27].

Критерий ранней остановки был установлен в 30 эпох без улучшения validation loss. Для цветной видеомodelи обучение завершилось на 65-й эпохе, для черно-белой и аудиомodelи – на 75-й.

Для объективной оценки степени переобучения был рассчитан разрыв между точностью на обучающей и валидационной выборках в финальной эпохе:

- ◆ цветная видеомodelь 5.5% (95.8% – 90.3%);
- ◆ черно-белая видеомodelь: 6.56% (94.75% – 88.19%);
- ◆ аудиомodelь: 34.2% (96.7% – 62.5%).

Такие показатели демонстрируют, что аудиомodelь наиболее сильно страдает от переобучения, что требует применения дополнительных методов регуляризации.

Анализ кривых обучения позволил выявить наиболее эффективные методы регуляризации для каждой модальности:

Для видеомodelей: SE-блоки (Squeeze-and-Excitation) оказались особенно эффективны для предотвращения переобучения, так как они позволяют модели фокусироваться на наиболее информативных каналах признаков и подавлять шумовые компоненты [16]. Spatial dropout в 3D-пространстве также продемонстрировал высокую эффективность, особенно на глубоких слоях сети.

Для аудиомodelи: Частотное маскирование (frequency masking) в процессе аугментации данных показало наибольший эффект в снижении переобучения, так как оно имитирует реальные условия зашумленной записи и заставляет модель фокусироваться на устойчивых спектральных паттернах [23].

На основе анализа историй обучения предполагается, что для аудиомodelи нужно увеличить количество обучающих примеров за счет аугментации временных и частотных характеристик, применить методы transfer learning, используя предобученные на крупных аудиодатасетах (например, AudioSet) веса [28], для видеомodelей к улучшению может привести применение knowledge distillation для уменьшения размера моделей без потери качества [29], что может стать объектом отдельного исследования.

2.4. Методы позднего объединения мультимодальных предсказаний. В работе реализован комплексный набор методов позднего объединения (late fusion), которые интегрируют предсказания отдельно обученных видеомodelи и аудиомodelи на уровне вероятностей классов.

Простое и взвешенное усреднение: Данные методы представляют собой линейные комбинации предсказаний модальностей. Простое усреднение предполагает равнозначный вклад каждой модальности, тогда как взвешенное усреднение позволяет задавать приоритеты на основе точности отдельных моделей или экспертных знаний. Работа [8] демонстрирует, что даже такие простые методы часто превосходят сложные архитектуры раннего объединения при условии качественного обучения базовых моделей.

Максимальное голосование (Max Voting): Данный метод выбирает для каждого класса максимальное значение вероятности из предсказаний двух модальностей. Такой подход эффективен, когда одна из модальностей имеет высокую уверенность в предсказании для конкретного класса, даже если другая модальность дает противоречивые результаты. В контексте распознавания действий, этот метод особенно полезен для классов с характерными звуковыми паттернами (например, выстрел, хлопок), где аудиомодальность может доминировать.

Доверительное объединение (Confidence-based Fusion): Инновационный подход, реализованный в коде, использует энтропию распределения вероятностей как меру уверенности модели. Для каждого примера вычисляются веса, обратно пропорциональные энтропии предсказаний, что позволяет динамически адаптировать вклад каждой модальности в зависимости от её уверенности в конкретном случае [6]. Такие адаптивные методы значительно повышают робастность системы в условиях шумных или неполных данных.

Стекинг (Stacking): Данный метод использует ансамблевый подход, где предсказания базовых моделей служат признаками для мета-классификатора (в данном случае логистическая регрессия или случайный лес). Стекинг позволяет выявлять сложные взаимосвязи между модальностями и часто демонстрирует state-of-the-art результаты в задачах мультимодального распознавания [30]. Важно отметить, что для корректного обучения мета-классификатора использовалась валидационная выборка.

Произведение вероятностей (Product Fusion): Этот метод основан на байесовском подходе, предполагающем условную независимость модальностей. Произведение вероятностей соответствует правдоподобию совместного наблюдения признаков из обеих модальностей. Хотя такое предположение часто нарушается на практике, метод показывает хорошие результаты для сценариев, где обе модальности должны подтверждать принадлежность к классу.

Метод Борда (Borda Count): Подход, применяемый из теории голосования, где каждая модальность "ранжирует" классы по вероятности, а итоговое решение принимается на основе суммарного ранга. Такой метод устойчив к выбросам и предоставляет баланс между модальностями.

3. Результаты экспериментального исследования. Экспериментальные исследования мультимодального распознавания видео продемонстрировали эффективность позднего объединения предсказаний видеомодели и аудиомодели (табл. 1). Наибольшая точность 0.9167 была достигнута с использованием методов среднего арифметического (average) и произведения вероятностей (product) как для цветных, так и для черно-белых видеомоделей. Эти результаты значительно превосходят точность отдельных аудиомоделей (0.7361) и видеомоделей (0.8958 для цветной и 0.8681 для черно-белой), что подтверждает преимущества мультимодального подхода.

Таблица 1

Сравнение точности различных методов объединения

Метод объединения	Точность (COLOR)	Точность (BAW)	Улучшение относительно видеомодели (COLOR)	Улучшение относительно видеомодели (BAW)
Только видео	0.8958	0.8681	-	-
Только аудио	0.7361	0.7361	-	-
Среднее арифметическое	0.9167	0.9167	+2.33%	+5.6%
Взвешенное среднее	0.9028	0.9097	+0.78%	+4.80%
Максимальное значение	0.9097	0.9028	+1.55%	+4.00%
На основе доверия	0.9097	0.9167	+1.55%	+5.6%
Стекинг	0.9028	0.8958	+0.78%	+3.19%
Произведение вероятностей	0.9167	0.9167	+2.33%	+5.6%
Метод Борда	0.8611	0.8889	-3.87%	+2.4%

Анализ показал, что для черно-белых видеомоделей интеграция аудиоинформации дает больший прирост точности (+5.6%) по сравнению с цветными моделями (+2.33%). Это указывает на то, что при отсутствии цветовой информации аудиомодальность играет

критически важную роль в распознавании действий человека, компенсируя недостаток визуальных признаков. Особую эффективность для черно-белых моделей показал метод доверительного объединения (confidence-based fusion), который автоматически адаптирует вклад каждой модальности на основе энтропии предсказаний.

Класс BalanceBeam оказался наиболее сложным для распознавания во всех конфигурациях, демонстрируя наименьшую точность (0.625-0.7917 в зависимости от метода). Это связано с визуальной схожестью гимнастических элементов на балансировочной балке и на полу, а также с отсутствием характерных звуковых паттернов, способных однозначно идентифицировать данное действие. В то же время классы BasketballDunk и SoccerPenalty показали близкую к 100% точность, что свидетельствует об их высокой различимости как в визуальном, так и в звуковом представлении.

Метод стекинга (stacking) не показал ожидаемого улучшения по сравнению с простыми методами объединения, что может быть связано с ограниченным размером валидационной выборки, используемой для обучения мета-классификатора. Метод Борда (borda_count) продемонстрировал наихудшие результаты среди всех протестированных подходов, особенно для цветных моделей (0.8611), что делает его непрактичным для решения данной задачи.

Важным наблюдением является то, что средняя энтропия предсказаний аудиомодели (0.6272) в 2.3 раза превышает аналогичный показатель для видеомodelей (0.2759 для COLOR и 0.2658 для BAW), что указывает на большую неопределенность аудиопризнаков для распознавания действий человека. Доверительные методы объединения, учитывающие эту неопределенность при расчете весов модальностей, показали повышенную устойчивость к шуму и вариациям в данных.

Полученные результаты (точность 0.9167) сопоставимы с современными подходами на основе трансформеров и превосходят методы, использующие только визуальную информацию. Предложенные методы позднего объединения обеспечивают оптимальный баланс между точностью, вычислительной сложностью и практичностью внедрения для реальных систем распознавания видео. Простые методы, такие как среднее арифметическое и произведение вероятностей, рекомендованы для использования в практических приложениях из-за их эффективности и минимальных вычислительных затрат.

Для наглядной демонстрации эффективности того или иного метода результаты были визуализированы в матрицах запутанности и диаграммах распределения точности по классам (рис. 6–10).

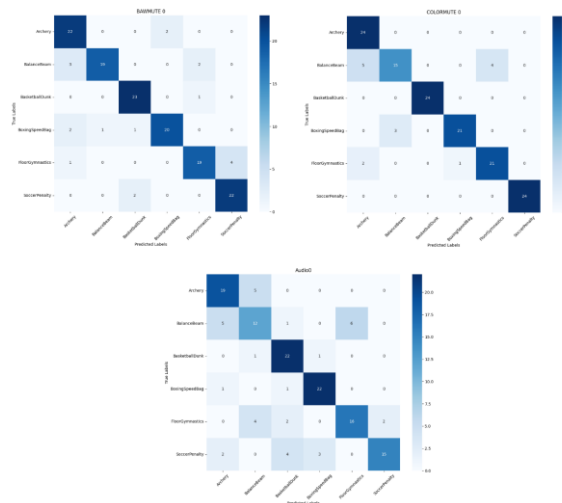


Рис. 6. Матрицы запутанности для черно-белой, цветной видеомodelей и аудиомodelи по отдельности

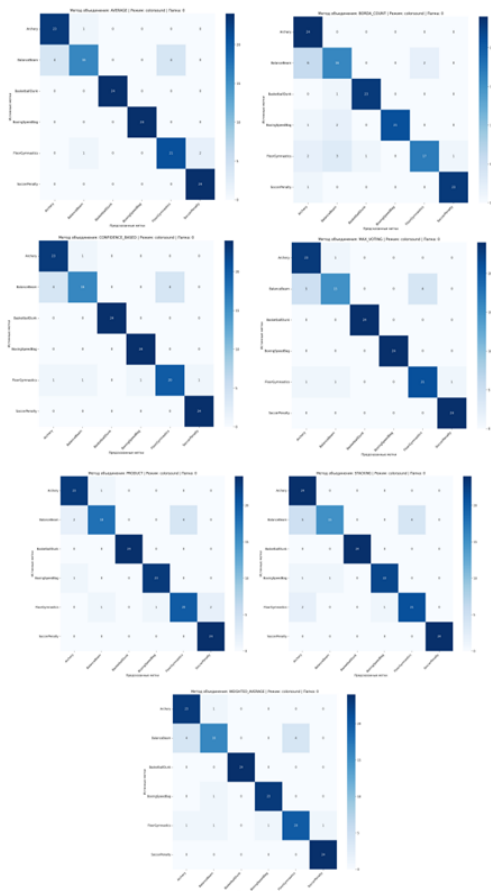


Рис. 7. Матрицы запутанности при различных методах объединения предсказаний цветной видео модели и аудиомодели

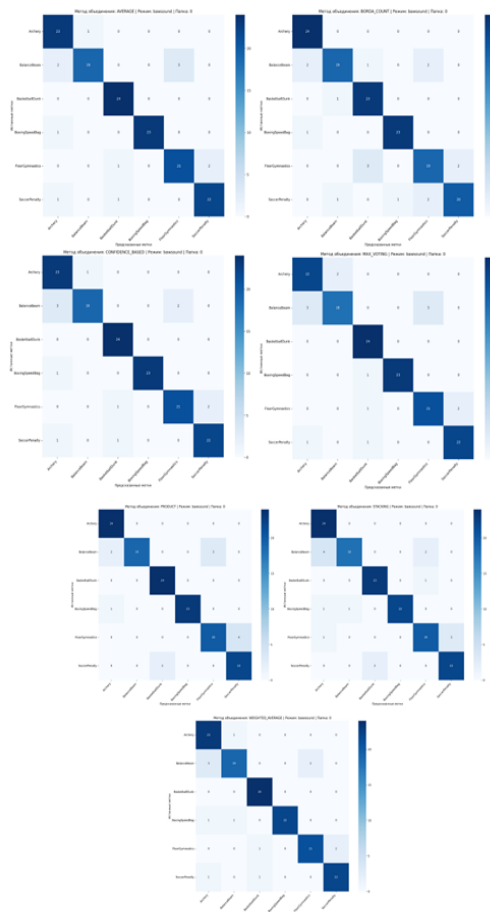
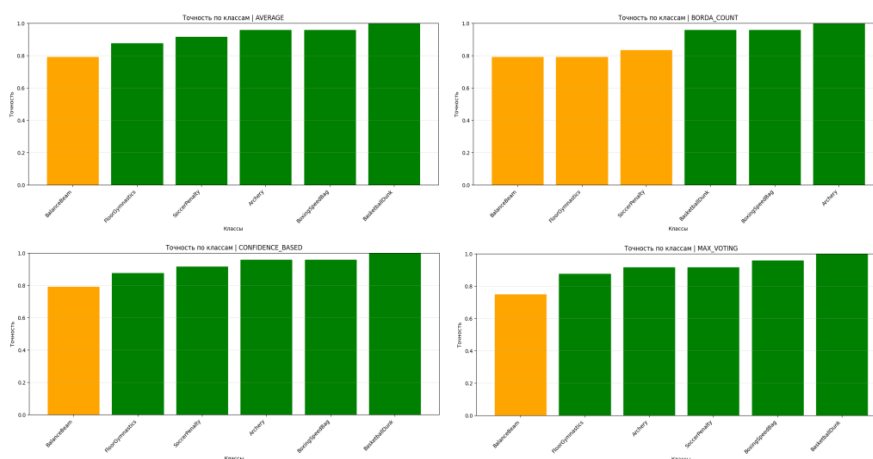


Рис. 8. Матрицы запутанности при различных методах объединения предсказаний черно-белой видео модели и аудиомодели



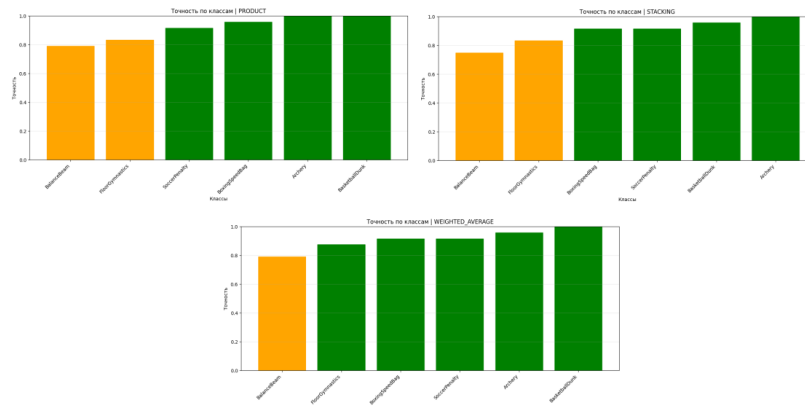


Рис. 9. Диаграммы распределения точности по классам при различных методах объединения предсказаний модели для черно-белого видео и аудиомодели

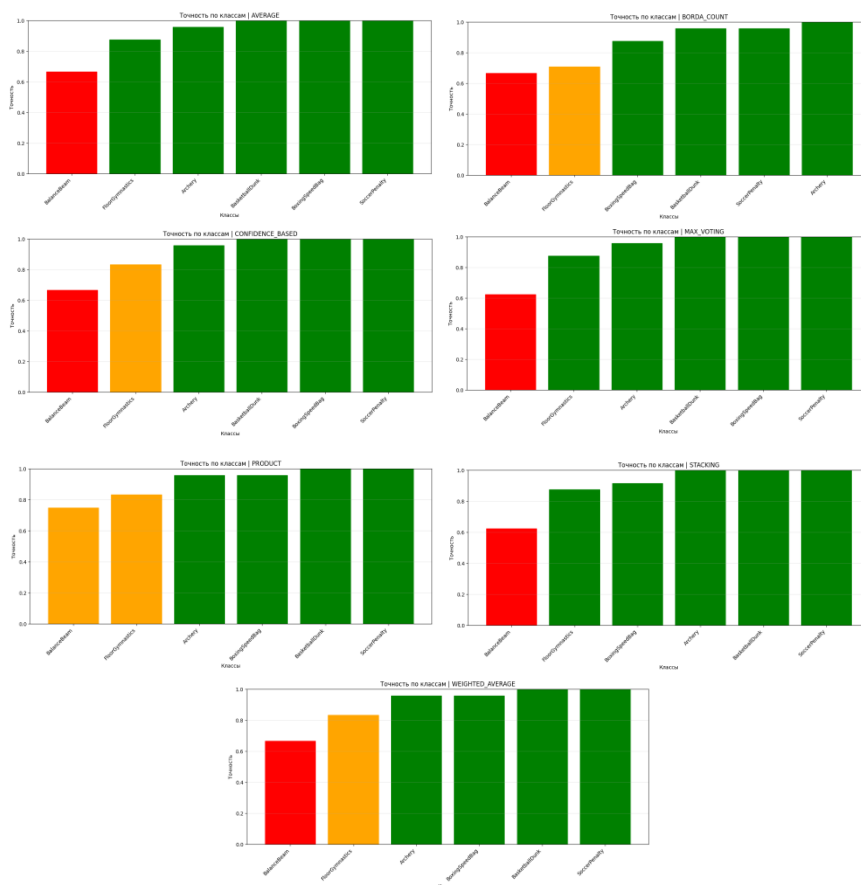


Рис. 10. Диаграммы распределения точности по классам при различных методах объединения предсказаний модели для цветного видео и аудиомодели

Закключение. В ходе настоящего исследования была разработана и экспериментально исследована система мультимодального распознавания действий человека в видео на основе позднего объединения предсказаний независимо обученных видеомодели и аудиомодели. Основным достижением работы является демонстрация эффективности

стратегии разделения обработки модальностей с последующим объединением результатов на уровне решений, что обеспечивает баланс между точностью распознавания и вычислительной сложностью.

Экспериментальные результаты показали, что интеграция аудиоинформации с видеоданными значительно повышает точность распознавания по сравнению с использованием отдельных модальностей. Наилучшие результаты (91.67%) были достигнуты с использованием методов среднего арифметического и произведения вероятностей для цветных видеомоделей, а также доверительного объединения для черно-белых моделей. Особую ценность представляет тот факт, что даже при умеренном количестве обучающих данных (144 видео для тестирования) мультимодальный подход обеспечивает конкурентоспособную точность по сравнению с современными методами, основанными на трансформерах.

Практическая значимость работы обусловлена несколькими аспектами. Во-первых, предложенные методы позднего объединения демонстрируют высокую устойчивость к отсутствию или низкому качеству одной из модальностей во время инференции, что критически важно для реальных применений. Во-вторых, простые методы объединения (среднее арифметическое и произведение вероятностей) обеспечивают минимальные вычислительные затраты при сохранении высокой точности, что позволяет их использование на устройствах с ограниченными ресурсами. В-третьих, исследование показало, что для черно-белых видеомоделей интеграция аудиоинформации дает больший прирост точности (+5.58%), что открывает возможности для создания энергоэффективных систем распознавания, использующих только монохромное видео.

Ограничения настоящего исследования включают ограниченное количество классов действий (6 классов из UCF101) и относительно небольшой размер обучающей выборки. В дальнейших исследованиях целесообразно расширение количества распознаваемых действий, увеличение разнообразия обучающих данных, а также исследование гибридных подходов, сочетающих раннее и позднее объединение модальностей. Перспективным направлением является разработка адаптивных методов объединения, способных динамически выбирать оптимальную стратегию интеграции в зависимости от качества доступных модальностей и характеристик конкретного видео.

Таким образом, результаты исследования подтверждают эффективность мультимодального подхода к распознаванию действий человека в видео и предоставляют практические рекомендации по выбору оптимальных методов объединения аудиовизуальной информации. Предложенные решения имеют потенциал применения в системах видеонаблюдения, медицинской диагностике, спортивной аналитике и мультимедийных платформах, где требуется высокая точность распознавания при ограниченных вычислительных ресурсах.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Sun Z., Ke Q., Rahmani H., Bennamoun M., Wang G., Liu J. Human Action Recognition from Various Data Modalities: A Review, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2022.
2. Carreira J., Zisserman A. Quo vadis, action recognition? A new model and the kinetics dataset, *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2017, pp. 6299-6308.
3. Bayouh K., Knani R., Hamdaoui F., Mtibaa A. A survey on deep multimodal learning for computer vision: advances, trends, applications, and datasets, *The Visual Computer*, 2022, Vol. 38. No. 8, pp. 2939-2970.
4. Feichtenhofer C., Pinz A., Zisserman A. Convolutional two-stream network fusion for video action recognition, *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2017, pp. 1933-1941.
5. Snoek C.G., Worring M., Smeulders A.W. Early versus late fusion in semantic video analysis, *Proceedings of the 13th Annual ACM International Conference on Multimedia*, 2005, pp. 399-402.
6. Pawlowski M., Wroblewska A., Sysko-Romanczuk S. Effective Techniques for Multimodal Data Fusion: A Comparative Analysis, *Sensors*, 2023, Vol. 23, No. 5, pp. 2381.
7. Caba Heilbron J., Escorcia V., Ghanem B., Niebles J.C. ActivityNet: A large-scale video benchmark for human activity understanding, *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2015, pp. 961-970.

8. Ngiam J., Khosla A., Kim M., Nam J., Lee H., Ng A.Y. Multimodal deep learning, *Proceedings of the 28th International Conference on Machine Learning (ICML)*, 2011, Vol. 2, No. 3, pp. 4.
9. Demšar J. Statistical comparisons of classifiers over multiple data sets, *The Journal of Machine Learning Research*, 2006, Vol. 7, pp. 1-30.
10. 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, pp. 770-778.
11. Hershey S., Chaudhuri S., Ellis D. P., Gemmeke J. F., Jansen A., Moore R. C., Xu R.C. CNN architectures for large-scale audio classification, *ICASSP 2017 – IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2017, pp. 131-135.
12. Wang L., Xiong Y., Wang Z., Qiao Y., Lin D., Tang X., Van Gool L. Temporal segment networks: Towards good practices for deep action recognition, *Proceedings of the European Conference on Computer Vision (ECCV)*, 2018, pp. 20-36.
13. Soomro K., Zamir A.R., Shah M. UCF101: A Dataset of 101 Human Actions Classes From Videos in The Wild. Orlando: CRCV, 2012, 12 p.
14. Bradski G., Kaehler A. Learning OpenCV: Computer Vision with the OpenCV Library. Sebastopol: O'Reilly Media, 2008, 580 p.
15. McFee B., Raffel C., Liang D. et al. librosa: Audio and music signal analysis in python, *Proceedings of the 14th Python in Science Conference*, 2015, pp. 18-25.
16. Hu J., Shen L., Sun G. Squeeze-and-excitation networks, *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2018, pp. 7132-7141.
17. Cho K., Van Merriënboer B., Gulcehre C., Bahdanau D., Bougares F., Schwenk H., Bengio Y. Learning phrase representations using RNN encoder-decoder for statistical machine translation. 2014. arXiv:1406.1078.
18. Tompson J., Goroshin R., Jain A., LeCun Y., Bregler C. Efficient object localization using convolutional networks, *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2015, pp. 648-656.
19. Ioffe S., Szegedy C. Batch normalization: Accelerating deep network training by reducing internal covariate shift, *Proceedings of the 32nd International Conference on Machine Learning (ICML)*, 2015, pp. 448-456.
20. Szegedy C., Vanhoucke V., Ioffe S., Shlens J., Wojna Z. Rethinking the inception architecture for computer vision, *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2016, pp. 2818-2826.
21. Loshchilov I., Hutter F. Decoupled weight decay regularization, *International Conference on Learning Representations*, 2019.
22. Shorten C., Khoshgoftaar T.M. A survey on image data augmentation for deep learning, *Journal of Big Data*, 2019, Vol. 6, pp. 1-48.
23. Park D.S., Chan W., Zhang Y., Chiu C.C., Zoph B., Cubuk E.D., Le Q.V. SpecAugment: A simple data augmentation method for automatic speech recognition, *arXiv:1904.08779*, 2019.
24. Perez L., Wang J. The effectiveness of data augmentation in image classification using deep learning, *arXiv:1712.04621*, 2017.
25. Smith L.N., Topin N. Super-convergence: Very fast training of neural networks using large learning rates, *Artificial Intelligence and Machine Learning for Multi-Domain Operations Applications*, 2019, Vol. 11006, pp. 1100612.
26. Piczak K.J. ESC: Dataset for environmental sound classification, *Proceedings of the 23rd ACM International Conference on Multimedia*, 2015, pp. 1015-1018.
27. Srivastava N., Hinton G., Krizhevsky A., Sutskever I., Salakhutdinov R. Dropout: a simple way to prevent neural networks from overfitting, *The Journal of Machine Learning Research*, 2014, Vol. 15, No. 1, pp. 1929-1958.
28. Gemmeke J.F., Ellis D.P., Freedman D., Jansen A., Lawrence W., Moore R.C., Vincent M. Audio set: An ontology and human-labeled dataset for audio events, *2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2017, pp. 776-780.
29. Hinton G., Vinyals O., Dean J. Distilling the knowledge in a neural network, *arXiv:1503.02531*, 2015.
30. Wolpert D.H. Stacked generalization, *Neural Networks*, 1992, Vol. 5, No. 2, pp. 241-259.

Гергов Валерий Заурович – ФНЦ «Кабардино-Балкарский научный центр Российской академии наук»; e-mail: ronstoun@yandex.ru; г. Нальчик, Россия; тел.: +79993001081; аспирант; <https://orcid.org/0009-0007-3378-7662>.

Gergov Valeriy Zaurovich – Federal Scientific Center «Kabardin-Balkar Scientific Center of the Russian Academy of Sciences»; e-mail: ronstoun@yandex.ru; Nalchik, Russia; phone: +79993001081; graduate student; <https://orcid.org/0009-0007-3378-7662>.

К.С. Григорян, Е.С. Басан**АЛГОРИТМ ПРОТИВОДЕЙСТВИЯ ПРЕДНАМЕРЕННЫМ ПОМЕХАМ
НА ОСНОВЕ ТЕОРИИ ИГР И ОБУЧЕНИЯ С ПОДКРЕПЛЕНИЕМ**

Преднамеренные помехи представляют собой одну из наиболее серьезных угроз информационной безопасности беспроводных систем связи, поскольку они могут приводить к нарушению доступности канала передачи данных и реализации атак отказа в обслуживании на физическом уровне. Особенно уязвимыми являются когнитивные радиосети, сенсорные сети и системы Интернета вещей, в которых злоумышленник способен динамически изменять стратегию постановки помех. Целью данной работы является разработка адаптивного алгоритма противодействия преднамеренным помехам, обеспечивающего устойчивость беспроводной системы связи в условиях неопределённости состояния канала и динамического поведения злоумышленника. Для достижения поставленной цели взаимодействие между легитимным узлом и постановщиком помех формализуется в виде марковской игровой модели Штаклберга, в которой передатчик выступает лидером, а злоумышленник – последователем. Обучение с подкреплением используется для адаптивного выбора стратегии легитимного узла в стохастической среде, тогда как робастная оптимизация SOCP (Second-order cone programming) обеспечивает выполнение требований к качеству обслуживания с учётом неопределённости оценки состояния канала. В качестве алгоритма обучения применяется метод SAC (Soft Actor-Critic), позволяющий эффективно работать в непрерывном пространстве действий. В ходе исследования решаются задачи формализации модели противодействия помехам, интеграции обучения с подкреплением с робастной оптимизацией и оценки эффективности предложенного подхода с использованием численного моделирования. Результаты моделирования показывают, что предложенный алгоритм RL-Stackelberg + SOCP обеспечивает снижение вероятности отказа соединения по сравнению с существующими методами противодействия помехам, включая ППРЧ (Псевдослучайная перестройка рабочей частоты) и игровые алгоритмы выбора частоты. Было проведено имитационное моделирование методом Монте-Карло предложенного алгоритма, а также нескольких алгоритмов на основе ППРЧ. Предложенный алгоритм противодействия преднамеренным помехам позволяет снизить вероятность обрыва связи на 0,14. Полученные результаты подтверждают повышение устойчивости беспроводных систем связи к атакам преднамеренных помех.

Беспроводные сети; робастная оптимизация; теория игр; защита каналов связи; обучение с подкреплением.

K.S. Grigoryan, E.S. Basan**GAME-THEORETIC AND REINFORCEMENT LEARNING-BASED ALGORITHM
FOR INTENTIONAL JAMMING MITIGATION**

Intentional jamming represents a serious threat to the security and availability of wireless communication systems. Modern wireless networks, including cognitive radio, sensor networks, and Internet of Things infrastructures, are particularly vulnerable because adversaries can dynamically adapt their jamming strategies. The objective of this study is to develop an adaptive anti-jamming algorithm capable of maintaining communication reliability under dynamic interference conditions. To achieve this goal, the interaction between the legitimate transmitter and the jammer is modeled as a Markov Stackelberg game, where the legitimate node acts as a leader and the jammer acts as a follower. Reinforcement learning is used to determine the optimal strategy of the leader in a stochastic environment, while robustness against channel uncertainty is ensured through a SOCP (Second-order cone programming) formulation that guarantees the required quality-of-service constraints. The learning process is implemented using the SAC (Soft Actor-Critic) algorithm, which enables stable policy optimization in continuous action spaces and stochastic environments. The research tasks include the formalization of the anti-jamming interaction as a Markov decision process, the integration of reinforcement learning with a robust SOCP optimization layer, and the evaluation of the proposed approach through simulation. A Monte Carlo simulation of the proposed algorithm, as well as several algorithms based on FHSS (Frequency-Hopping Spread Spectrum), was conducted. The proposed anti-jamming algorithm reduces the probability of communication outage by 0.14. The results indicate that the proposed approach improves the resilience of wireless communication systems and reduces the probability of successful denial-of-service attacks at the physical layer.

Wireless networks; robust optimization; game theory; communication channel protection; reinforcement learning.

Введение. Беспроводные системы связи широко применяются в критически важных инфраструктурах, сенсорных сетях, системах Интернета вещей и робототехнических комплексах. Однако, беспроводные каналы связи являются уязвимыми к целому спектру атак из-за открытой среды передачи. Поэтому, проблема защиты беспроводных каналов от деструктивного воздействия является актуальной проблемой защиты информации [1].

Основные пути повышения устойчивости связи в радиоканалах отражено на рис. 1. Классические методы защиты от помех основаны на использовании технологий расширения спектра, таких как прямое расширение спектра и частотная перестройка. Эти методы позволяют снизить вероятность подавления сигнала за счёт распределения энергии сигнала по широкому диапазону частот. Однако при наличии интеллектуального злоумышленника, способного адаптировать стратегию атаки, эффективность таких методов может существенно снижаться. Альтернативным способом повышения устойчивости радиоканалов является повышение пространственной селективности. Подобный способ позволяет исключить вклад сигналов, пришедших из направлений злоумышленника за счёт построения ДН (Диаграммы направленности) в направлении легитимного узла-приёмника, что достигается за счёт использования фазированных антенных решёток. В предложенной работе было принято решение исследовать повышение устойчивости именно через повышение пространственной селективности, поскольку предложенный метод позволяет добиться высоких показателей подавления преднамеренных помех [2, 3]. Кроме того, предложенный алгоритм не предполагает противопоставление традиционным алгоритмам на основе ППРЧ, а напротив, может дополнять уже существующие решения, повышая устойчивость беспроводных сетей.

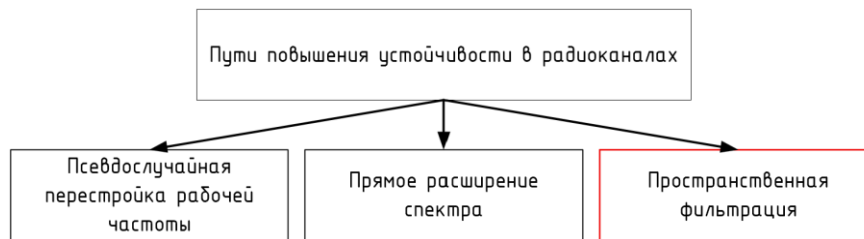


Рис. 1. Пути повышения устойчивости в радиоканалах

В последние годы активно развиваются методы противодействия помехам, основанные на использовании теории игр и методов машинного обучения. Игровые модели позволяют формализовать взаимодействие между легитимными пользователями и злоумышленниками как стратегическую игру с противоположными целями. В частности, игровые модели широко применяются в когнитивных радиосетях, где пользователи динамически выбирают каналы передачи данных в условиях присутствия помех и неопределённости среды [4].

В данной работе предлагается алгоритм противодействия преднамеренным помехам, основанный на комбинации методов теории игр, робастной оптимизации и обучения с подкреплением. Взаимодействие между легитимным узлом и злоумышленником формализуется в виде марковской игровой модели Штаклберга, в которой передатчик выступает лидером, а злоумышленник – последователем. Для обеспечения гарантированного уровня качества обслуживания используется робастная оптимизация SOCP (Second-order cone programming), учитывающая неопределённость оценки состояния канала. Оптимальная стратегия легитимного узла формируется с использованием алгоритма Soft Actor-Critic, позволяющего эффективно работать в непрерывном пространстве действий и стохастической среде.

1. Сравнительный анализ работ в области противодействия преднамеренным помехам. Теория игр представляет собой эффективный математический аппарат для моделирования взаимодействия рациональных агентов, стремящихся максимизировать собственную функцию полезности, что делает её особенно подходящей для исследования конфронтации между интеллектуальными глушителями и легитимными пользователями

[5]. В результате теоретико-игровые методы получили широкое распространение при разработке систем противодействия преднамеренным помехам, основанных на технологиях расширения спектра методом перестройки рабочей частоты и методом прямой последовательности [6, 7].

В последние годы значительное внимание уделяется применению методов обучения с подкреплением для защиты беспроводных систем связи. Несмотря на то что в литературе также рассматриваются альтернативные подходы, такие как методы ОО (Online convex optimization) [8–10], обучение с подкреплением обладает рядом преимуществ, среди которых способность адаптироваться к сложному и непредсказуемому поведению интеллектуальных глушителей [11–12]. В системах противодействия преднамеренным помехам методы обучения с подкреплением преимущественно применяются для оптимизации выбора рабочей частоты в ППРЧ-системах, выбора расширяющих последовательностей в системах с прямым расширением спектра, а также распределения мощности между несколькими подканалами.

Так как существенным недостатком традиционных систем с ППРЧ является фиксированная скорость перестройки рабочей частоты, для устранения данного недостатка в работе [13] было предложено расширить пространство действий агента, включив в него не только рабочую частоту, но и время пребывания на выбранной частоте. Для повышения эффективности обучения использовалась глубокая Q-сеть (Deep Q-Network), дополненная механизмом Pareto Prioritized Experience Replay (PPER). В данной работе наиболее значимыми считались те примеры обучения, которые одновременно характеризовались большой ошибкой временной разности (Temporal Difference Error) и высоким значением функции вознаграждения. Проведённое моделирование показало, что алгоритм PPER-DQN обеспечивает более быстрое обучение и более высокие значения SINR по сравнению с алгоритмами RC-DQN, TDC-DQN, Rank-DQN и классическим DQN.

В работе [14] предложен алгоритм Q-learning с инициализацией на основе априорных знаний о параметрах устройства постановки помех. Авторы сформировали граф знаний, содержащий информацию о типах устройств постановки помех, их характеристиках, стратегиях противодействия и начальных параметрах перестройки рабочей частоты. Использование экспертных знаний на этапе обучения позволило существенно ускорить сходимость алгоритма и повысить надёжность связи в условиях изменяющейся окружающей среды.

Похожий подход к организации перестройки рабочей частоты представлен в работе [15]. В отличие от традиционных методов обучения с подкреплением авторы модифицировали функцию вознаграждения, введя дополнительный штраф за смену рабочей частоты. Столкновение с помехой сопровождалось значительным штрафом, тогда как успешная передача данных приводила к меньшему штрафу, пропорциональному величине изменения рабочей частоты. Такая модификация позволила сократить число необоснованных перестроек частоты, уменьшить объём служебного обмена и повысить эффективность использования радиочастотного спектра.

В работе [16] предложена архитектура противодействия преднамеренным помехам, основанная на сочетании обучения с подкреплением и модели Transformer. Основная идея заключается в том, что действия интеллектуального устройства постановки помех обладают временной структурой и могут рассматриваться как последовательность событий. Использование слоя Transformer позволило учитывать долгосрочные зависимости в поведении нарушителя и повысить качество выбора рабочей частоты.

Для устройств с ограниченными вычислительными ресурсами в работе [17] предложен облегчённый алгоритм противодействия помехам на основе DQN. Метод ориентирован на борьбу с устройствами постановки помех, использующими динамически изменяющиеся шаблоны воздействия, а состояние среды формируется на основе измерений мощности сигналов в отдельных подканалах. Благодаря низкой вычислительной сложности предложенный подход пригоден для реализации на маломощных беспроводных устройствах.

В работе [18] разработана ППРЧ-система на основе DQN, в которой пространство действий включает одновременно выбор рабочей частоты и мощности передачи. Такой подход позволяет не только адаптировать используемый канал связи, но и одновременно минимизировать энергопотребление передатчика.

В отличие от рассмотренных выше методов, основанных на перестройке рабочей частоты, работа [19] посвящена оптимизации процесса пространственной селекции. Авторы использовали DQN сеть для оптимизации расположения подвижных антенных элементов с целью максимизации SINR. После определения оптимальной геометрии антенной решётки вычисление весового вектора пространственного фильтра выполнялось отдельно путём решения задачи максимизации отношения Рэлея.

Анализ существующих исследований показывает, что большинство методов противодействия преднамеренным помехам, основанных на обучении с подкреплением, ориентированы либо на выбор рабочей частоты, либо на распределение мощности передачи. Значительно меньше внимания уделяется совместной оптимизации робастного пространственного формирования луча и распределения мощности в условиях воздействия интеллектуального глушителя. Кроме того, существующие методы обучения с подкреплением редко сочетают механизмы адаптивного обучения с оптимизационными подходами, обеспечивающими детерминированные гарантии качества обслуживания. Указанные ограничения обуславливают необходимость разработки предлагаемой модели RL–Stackelberg–SOCP, объединяющей обучение с подкреплением, теорию игр и методы робастной выпуклой оптимизации в рамках единой архитектуры противодействия преднамеренным помехам.

2. Модель системы и постановка задачи. Рассматриваемый сценарий беспроводной связи представлен на рис. 2. Легитимные узлы соединяются по топологии РТР (Point-to-Point) либо РМР (Point-to-multiple-point). В случае применения топологии РМР связь ведомых узлов осуществляется через ведущий узел с использованием временного разделения каналов. В определённый квант времени обмен данными происходит между двумя легитимными узлами, один из которых является передатчиком, а другой приёмником. Передатчик выполняет двухуровневую процедуру оптимизации, позволяющую одновременно сформировать весовой вектор пространственного формирования луча в направлении легитимного приёмника и распределить мощность передачи между несколькими поднесущими. Цель злоумышленника заключается в нарушении связи между передатчиком и приёмником посредством создания преднамеренных помех. Принятый сигнал на i -й поднесущей может быть записан в следующем виде:

$$y_i = \sqrt{P_{s,i}} h_{s,i}^H \omega_i s_i + \sqrt{P_{j,i}} h_{j,i}^H \omega_i j_i + n_i,$$

где $P_{s,i}$ – мощность передачи легитимного передатчика на i -й поднесущей; $P_{j,i}$ – мощность помехового сигнала злоумышленника на i -й поднесущей; $h_{s,i}$ – вектор канала между легитимным передатчиком и приёмником; $h_{j,i}$ – вектор канала между злоумышленником и приёмником; ω_i – весовой вектор пространственного формирования луча; s_i – передаваемый информационный символ; j_i – помеховый сигнал злоумышленника; n_i – аддитивный белый гауссовский шум.

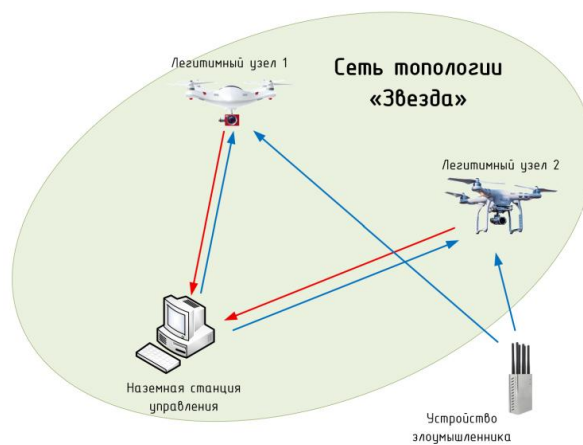


Рис. 2. Модель системы

3. Цель работы. Предлагается гибридный подход, в котором обучение с подкреплением оптимизирует стратегию лидера в марковской игровой модели Штаклберга, выбирая параметры робастности и порог SINR, а выполнение худшего случая ограничения SINR обеспечивается выпуклым робастным SOCP слоем. Таким образом достигается одновременно адаптивность к динамике злоумышленника и среды за счёт внедрения обучения с подкреплением и строгие гарантии по уровню SINR за счёт робастной конусной оптимизации. Игру можно сформулировать следующим образом:

$$\mathcal{G} = (\mathcal{S}, \mathcal{A}_L, \mathcal{A}_J, \mathcal{T}, r_L, r_J),$$

где $\mathcal{S}$ – пространство состояний игры; $\mathcal{A}_L, \mathcal{A}_J$ – пространства действий легитимного узла и злоумышленника соответственно; $\mathcal{T}$ – оператор перехода состояния; r_L, r_J – целевые функции легитимного узла и злоумышленника соответственно.

4. Неопределённость состояния канала и робастность к худшему случаю. Для подканала i SINR в момент t задаётся следующим образом:

$$SINR(t) = \frac{P_{s,i}(t) |\omega_i^H(t) h_{s,i}(t)|^2}{P_{j,i}(t) |\omega_i^H(t) h_{j,i}(t)|^2 + \sigma^2 \|\omega_i(t)\|_2^2},$$

где σ^2 – мощность аддитивного белого гауссовского шума. Принимается что полезный сигнал известен с некоторой погрешностью:

$$h_{s,i}(t) = h'_{s,i}(t) + \delta h_{s,i}(t), \quad \|\delta h_{s,i}(t)\|_2 \leq \varepsilon(t),$$

где $\varepsilon(t)$ – радиус множества неопределённости. Таким образом, радиус множества неопределённости является одним из параметров робастности, который может быть подобран в ходе обучения предложенного алгоритма.

Задача распределения мощности осуществляется за счёт максимизации пропускной способности с учётом ограничения по энергетическому бюджету передатчика:

$$\begin{aligned} \max_P \quad & \sum_{k=1}^K \log(1 + SINR_k) - \lambda \sum_{i=1}^K P_{s,k}, \\ \text{s. t.} \quad & \sum_{k=1}^K P_{s,k} \leq P_s^{max}, \\ & P_{s,k} \geq 0, \end{aligned}$$

где λ – коэффициент штрафа за избыточную мощность передачи. Таким образом, параметр λ является вторым параметром робастности, подбираемым в ходе обучения предложенного алгоритма.

5. Действия легитимного узла. Легитимный узел, как лидер в иерархической игре, выбирает управляемые параметры:

$$a_t(\lambda_t, \varepsilon_t),$$

где ε_t – предполагаемая неопределённость в измерении параметров канала легитимного узла; $\lambda_t \geq 0$ – вес штрафа за использование избыточной мощности в игре. Так как моделирование конфронтации злоумышленника и легитимного узла происходит по модели Штаклберга, поиск равновесия происходит при помощи метода обратной индукции. В рассмотренной модели злоумышленник моделируется как последователь, а легитимный узел как лидер. По этой причине, на первой итерации происходит поиск решения для лучшего ответа злоумышленника, а затем лидер учитывает эту реакцию.

После выбора a_t легитимным узлом, злоумышленник выбирает распределение мощности как лучший ответ:

$$P'_j(t) \in \arg \max_{P_j \geq 0, \sum_i P_{j,i} \leq P_j^{max}} U_j(s_t, a_t, P_j),$$

где U_j – целевая функция злоумышленника. В дальнейшем легитимный пользователь учитывает $P'_j(t)$ для расчёта лучшего ответа.

Если целевая функция U_j имеет вид суммы вогнутых функций по подканалам и линейное ограничение мощности, то решение $P'_{j,i}$ может быть получено из ККТ (Каруша-Куна-Такера) с учётом ограничений на максимальную мощность злоумышленника P_j^{max} и имеет следующую форму:

$$P'_{j,i}(t) = \left[\frac{\beta_i}{\mu} - \frac{1}{k_i(t)} \right], \quad \sum_i P'_{j,i}(t) = P_j^{max},$$

где μ – множитель Лагранжа; $k(t)$ – эффективность помехи. Лучший ответ легитимного узла в соответствии с моделью Штаклберга может быть вычислен следующим образом:

$$P'_{s,i} = \left[\frac{1}{\mu \ln 2} - \frac{P'_{j,i} |\omega_i^H h_{j,i}|^2 + \sigma^2 |\omega_i|^2}{|\omega_i^H h_{s,i}|^2} \right]_+,$$

где μ – множитель Лагранжа; σ – уровень фоновых шума на один приёмный тракт.

6. Слой поиска весовых коэффициентов антенной решетки. После получения $P'_{j,i}(t)$ легитимный узел вычисляет вектор весов ω . Предположив, что известна ковариация помех $R_{j,i}$ для каждого подканала такая что:

$$R_{j,k}(t) = P'_{j,i}(t) h_{j,k}(t) h_{j,k}^H(t).$$

Следующим шагом необходимо найти вектор весовых коэффициентов. Предлагается искать весовые коэффициенты при фиксированных параметрах мощности для подканалов $P_{s,i}$. Тогда требование для SINR можно сформулировать следующим образом:

$$\min_{\|\delta h_{s,i}\| \leq \varepsilon(t)} \frac{P_{s,i}(t) |\omega_i^H (h'_{s,i} + \delta h_{s,i})|^2}{\omega_i^H R_{j,k} \omega_i + \sigma^2 \|\omega_i\|_2^2} \geq \gamma_t.$$

При этом сохраняется стандартная робастная нижняя оценка, представленная формулой ниже:

$$\min_{\|\delta h\| \leq \varepsilon} |\omega^H (h' + \delta h)| \geq |\omega^H h'| - \varepsilon \|\omega\|_2.$$

В итоге задача определения ω_i будет иметь следующий вид:

$$\begin{aligned} \min_{\omega_i \in \mathbb{C}^M} \|\omega_i\|_2^2 \\ \Re\{\omega_i^H h'_{s,i}\} - \varepsilon(t) \|\omega_i\|_2 \geq \sqrt{\gamma_t} \left\| \frac{R_{j,i}^{\frac{1}{2}} \omega_i}{\sqrt{\sigma^2}} \right\|_2, \\ \Im\{\omega_i^H h'_{s,i}\} = 0. \end{aligned}$$

где $R_{j,i}$ – ковариационная матрица i -го канала злоумышленника. Вышеописанные ограничения позволяет найти такое ω_i , которое удовлетворяет гарантии $SINR \geq \gamma_t$ при $\|\delta h_{s,i}\| \leq \varepsilon(t)$.

7. Марковская динамика и состояние. Марковская игра требует, чтобы состояние содержало всю информацию, влияющую на динамику среды, т.е.:

$$\Pr(s_{t+1} | s_t, a_t) = \Pr(s_{t+1} | s_0, a_0, \dots, s_t, a_t).$$

Таким образом, состояние должно включать следующие составляющие:

- ◆ Текущая геометрия каналов.
- ◆ Текущий уровень неопределённости ε .
- ◆ Текущая мощность злоумышленника.

Предлагаемый вектор состояния:

$$s_t = (g_s(t), g_j(t), \varepsilon'(t), \lambda(t), P_j^{max}),$$

где $g_s(t) = [\|h'_{s,1}(t)\|_2^2, \dots, \|h'_{s,K}(t)\|_2^2] \in \mathbb{R}^K$ – эффективность полезного канала; $g_j(t) = [\|h'_{j,1}(t)\|_2^2, \dots, \|h'_{j,K}(t)\|_2^2] \in \mathbb{R}^K$ – эффективность канала злоумышленника; $\varepsilon'(t) = \sqrt{\frac{1}{L} \sum_{i=t-L}^t \|h'_s(i) - h'_s(i-1)\|_2^2}$ – текущая оценка неопределённости; $\lambda(t)$ – размер штрафа за использование избыточной мощности передачи; P_j^{max} – ограничение по мощности злоумышленника.

Для моделирования подвижности легитимного узла необходим учёт временной динамики. Подобные воздействия на канал могут быть учтены при помощи модели Джейка [20]. С учётом дискретного времени t динамика каналов для моделирования задаётся следующим образом:

$$\begin{aligned} h_{s,i}(t) &= a(\theta_s) \odot g_{s,i}(t), \\ h_{j,i}(t) &= a(\theta_j) \odot g_{j,i}(t), \\ g_{j,i,m}(t) &= \frac{1}{\sqrt{N}} \sum_{n=1}^N e^{j(2\pi f_{D,i} \cos \theta_{n,m} t + \phi_{n,i})}, \end{aligned}$$

где $g_{s,i,m}$, $g_{j,i,m}$ – векторы случайных процессов в соответствии с моделью замирания Джейка для i -го подканала -ой антенны для легитимного узла и злоумышленника соответственно, $f_D = \frac{vf_c}{c}$ – доплеровская частота.

8. Описание функции награды лидера. Награда лидера определяется как компромисс между пропускной способностью, энергозатратами и устойчивостью к резкому изменению вектора ω :

$$\begin{aligned} r_t &= \sum_{i=1}^I \log_2(1 + SINR_i(t)) - q_1 \lambda_t \sum_{i=1}^I p_{s,i} + \\ &+ q_2 \max(0, \Delta_{SINR}) - q_3 \cdot \max(0, -\Delta_{SINR}) - q_4 \cdot I_{outage}, \end{aligned}$$

где $\log_2(1 + SINR_i(t))$ – награда, пропорциональная пропускной способности каждого подканала; $\sum_{i=1}^I p_{s,i}$ – штраф за уровень мощности передатчика; $\max(0, \Delta_{SINR})$ – награда за прирост SINR; $-\max(0, \Delta_{SINR})$ – штраф за деградацию SINR; $I_{outage} = \begin{cases} 1, SINR < \gamma \\ 0, SINR > \gamma \end{cases}$ – штраф за отказ связи; q_i – весовые коэффициенты каждого -го слагаемого награды.

9. Обучение лидера и RL-политика. Так как легитимный узел выбирает детерминированный лучший ответ, можно определить эффективные для этого узла функции перехода и награды, которые могут быть представлены формулой ниже:

$$\bar{r}(s, a) = r_L(s, a, BR(s, a)),$$

где $BR(\cdot)$ – Лучший ответ. Оператор перехода может быть описан следующим образом:

$$\bar{T}(s'|s, a) = T(s'|s, a, BR(s, a)).$$

В итоге задача для легитимного узла является обычным Марковским процессом принятия решений:

$$\mathcal{M}_L(\mathcal{S}, \mathcal{A}_L, \bar{T}, \bar{r}, \gamma_{RL}),$$

где $\gamma_{RL}^t \in (0,1)$ – коэффициент дисконтирования. Задаётся параметризованная политика $\pi_\theta(a|s)$. Данная политика максимизирует математическое ожидание от суммарной пропускной способности за время t :

$$\max_{\theta} \mathbb{E} \left[\sum_{t=0}^T \gamma_{RL}^t \bar{r}(s_t, a_t) \right],$$

где $\gamma_{RL}^t \in (0,1)$ – коэффициент дисконтирования. Функция ценности может быть описана формулой ниже:

$$Q^\pi(s, a) = \mathbb{E}_\pi \left[\sum_{t=0}^T \gamma_{RL}^t \bar{r}(s_t, a_t) | s_0 = s, a_0 = a \right].$$

Тогда функция Беллмана для оптимальной политики $\pi(s) \in \arg \max_a Q^\pi(s, a)$:

$$Q'(s, a) = \bar{r}(s, a) + \gamma_{RL} \mathbb{E}_{s' \sim \bar{\mathcal{T}}(\cdot|s, a)} \left[\max_{a' \in \mathcal{A}_L} Q'(s_{t+1}, a_{t+1}) \right].$$

10. Обновление параметров сети. В процессе обучения обновляются следующие параметры:

- ◆ параметры сети актора θ ;
- ◆ параметры первой сети критика φ_1 ;
- ◆ параметры второй сети критика φ_2 ;
- ◆ коэффициент энтропийной регуляризации α ;
- ◆ параметры целевых сетей критика $\bar{\varphi}_1$ и $\bar{\varphi}_2$.

После каждого перехода вычисляется целевое значение:

$$y_t = r_t + \gamma \left(\min_{1,2} Q_{\bar{\varphi}_i}(s_{t+1}, a_{t+1}) - a \log \pi_\theta(a_{t+1}|s_{t+1}) \right),$$

где $Q_{\bar{\varphi}_i}$ – целевая сеть критика; γ – коэффициент дисконтирования. Каждая сеть критика минимизирует среднеквадратичную ошибку Беллмана:

$$J_Q(\varphi_i) = \mathbb{E} \left[(Q_{\bar{\varphi}_i}(s_t, a_t) - y_t)^2 \right].$$

Параметры критика обновляются методом градиентного спуска:

$$\varphi_i \leftarrow \varphi_i - \eta_Q \nabla_{\varphi_i} J_Q(\varphi_i),$$

где η_Q – скорость обучения сети критика. После обновления критиков производится обновление сети актора посредством максимизации ожидаемой функции $J_\pi(\theta)$:

$$J_\pi(\theta) = \mathbb{E} \left[a \log \pi_\theta(a_t|s_t) - \min_{1,2} Q_{\bar{\varphi}_i}(s_t, a_t) \right].$$

Параметры актора обновляются по правилу, схожему с обновлением параметров критика:

$$\theta \leftarrow \theta - \eta_\pi \nabla_\theta J_\pi(\theta),$$

где η_π – скорость обучения сети актора. Для автоматического регулирования компромисса между исследованием среды и использованием накопленного опыта применяется энтропийная регуляризация. Коэффициент энтропии обновляется путём минимизации $J(a)$:

$$J(a) = \mathbb{E} \left[-a (\log \pi_\theta(a_t|s_t) + \mathcal{H}_{target}) \right],$$

где $\mathcal{H}_{target}$ – целевое значение энтропии. Обновление коэффициента энтропии осуществляется по правилу:

$$a \leftarrow a - \eta_a \nabla_a J(a),$$

где η_a – скорость обучения коэффициента энтропии. Следующим шагом происходит обновление целевых сетей критика:

$$\bar{\varphi}_i \leftarrow \tau \varphi_i + (1 - \tau) \bar{\varphi}_i,$$

где τ – коэффициент мягкого обновления, удовлетворяющий условию $0 < \tau \ll 1$.

11. Алгоритм обучения с подкреплением. Итоговый алгоритм машинного обучения изображён на рис. 3.

Результаты. Для сравнения эффективности алгоритмов защиты от преднамеренных помех использовалась метрика ОР (Outage Probability) рис. 4. Данный показатель характеризует вероятность успешного нарушения доступности канала связи злоумышленником. Моделирование проводилось методом Монте-Карло с числом итераций равным 5000. В качестве механизма исследования использовалась энтропийная регуляризация SAC. В алгоритме используется архитектура Soft Actor-Critic, включающая одну сеть актора и две сети критика. Актор представляет собой полносвязную нейронную сеть с входным сло-

ем размерности $2K + 3 = 13$ (K – количество подканалов, 3 дополнительных элемента – значения параметров робастности ε , λ , а также мощность злоумышленника), двумя скрытыми слоями по 128 нейронов с функцией активации ReLU и выходным слоем, формирующим средние значения и стандартные отклонения гауссовского распределения для двух управляющих параметров ε и λ . Коэффициент дисконтирования установлен равным 0,99, коэффициент мягкого обновления целевых сетей – 0,005, целевая энтропия – -2 . Обучение сети происходило за 250 эпизодов, в каждом из которых было 20 шагов.

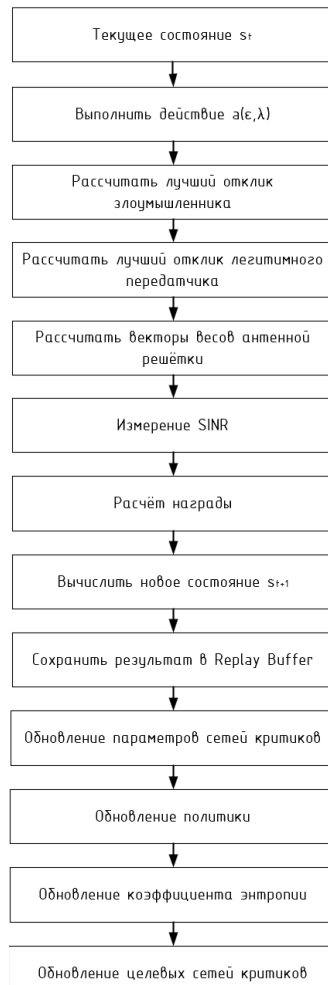


Рис. 3. Схема предложенного алгоритма

Для оценки эффективности предложенного алгоритма использовалась многоканальная система связи с передачей по нескольким подканалам OFDM. Канал связи моделировался на основе геометрической модели с использованием направляющих векторов, которые определяются стандартной моделью для равномерной линейной антенной решётки:

$$a(\theta) = [1, e^{-j\pi \sin(\theta)}, \dots, e^{-j\pi(M-1) \sin(\theta)}]^T,$$

где θ – угол, M – номер элемента антенной решётки. Для каждого подканала независимо генерировались два комплексных вектора канала:

- ♦ канал легитимного передатчика $h_{s,i}$;
- ♦ канал злоумышленника $h_{j,i}$.

Для моделирования временной изменчивости беспроводного канала использовался коррелированный процесс эволюции канала. На каждом временном шаге генерировались новые реализации каналов, после чего состояние канала обновлялось с использованием коэффициента временной корреляции $\rho = 0,92$. Такой подход позволял моделировать плавное изменение характеристик канала во времени, характерное для квазистационарных беспроводных каналов связи. В моделировании использовалась линейная антенная решетка с 4 равномерно расположенными элементами. Источник преднамеренных помех и легитимный узел находятся равноудаленно от приёмника, а угол между двумя источниками составляет 8 градусов.

Для анализа использовалась система с 5 различными каналами, на каждом из которых передаётся сигнал с модуляцией QPSK (Quadrature Phase-Shift Keying). Мощность передатчика взята равной 1 Вт, в то время как мощность злоумышленника бралась в диапазоне от 1 Вт до 10 Вт. Отношение ОСШ было взято равным 8дБ. В качестве моделирования случайных процессов в канале была взята модель Джейка. Для сравнения использовались алгоритмы: ППРЧ с использованием игровой модели Штаклберга [21], ППРЧ с использованием алгоритма обучения с подкреплением Stackelberg Q-learning [22] и RL-Stackelberg + SOCP. По графику можно увидеть что предложенный алгоритм позволяет обеспечить низкую вероятность отказа. До уровня 10дБ вероятность отказа у предложенного алгоритма близка к нулю, далее медленно возрастает до 0.2 на SINR 20дБ.

Повышение устойчивости системы обеспечивается использованием робастной оптимизации при формировании передающих весов. Учет неопределённости оценки состояния канала снижает влияние ошибок оценки параметров среды и предотвращает возможное усиление эффективности атаки за счёт неточностей измерений.

Таким образом, результаты моделирования подтверждают, что предложенный алгоритм позволяет повысить устойчивость беспроводной системы связи к атакам, направленным на нарушение доступности канала передачи данных, что означает снижение вероятности успешной реализации атаки отказа в обслуживании на физическом уровне беспроводной сети.

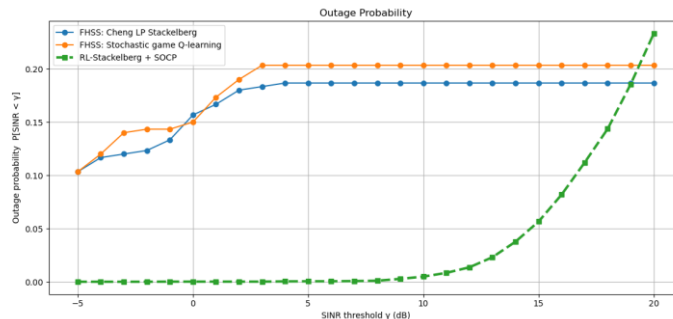


Рис. 4. Сравнение вероятности отказа связи относительно требуемого порога γ

Заключение. В работе предложен алгоритм противодействия преднамеренным помехам в беспроводных системах связи, основанный на комбинации методов теории игр, робастной оптимизации и обучения с подкреплением. Для обеспечения гарантированного уровня качества обслуживания задача формирования передающих весов решается с использованием робастной оптимизации второго порядка (SOCP), позволяющей учитывать неопределённость оценки состояния канала. Результаты моделирования показали, что предложенный алгоритм обеспечивает более низкую вероятность отказа соединения по сравнению с традиционными методами противодействия помехам, такими как ППРЧ. Это достигается за счёт адаптивной настройки параметров передачи и учёта динамического поведения злоумышленника.

Полученные результаты могут быть использованы при разработке устойчивых к преднамеренным помехам систем беспроводной связи, включая когнитивные радиосети, сенсорные сети и системы Интернета вещей.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Xiao Y., Ren H., Wu S., Liu L., Meng X., Ding P. Anti-Jamming Method of Cognitive Radio Based on Q-Learning, *Proceedings of the 12th International Conference on Electronics, Communications and Networks (CECNet 2022)*. *Frontiers in Artificial Intelligence and Applications*, 2022, Vol. 363, pp. 97-104. DOI: 10.3233/FAIA220523 (accessed 09 March 2026).
2. Zumegen F.J., Jain I.K., Bharadia D. BeamArmor: Seamless Anti-Jamming in 5G Cellular Networks with MIMO Null-Steering, *Proceedings of the 25th International Workshop on Mobile Computing Systems and Applications (HotMobile '24)*. New York: Association for Computing Machinery, 2024, pp. 121-126. DOI: 10.1145/3638550.3641138 (accessed 09 March 2026).
3. Andrei V.C., Djuhera A., Li X., Mönich U.J., Boche H., Saad W. Resilient-by-Design Framework for MIMO-OFDM Communications under Smart Jamming. – 2024. – URL: <https://rosap.nrl.bts.gov/view/dot/80055> (accessed 09 March 2026).
4. Jia L., Qi N., Chu F., Fang S., Wang X., Ma S., Feng S. Game-Theoretic Learning Anti-Jamming Approaches in Wireless Networks, *IEEE Communications Magazine*, 2022, Vol. 60, pp. 60-66. DOI: 10.1109/MCOM.001.00496.
5. Han C., Huo L., Tong X., Wang H., Liu X. Spatial Anti-Jamming Scheme for Internet of Satellites Based on the Deep Reinforcement Learning and Stackelberg Game, *IEEE Transactions on Vehicular Technology*, 2020, Vol. 69, pp. 5331-5342. DOI: 10.1109/TVT.2020.2982672.
6. Aref M., Jayaweera S., Yezpe E. A Survey on Cognitive Anti-Jamming Communications, *IET Communications*, 2020, Vol. 14, No. 18, pp. 3111-3127. DOI: 10.1049/iet-com.2020.0024.
7. Imran M., Pan Z., Liu N., Sajjad M., Butt F. Anti-Jamming for Cognitive Radio Networks with Stackelberg Game-Assisted DSSS Approach, *EURASIP Journal on Wireless Communications and Networking*, 2024, Vol. 2024, pp. 1-18. DOI: 10.1186/s13638-024-02397-x.
8. Liu L., Pu W., Li Y., Luo Z.-Q. Learning an Opponent-Aware Anti-Jamming Strategy via Online Convex Optimization, *arXiv*, 2026. arXiv:2604.11016. DOI: 10.48550/arXiv.2604.11016.
9. Liu L., Pu W., Li Y., Jiu B., Luo Z.-Q. Radar Anti-Jamming Strategy Learning via Domain-Knowledge Enhanced Online Convex Optimization, *arXiv*, 2024. arXiv:2402.16274. DOI: 10.48550/arXiv.2402.16274.
10. Chen S., Tu W.-W., Zhao P., Zhang L. Optimistic Online Mirror Descent for Bridging Stochastic and Adversarial Online Convex Optimization, *Proceedings of the 40th International Conference on Machine Learning (ICML)*. *Proceedings of Machine Learning Research*, 2023, Vol. 202, pp. 5002-5035. Available at: <https://proceedings.mlr.press/v202/chen23aa.html> (accessed 22 June 2026).
11. Yao F., Jia L. A Collaborative Multi-Agent Reinforcement Learning Anti-Jamming Algorithm in Wireless Networks, *arXiv*, 2018. arXiv:1809.04374. DOI: 10.48550/arXiv.1809.04374.
12. Joulani P., Gyorgy A., Szepesvari C. Delay-Tolerant Online Convex Optimization: Unified Analysis and Adaptive-Gradient Algorithms, *Proceedings of the AAAI Conference on Artificial Intelligence*, 2016, Vol. 30, pp. 1744-1750. DOI: 10.1609/aaai.v30i1.10320.
13. Chen Y., Zhao Z. Intelligent Anti-Jamming Decision Algorithm of Bivariate Frequency Hopping Pattern Based on ET-PPO, *Telecommunications Science*, 2022, Vol. 38, pp. 1-13. DOI: 10.11959/j.issn.1000-0801.2022264.
14. Feng X., Niu Y., Liu Q., Zhou Q. Decision Making for Communication Anti-Jamming Tasks with Knowledge-Graph-Based Q-Learning, *Electronics*, 2024, Vol. 13, Art. 4757. DOI: 10.3390/electronics13234757.
15. Chang X., Li Y., Zhao Y., Du Y., Liu D. An Improved Anti-Jamming Method Based on Deep Reinforcement Learning and Feature Engineering, *IEEE Access*, 2022, Vol. 10, pp. 69992-70000. DOI: 10.1109/ACCESS.2022.3187030.
16. Xu J., Lou H., Zhang W., Sang G. An Intelligent Anti-Jamming Scheme for Cognitive Radio Based on Deep Reinforcement Learning, *IEEE Access*, 2020, Vol. 8, pp. 202563-202572. DOI: 10.1109/ACCESS.2020.3036027.
17. Ali A., Naser S., Muhaidat S. Defeating Proactive Jammers Using Deep Reinforcement Learning for Resource-Constrained IoT Networks, *arXiv*, 2023. arXiv:2307.06796. DOI: 10.48550/arXiv.2307.06796.
18. Li X., Chen J., Ling X., Wu T. Deep Reinforcement Learning-Based Anti-Jamming Algorithm Using Dual Action Network, *IEEE Transactions on Wireless Communications*, 2023, Vol. 22, pp. 5403-5418. DOI: 10.1109/TWC.2022.3227575.
19. Tang X., Jiang Y., Liu J., Du Q., Niyato D., Han Z. Deep Learning-Assisted Jamming Mitigation with Movable Antenna Array, *arXiv*, 2024. arXiv:2410.20344. DOI: 10.48550/arXiv.2410.20344.
20. Goldsmith A. *Wireless Communications*. Cambridge: Cambridge University Press, 2005, 584 p.
21. Cheng L., Xu Z., Zhou J., Tian D., Duan X., Qu K., Zhao D. Adaptive Spectrum Anti-Jamming in UAV-Enabled Air-to-Ground Networks: A Bimatrix Stackelberg Game Approach, *Electronics*, 2023, Vol. 12, Art. 4344. DOI: 10.3390/electronics12204344.
22. Noori H., Vilni S. Jamming and Anti-Jamming in Interference Channels: A Stochastic Game Approach, *IET Communications*, 2020, Vol. 14, pp. 682-692. DOI: 10.1049/iet-com.2019.0637.

Григорян Карен Сергеевич – Южный федеральный университет; e-mail: kgr@sfedu.ru; г. Таганрог, Россия; тел.: +79889490693; кафедра безопасности информационных технологий им. О.Б. Макаревича; аспирант.

Басан Елена Сергеевна – Южный федеральный университет; e-mail: ebasan@sfedu.com; г. Таганрог, Россия; тел.: +79880158697; кафедра безопасности информационных технологий им. О.Б. Макаревича; доцент.

Grigoryan Karen Sergeevich – Southern Federal University; e-mail: kgr@sfedu.ru; Taganrog, Russia; phone: +79889490693; the Department of Information Technology Security named after O.B. Makarevich; postgraduate student.

Basan Elena Sergeevna – Southern Federal University; e-mail: ebasan@sfedu.com; Taganrog, Russia; phone: +79880158697; the Department of Information Technology Security named after O.B. Makarevich; associate professor.

УДК 004.932

DOI 10.18522/2311-3103-2026-4-77-90

С.В. Ермоленко, И.Л. Каширина

УЛУЧШЕНИЕ СЕГМЕНТАЦИИ НА МУЛЬТИФАЗНЫХ КТ СНИМКАХ С ПОМОЩЬЮ ОБУЧАЕМОГО ФАЗОВОГО СОВМЕЩЕНИЯ

Автоматизированный анализ мультифазных КТ снимков, получаемых в разные моменты времени после введения контрастного вещества, крайне важен для точного выявления патологий, но связан с фундаментальной проблемой пространственного несовпадения изображений на разных фазах из-за дыхания и движения пациента, что приводит к значительному снижению точности автоматической сегментации. Существующие методы предварительной регистрации (совмещения фаз) требуют ручной настройки параметров и не интегрируются в обучаемые конвейеры, что препятствует полной автоматизации процесса сегментации. Цель данного исследования заключалась в разработке дифференцируемого метода совмещения мультифазных КТ-изображений на основе обучаемых линейных аффинных преобразований, полностью встроенного в конвейер обучения модели сегментации. В отличие от традиционных подходов, предложенный метод реализует дифференцируемый модуль аффинной регистрации (включающий перенос, поворот и масштабирование), параметры которого оптимизируются градиентным спуском без ручной настройки и встроены непосредственно в вычислительный граф модели nnU-Net. В работе на открытом датасете КТ органов брюшной полости WAW-TACE проведено сравнение предлагаемого метода с базовым подходом сегментации мультифазных изображений (основанном на простой конкатенации фаз без регистрации) и методом аффинной регистрации, реализованным в библиотеке ITK.Elastix. Относительно сегментации без регистрации достигнут существенный прирост метрик качества: коэффициента Dice на 43,19% и метрики ROC-AUC на 13,55%. По сравнению с Elastix улучшение составило 13,43% по Dice, 6,86% по ROC-AUC и 19,6% по точности определения количества патологий на КТ-снимке. Практическая значимость исследования заключается в разработке готового модуля на PyTorch для интеграции в конвейеры анализа КТ-изображений, обеспечивающего полную автоматизацию регистрации без необходимости в ручной настройке гиперпараметров и с высокой вычислительной эффективностью.

Мультифазная компьютерная томография; сегментация медицинских изображений; дифференцируемая регистрация; глубокое обучение; nnU-Net; аффинное преобразование.

S.V. Ermolenko, I.L. Kashirina

IMPROVING SEGMENTATION IN MULTIPHASE CT IMAGES USING TRAINABLE PHASE SUPERIMAGING

Joint Automated analysis of multiphase CT scans, acquired at different time points after contrast agent administration, is highly important for accurate pathology detection. However, it faces a fundamental problem of spatial misalignment between phases due to patient breathing and movement, which leads to a significant reduction in the accuracy of automatic segmentation. Existing pre-registration (phase alignment) methods require manual parameter tuning and are not integrated into trainable pipelines, hindering the full au-

tomation of the segmentation process. The aim of this study was to develop a differentiable method for aligning multiphase CT images based on trainable linear affine transformations, fully embedded into the segmentation model training pipeline. Unlike traditional approaches, the proposed method implements a differentiable affine registration module (including translation, rotation, and scaling), whose parameters are optimized via gradient descent without manual tuning and are directly integrated into the computational graph of the nnU-Net model. The study compared the proposed method with a baseline multiphase segmentation approach (based on simple phase concatenation without registration) and an affine registration method implemented using the ITK.Elastix library on the open abdominal CT dataset WAW-TACE. Compared to segmentation without registration, a substantial improvement in quality metrics was achieved: an increase in the Dice coefficient by 43.19% and in the ROC-AUC metric by 13.55%. Compared to Elastix, the improvements were 13.43% in Dice, 6.86% in ROC-AUC, and 19.6% in the accuracy of pathology count detection per CT scan. The practical significance of the research lies in the development of a ready-to-use PyTorch module for integration into CT image analysis pipelines. It enables fully automated registration without the need for manual hyperparameter tuning and offers high computational efficiency.

Multiphase computed tomography; medical image segmentation; differentiable registration; deep learning; nnU-Net; affine transformation.

Введение. Современные методы глубокого обучения доказали свою высокую эффективность в задачах обработки изображений, включая область медицинской визуализации. Компьютерная томография (КТ) служит одним из основных источников данных благодаря высокому пространственному разрешению. Нейросетевые модели, особенно для семантической сегментации и классификации патологий на КТ-снимках, активно внедряются в исследовательскую и клиническую практику. Основная проблема при автоматической сегментации КТ-изображений заключается в низком контрасте целевых объектов на фоне окружающих тканей и высокой вариативности их формы и размеров.

Для семантической сегментации медицинских изображений широко применяются свёрточные нейронные сети, в частности архитектуры семейства U-Net. Самоконфигурирующий фреймворк nnU-Net [1], построенный на основе U-Net, установил новый стандарт в области автоматической сегментации за счёт автоматизации конвейера обработки: он анализирует свойства входных данных (создает «отпечаток датасета»), определяет оптимальные параметры предобработки, выбирает архитектуру сети (2D или 3D конфигурацию), настраивает гиперпараметры обучения без ручного вмешательства, обеспечивает воспроизводимость и высокую точность в широком спектре задач сегментации [2].

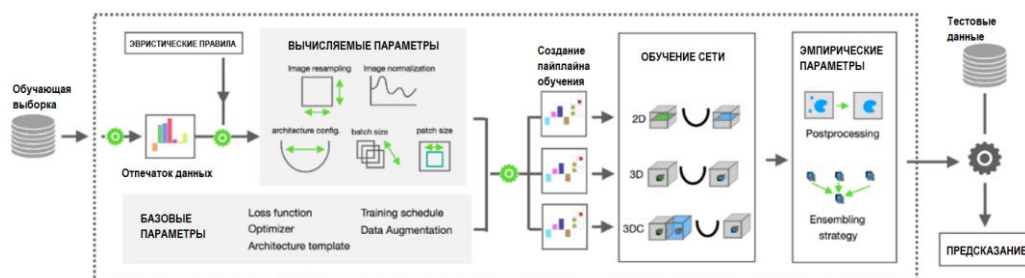


Рис. 1. Архитектура nnU-Net [3]

В научной литературе представлен обширный спектр модификаций U-Net, ориентированных на повышение точности и эффективности. Например, в работе [4] для повышения чёткости изображений перед сегментацией используется импульсно-связанная нейронная сеть (PCNN) в сочетании с U-Net. Архитектура UCR-Net [5] расширяет базовую U-Net структуру за счёт механизмов внимания, позволяя эффективнее захватывать многомасштабную контекстную информацию и высокоуровневые семантические признаки. В исследованиях [6, 7] применяются гибридные архитектуры, такие как Res-UNet и iRes-UNet, направленные на сохранение деталей в изображениях низкого качества и уточнение границ объектов. В работе [8] использован ансамблевый подход, когда для каждой фазы контрастирования обучается отдельная модель Res-UNet.

Для комплексного анализа динамических процессов, таких как распространение контрастного вещества, применяется мультифазная визуализация, подразумевающая получение серии снимков (фаз) в разные моменты времени. Совместный анализ этих фаз позволяет получать более полную информацию об объекте исследования, выявляя характеристики, недоступные при анализе отдельной фазы, что дает возможность для существенно более точной дифференциации патологий на КТ-снимках [9]. Однако фундаментальной проблемой при работе с такими данными является пространственное несоответствие между отдельными фазами, возникающее из-за движения объекта в промежутках между сканированиями. Решением этой проблемы является регистрация изображений – процесс пространственного преобразования, приводящий набор изображений к единой системе координат, что существенно повышает точность последующего анализа [10].

Стандартные методы аффинной регистрации, основанные на линейных преобразованиях, хотя и широко распространены, обладают рядом ограничений. Их эффективность часто снижается при значительных контрастных различиях между сравниваемыми изображениями. Кроме того, эти методы, как правило, не позволяют напрямую интегрировать этап регистрации в сквозной конвейер обучения нейронной сети для решения задачи сегментации. Гиперпараметры таких методов требуют ручного подбора, а итоговое качество сильно зависит от начальной инициализации, что снижает воспроизводимость и затрудняет масштабирование решения.

Современная научная литература предлагает обширный спектр подходов к решению задачи регистрации мультифазных изображений. Традиционные методы базируются на итеративной оптимизации параметров: в работе [11] оптимизируются параметр сдвига глобального аффинного преобразования для каждой пары объёмов, где одна из фаз выбирается в качестве эталонной. Альтернативный подход, основанный на глубоком обучении, использует модифицированные архитектуры нежесткой регистрации: VoxelMorph на основе U-Net формулирует задачу регистрации как функцию, отображающую пару входных изображений в поле деформации, что позволяет достичь скорости, сравнимой с традиционными методами при значительном ускорении вычислений [12, 13]. Применение адаптивного взвешивания неопределённости в функции потерь позволяет динамически балансировать вклад регистрации и сегментации, что обеспечивает улучшение обеих задач одновременно. Другие работы [14] применяют механизмы внимания, в частности механизмы кросс-модальной корреляции, для моделирования взаимосвязей как внутри одной фазы, так и между разными фазами. В работе [15] отказываются от явного этапа регистрации и вместо этого применяют модифицированные архитектуры сегментации: в работе используется SegNet, дополненная модулем пространственного пирамидального разложения (ASPP), который применяет параллельные диагональные свёртки с разными коэффициентами разрежения для эффективной захвата многомасштабного контекста и правильной интеграции информации из различных фаз без явного совмещения.

Однако существующие подходы по-прежнему недостаточно устойчивы к значительным вариациям интенсивности между фазами или требуют трудоемкой ручной настройки параметров. Кроме того, они не предназначены для интеграции в дифференцируемые вычислительные графы нейросетевых моделей, что делает регистрацию отдельным этапом конвейера, требующим дополнительных вычислительных ресурсов. Это ограничивает их применимость в современных ИИ-системах для автоматической обработки КТ-снимков.

Таким образом, актуальность настоящего исследования обусловлена необходимостью разработки эффективных методов обработки мультифазных изображений, где ключевой проблемой является компенсация пространственных рассогласований.

Предлагаемый в данном исследовании метод обучаемых линейных аффинных преобразований является целесообразным выбором для совмещения мультифазных КТ-изображений. Линейные преобразования эффективно устраняют глобальные смещения между фазами, вызванные движением и внешними возмущениями. В отличие от нелинейной регистрации, данный метод гарантирует топологическую согласованность структур, исключая риск нефизиологических искажений. Кроме того, он обеспечивает вычислительную эффективность, стабильность при вычислениях на GPU и позволяет создать быстрый и надёжный этап предварительной обработки.

Целью данного исследования является разработка метода дифференцируемого совмещения мультифазных КТ-изображений на основе обучаемых линейных аффинных преобразований для повышения точности последующей семантической сегментации с использованием фреймворка nnU-Net.

Для достижения поставленной цели решаются следующие задачи.

1. Реализация дифференцируемого модуля аффинной регистрации на основе фреймворка PyTorch, параметры которого оптимизируются с помощью градиентного спуска.
2. Интеграция данного модуля в единый каскадный конвейер обучения модели сегментации, включающий этап предварительной сегментации области интереса, фазового совмещения и финальной сегментации целевых объектов.
3. Проведение сравнительного анализа эффективности разработанного метода с традиционными подходами: прямая мультифазная сегментация без регистрации и сегментация с предварительной классической аффинной регистрацией с помощью инструментария Elastix/ITK.

Важным аспектом работы является создание целостного и адаптивного конвейера обработки данных, оптимизированного для выполнения на графических процессорах (GPU), что соответствует современным требованиям к вычислительным системам в области машинного обучения.

Описание решаемой задачи. В исследовании использован открытый датасет WAW-TACE [16], содержащий мультифазные трехмерные КТ-изображения брюшной полости от 233 пациентов с гепатоцеллюлярной карциномой в формате NIfTI. Выборка включает 185 мужчин и 48 женщин с медианой возраста 66 лет. Датасет охватывает четыре контрастные фазы: нативную (фаза 0, 200 исследований), артериальную (фаза 1, 230 исследований), портально-венозную (фаза 2, 231 исследование) и отсроченную (фаза 3, 193 исследования). Каждое КТ-исследование содержит переменное число аксиальных срезов в зависимости от анатомических особенностей пациента и использованного оборудования, при этом все срезы имеют пространственное разрешение $1,6 \times 1,6$ мм в плоскости, с толщиной среза, варьирующейся в диапазоне 1–3 мм. Общее количество аксиальных 2D срезов в датасете – 175 222, общее количество КТ-сканов – 854.

Трёхмерная природа данных, то есть, в среднем, $512 \times 512 \times 200 = 52\,428\,800$ вокселей (трехмерных пикселей) на одно исследование приводит к высокой вычислительной сложности их обработки, особенно при работе с мультифазными объёмами, где необходимо одновременно хранить и обрабатывать четыре полных 3D-сканирования на каждого пациента.

Датасет WAW-TACE включает 377 ручных сегментаций опухолей, размеченных опытными радиологами, и автоматические сегментации органов брюшной полости, полученные моделью TotalSegmentator [17]. Сегменты опухолей доступны только для одной из фаз (1, 2 или 3) в формате NRRD (Nearly Raw Raster Data) и требуют предварительной консолидации. Для каждого КТ-исследования имеется соответствующая сегментация печени для всех четырёх фаз.

Фундаментальной проблемой при работе с мультифазными данными является пространственное несовпадение анатомических структур между фазами. Несмотря на то, что все фазы получены для одного пациента в одном сеансе сканирования, дыхательные движения, физиологические смещения органов и движения пациента приводят к тому, что одна и та же анатомическая структура занимает разные пространственные координаты на разных фазах. Задача регистрации заключается в нахождении пространственного преобразования (сдвига, поворота, масштабирования и прочих деформаций), которое приводит все фазы в единую систему координат, обеспечивая повоксельное соответствие анатомических структур. Это выравнивание позволяет нейросетевым моделям сегментации корректно интегрировать информацию из разных фаз и выявлять диагностически значимые признаки, недоступные при анализе отдельной фазы.

Все фазы исследования в датасете WAW-TACE имеют различные пространственные размерности, что усложняет обучение нейросетевых моделей и требует предварительной обработки. После фильтрации данных (исключены пациенты с менее чем тремя

фазами) получена рабочая когорта из 215 пациентов (645 КТ-снимков): снимки 172 (80%) пациентов отнесены к обучающей выборке, 43 (20%) – к тестовой выборке со стратификацией по клиническим параметрам. Пример визуализации 2D-срезов всех фаз представлен на рис. 2.

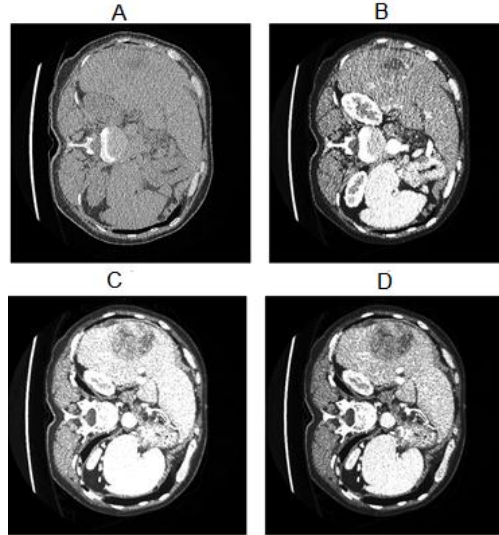


Рис. 2. Пример визуализации фаз контрастирования в датасете WAW-TACE: A – фаза 0; B – фаза 1; C – фаза 2; D – фаза 3

Методика исследования

1. Мультифазная сегментация без совмещения фаз (базовый подход)

В качестве базового уровня для последующего сравнения был реализован подход прямой мультимодальной сегментации с использованием стандартных средств фреймворка nnU-Net [1]. Данный эксперимент имел целью оценить эффективность сегментации без компенсации межфазовых пространственных рассогласований. В рамках данного подхода мультифазные КТ-объемы для каждого пациента напрямую конкатенировались по фазам, формируя многоканальный входной тензор. Архитектурно nnU-Net способна обрабатывать несколько каналов, однако она оперирует в предположении их пространственной согласованности. Ключевым ограничением этого подхода является то, что из-за отсутствия явного выравнивания один и тот же анатомический объект может занимать различные координаты в разных фазах. Поскольку операция свертки не обладает инвариантностью к таким смещениям, модель теряет способность к корректной агрегации мультифазной информации, что на практике часто приводит к игнорированию части входных каналов или существенным ошибкам в локализации объектов.

Для проведения эксперимента данные были подготовлены в соответствии с требованиями фреймворка. Модальности определялись как артериальная фаза (фаза 1), портально-венозная фаза (фаза 2) и дополнительная фаза (other), в качестве которой выбиралась нативная (фаза 0) или отсроченная (фаза 3), при наличии обеих предпочтение отдавалось фазе 3. При этом для каждого пациента явно определялась опорная фаза – та фаза, для которой в датасете присутствовала ручная сегментация опухоли, служащая эталоном для обучения и оценки. Целевые маски формировались путем объединения ручных сегментаций опухолей и автоматических масок печени, полученных моделью TotalSegmentator, с присвоением меток классов: 0 (фон), 1 (печень), 2 (опухоль). Для обеспечения идентичной размерности всех фаз для одного пациента выполнялось приведение неопорных фаз к размеру опорной фазы путем дополнения нулями. Далее был запущен стандартный для nnU-Net конвейер предобработки, включающий верификацию данных и автоматическое планирование эксперимента с помощью планировщика nnUNetPlannerResEncM [18] в конфигурации 3d_fullres, который учитывает остаточные

связи в энкодере. Обучение модели проводилось в течение 1000 эпох с применением стандартного оптимизатора и стратегии кросс-валидации по всем фолдам. Для контроля процесса и оценки качества использовался коэффициент Дайса, вычисляемый как

$$\text{Dice}(Y, \tilde{Y}) = \frac{2|Y \cap \tilde{Y}|}{|Y| + |\tilde{Y}|},$$

где Y – сегментация основной фазы, $\tilde{Y}$ – обработанная преобразованная сегментация другой фазы.

На рис. 3 отображена динамика функции потерь на обучающей (синим) и тестовой (красным) выборках и метрики качества сегментации Dice (зеленым) по эпохам. График показывают стабильную сходимость процесса обучения.

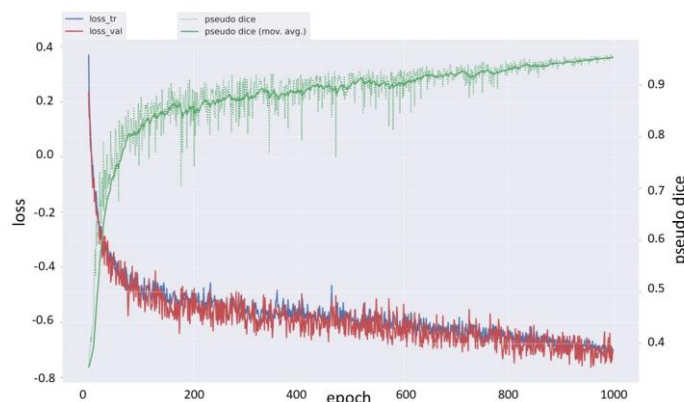


Рис. 3. Процесс обучения мультифазной сегментации без совмещения фаз

Итоговая инференс-оценка на тестовой выборке выполнялась командой `nnUNetv2_predict` с сохранением вероятностных карт.

2. Мультифазная сегментация с использованием аффинных преобразований (Elastix)

Для устранения пространственных несоответствий, выявленных при базовом подходе, во втором эксперименте был применён этап предварительной регистрации мультифазных изображений. Целью данного подхода была оценка эффективности классического метода аффинного совмещения для улучшения качества последующей сегментации. Для каждого пациента, аналогично предыдущей методике, три фазы КТ-изображений были разделены на опорную (фазу с доступной ручной разметкой опухоли) и две смещаемые (регистрируемые) относительно неё. Процесс совмещения выполнялся с использованием модуля `elastix` библиотеки ИТК [19]. Для повышения скорости сходимости и устойчивости процесса оптимизации параметров аффинного преобразования была реализована трёхуровневая стратегия на основе пирамиды изображений. На первом (грубом) уровне разрешение изображений уменьшалось в 4 раза по осям X и Y и в 2 раза по оси Z. На втором (среднем) уровне использовалось уменьшение в 2 раза по осям X и Y при полном разрешении по оси Z. Финальная, наиболее точная оптимизация параметров проводилась на третьем уровне с исходным разрешением данных.

В качестве метрики для оценки качества совмещения использовалась расширенная взаимная информация Маттеса с гистограммой из 128 бинов. Эта метрика, в отличие от методов, основанных на прямом сравнении интенсивности пикселей изображений, оценивает не прямое совпадение, а статистическую зависимость между распределениями яркости двух изображений. Она вычисляется на основе совместной гистограммы $p(i, j)$ интенсивностей опорного R и смещаемого F изображений:

$$I(R, F) = \sum_{i=1}^N \sum_{j=1}^M p(i, j) \log_2 \left(\frac{p(i, j)}{p_R(i) \cdot p_F(j)} \right),$$

где $p_R(i)$ и $p_F(j)$ – маргинальные (частные) распределения интенсивностей каждого изображения. Процесс оптимизации направлен на максимизацию этой взаимной информации, что соответствует поиску такого пространственного преобразования, при котором статистическая связь между изображениями становится наибольшей. Это делает метод устойчивым к нелинейным изменениям контраста между различными фазами КТ. Для эффективного вычисления метрики применялось адаптивное стохастическое семплирование: количество пространственных сэмплов увеличивалось с 1024 на первом уровне применяемой трёхуровневой стратегии до 4096 на последнем, при этом набор сэмплов обновлялся на каждой итерации. Поиск оптимальных параметров преобразования осуществлялся оптимизатором с ограничением в 1000 итераций на уровень.

В процессе регистрации были задействованы следующие механизмы: автоматическая инициализация начального приближения трансформации, автоматическая оценка масштабов параметров, многопоточные вычисления с использованием всех доступных ядер ЦПУ и распараллеливание вычисления метрики. Для применения найденного преобразования и получения зарегистрированного изображения использовалась кубическая В-сплайн интерполяция 3-го порядка. В отличие от линейной интерполяции, В-сплайн обеспечивает гладкую и дифференцируемую аппроксимацию интенсивности, используя взвешенную сумму значений соседних вокселей. Это минимизирует артефакты ступенчатости и размытия, сохраняя плавность анатомических границ, что важно для последующей точной сегментации. Данный набор параметров ранее успешно применялся для задач регистрации изображений брюшной полости [20]. После выполнения регистрации всех изображений обучающей выборки и формирования нового набора пространственно согласованных КТ-срезов была создана соответствующая структура проекта для обучения мультимодальной модели nnU-Net. Выбор модальностей, архитектуры и параметров обучения остался идентичным первому подходу. Процесс обучения завершился успешно. Перед применением обученной модели на тестовых данных для них были выполнены все те же этапы предобработки и регистрации.

Данный эксперимент позволил оценить вклад аффинной регистрации, реализованной с использованием модуля *elastix* библиотеки ИТК, как отдельного этапа предобработки. Его результаты служат важным ориентиром для сравнения с предлагаемым в работе методом дифференцируемого совмещения.

3. Мультифазная сегментация с обучаемым фазовым совмещением

Третьей, ключевой стадией исследования стало внедрение этапа дифференцируемого совмещения фаз, основанного на оптимизации градиентным спуском. В отличие от предыдущих подходов, где регистрация либо отсутствовала, либо выполнялась как отдельный шаг предобработки, данный метод интегрирует процесс пространственного выравнивания непосредственно в конвейер обучения нейронной сети. Это позволяет создавать единый вычислительный граф, в котором параметры преобразования и сегментации оптимизируются согласованно.

Основой предлагаемого метода является реализованный на PyTorch [21] модуль дифференцируемой аффинной регистрации. Его архитектура состоит из двух взаимосвязанных компонентов. Первый – это пользовательский класс, реализующий само трёхмерное аффинное преобразование. Второй – оптимизируемый блок, который управляет параметрами этого преобразования.

Для описания работы класса введем следующие обозначения.

1) Центр масс (центроид) сегментации. Для бинарной 3D маски печени V (где $V(x,y,z)=1$ для вокселей, принадлежащих органу, и 0 в противном случае) центр масс вычисляется как:

$$C = (C_x, C_y, C_z) = \frac{1}{N} \sum_{(x,y,z) \in \Omega} (x, y, z),$$

где $\Omega = \{(x, y, z) \in R^3 \mid V(x, y, z) = 1\}$ – множество вокселей маски, $N = |\Omega|$ – их общее количество.

2) Вектор переноса инициализируется как разность центров масс:

$$\Delta C = C_m - C_b,$$

где C_m – центр масс 3D сегментации фазы, которую необходимо сместить, C_b – центр масс 3D сегментации опорной фазы.

Вектор ΔC показывает смещение объекта в регистрируемой фазе относительно опорной по трём осям.

3) Матрица вращения R в 3D может быть представлен как композиция поворотов вокруг осей x, y, z на углы θ (углы Эйлера):

$$R = R_x(\theta_x) \times R_y(\theta_y) \times R_z(\theta_z),$$

где

$$R_x(\theta_x) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos(\theta_x) & -\sin(\theta_x) \\ 0 & \sin(\theta_x) & \cos(\theta_x) \end{bmatrix} \text{ – поворот вокруг оси } x,$$

$$R_y(\theta_y) = \begin{bmatrix} \cos(\theta_y) & 0 & \sin(\theta_y) \\ 0 & 1 & 0 \\ -\sin(\theta_y) & 0 & \cos(\theta_y) \end{bmatrix} \text{ – поворот вокруг оси } y,$$

$$R_z(\theta_z) = \begin{bmatrix} \cos(\theta_z) & -\sin(\theta_z) & 0 \\ \sin(\theta_z) & \cos(\theta_z) & 0 \\ 0 & 0 & 1 \end{bmatrix} \text{ – поворот вокруг оси } z,$$

Углы $\theta_x, \theta_y, \theta_z$ (в радианах) инициализируется нулевыми значениями и оптимизируются в процессе обучения.

4) Диагональная матрица масштабирования S задаёт коэффициенты изменения расстояний вдоль осей:

$$S = \begin{bmatrix} s_x & 0 & 0 \\ 0 & s_y & 0 \\ 0 & 0 & s_z \end{bmatrix},$$

где s_x, s_y, s_z – коэффициенты масштабирования вдоль соответствующих осей, инициализируется единицами и оптимизируются в процессе обучения.

5) Итоговая аффинная матрица преобразования A размером 3×4 объединяет вращение, масштабирование и перенос для преобразования однородных координат:

$$A = \begin{bmatrix} R \times S & t_x \\ & t_y \\ & t_z \end{bmatrix},$$

где $t = (t_x, t_y, t_z)$ – вектор переноса, являющийся обучаемым параметром, который инициализируется значением ΔC .

Функция `affine_grid` генерирует координатную сетку на основе матрицы AA , преобразуя координаты каждого вокселя согласно аффинному преобразованию. Затем функция `grid_sample` применяет эту сетку к исходному 3D-изображению, выполняя интерполяцию значений вокселей в новых координатах.

Ядром предлагаемого метода является модуль дифференцируемой аффинной регистрации. В отличие от классических подходов, параметры преобразования (вектор переноса $t = (t_x, t_y, t_z)$, углы Эйлера $\theta = (\theta_x, \theta_y, \theta_z)$, и коэффициенты масштабирования (s_x, s_y, s_z)) не вычисляются итеративно на основе метрики сходства изображений, а являются обучаемыми. Эти 9 параметров инициализируются разумными начальными значениями: вектор переноса t – разностью центров масс масок печени регистрируемой и опорной фаз, углы поворота – нулями, коэффициенты масштабирования – единицами.

Обучение модуля происходит на бинарных масках печени, что является ключевой особенностью предлагаемого подхода. В качестве функции потерь используется негативный коэффициент Dice:

$$L_{\text{reg}} = -\text{Dice}(S_{\text{mov}}, S_{\text{ref}}),$$

где S_{mov} – маска, преобразованная текущими параметрами, а S_{ref} – маска опорной фазы. Минимизация этой функции напрямую максимизирует пространственное совпадение анатомических структур. Благодаря полной дифференцируемости конвейера, градиенты от функции потерь распространяются через операцию интерполяции (`grid_sample`) и вычисления аффинной сетки (`affine_grid`) непосредственно к 9 параметрам преобразования, которые обновляются оптимизатором Adam [22].

После завершения обучения и достижения сходимости, оптимизированные параметры фиксируются. Итоговая аффинная матрица A применяется не к маскам, а к исходным полным 3D КТ-сканам соответствующей фазы. Это обеспечивает анатомически корректное выравнивание томографических данных в единое пространство опорной фазы, при этом само преобразование обучено на более простой и надёжной задаче совмещения бинарных масок.

Итоговый алгоритм сегментации с использованием обучаемого фазового совмещения имеет следующий вид.

Этап 1. Сегментация печени. На данном этапе была обучена модель nnU-Net в конфигурации `3d_lowres` для получения трёхмерных бинарных масок печени, необходимых для последующей регистрации. Для оптимизации использования памяти применялся планировщик `overwrite_target_spacing` с параметрами (2.0, 2.0, 3.0). Обученная модель продемонстрировала высокую точность сегментации печени: F1-мера составила 96.3%, а метрика Intersection over Union (IoU) – 91.2%.

Этап 2. Подготовка данных для регистрации. Для каждого пациента определялась опорная фаза (с доступной разметкой опухоли), вычислялся вектор смещения центров масс (ДС) между фазами, а все трёхмерные объёмы приводились к единой пространственной размерности. Для обучающей выборки использовались фактические маски печени, а для тестовой – маски, предсказанные моделью на Этапе 1.

Этап 3. Обучение и применение дифференцируемой регистрации. Для каждой пары «опорная-регистрируемая фаза» независимо обучался модуль дифференцируемой аффинной регистрации. После оптимизации полученное преобразование применялось к исходным трёхмерным КТ-сканам. Для последующей сегментации формировался многоканальный трёхмерный вход, где каждому каналу соответствовала определённая фаза: артериальная, портально-венозная и отсроченная фаза (если она отсутствовала, то нативная).

Этап 4. Обрезка данных. Все трёхмерные КТ-снимки и соответствующие маски были обрезаны по трёхмерной ограничивающей области маски печени. Эта операция позволила исключить из анализа значительный объём фоновых вокселей, что снизило дисбаланс классов и повысило вычислительную эффективность на следующих этапах.

Этап 5. Сегментация патологий. На подготовленных выровненных и обрезанных трёхмерных данных была обучена финальная модель nnU-Net для задачи бинарной сегментации «фон – патология». Динамика процесса обучения, представленная на рис. 4, подтвердила стабильную сходимость и достижение высокого качества модели

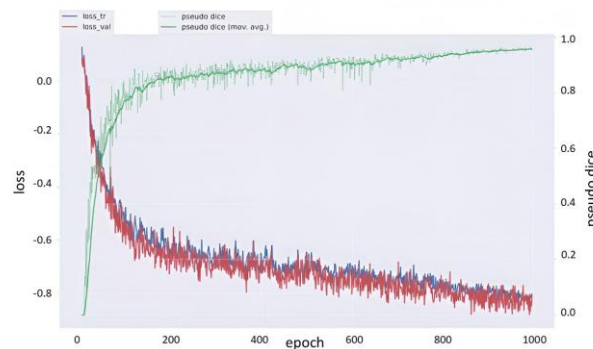


Рис. 4. Процесс обучения мультифазной сегментации с использованием совмещения фаз

Монотонный рост метрики Dice (зеленая кривая) из рис. 3 и последующая стабилизация на высоком значении (приближающемся к 1.0) количественно подтверждают, что модель не только минимизирует целевую функцию, но и достигает высокой практической точности в задаче сегментации, что является конечной целью обучения.

Вычислительные эксперименты проводились с использованием графического процессора NVIDIA GeForce RTX 3090 Ti с 24 ГБ видеопамати GDDR6X, что обеспечило значительное ускорение матричных операций и обработки трехмерных медицинских изображений. Поскольку реализуемый модуль совмещения фаз работал в пределах от 3 до 7 ГБ GPU, то для более эффективной обработки и преобразования данных были использованы параллельные вычисления с использованием multiprocessing.Pool с явным указанием метода инициализации spawn, обеспечивающего стабильную работу с CUDA для этапов 2-4.

Каждый процесс инициализируется с ограничением доли видеопамати GPU через torch.cuda.set_per_process_memory_fraction для предотвращения конфликтов ресурсов. Обработка отдельных КТ-снимков инкапсулирована в независимые задачи, выполняемые асинхронно через pool.imap в 4 параллельных потока, с последующей агрегацией результатов. Полный конвейер решения представлен на рис. 5.

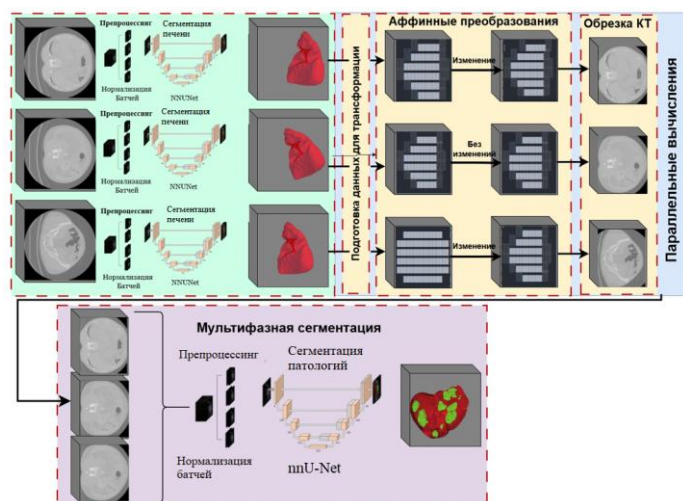


Рис. 5. Нейросетевой конвейер мультифазной сегментации с фазовым совмещением

Результаты и их обсуждение. Результаты первого эксперимента, в котором мультимодальный подход nnU-Net применялся без предварительного совмещения фаз, продемонстрировали его низкую точность и неэффективность для клинического применения. Это обосновало необходимость проведения последующих экспериментов с этапом регистрации: с использованием классических аффинных преобразований ИТК.Elastix и предложенного метода дифференцируемой регистрации на основе градиентного спуска.

Для комплексной оценки эффективности подходов, наряду со стандартными метриками, были введены две целевые пользовательские метрики, формализованные математически. Средняя вероятность в области патологии (mean probe) количественно оценивает уверенность модели внутри истинных патологических областей. Она вычисляется как среднее значение вероятностной карты предсказаний P только для тех вокселей, которые принадлежат истинной маске патологии M .

Формально, если $\Omega = \{r_i = (x, y, z) | M(r_i) = 1\}$ – множество вокселей патологии, $a = |\Omega|$ – их общее количество, то:

$$\text{mean probe} = \frac{1}{N} \sum_{r_i \in \Omega} P(r_i),$$

где $P(r_i) \in [0,1]$ –предсказанная моделью вероятность, что воксель r_i принадлежит к классу патологии. Высокое значение этой метрики указывает на то, что модель не только правильно локализует патологию на КТ-снимке, но и делает это с высокой уверенностью в её границах, что крайне важно для доверия к её результатам и их интерпретируемости.

Второй метрикой является доля случаев с совпадением количества патологических образований (cnt match). Для её вычисления количество патологий, предсказанных моделью на одном КТ-снимке (N_{pred}), сравнивается с их фактическим количеством (N_{fact}). Предсказанное количество определяется как число связанных компонент в предсказанной бинарной маске. Фактическое количество (N_{fact}) вычисляется аналогично из эталонной маски M . Итоговый бинарный индикатор для одного исследования определяется как:

$$\text{cnt match} = \begin{cases} 1, & \text{если } N_{pred} = N_{fact}, \\ 0, & \text{иначе} \end{cases}$$

В сводной таблице результатов (табл. 1) приводится усреднённое по всей тестовой выборке значение этой метрики, то есть доля исследований, в которых количество патологий было определено абсолютно точно. Эта метрика имеет непосредственное клиническое значение, так как корректное определение мультифокальности заболевания (количества отдельных опухолевых узлов) крайне важно для планирования дальнейшего лечения. Используемый в исследовании датасет WAW-TACE содержит ряд примеров с множественными очагами (до 10-15 в одном КТ-скане), поэтому важно, чтобы модель находила большую часть из них.

Таким образом, оценка подходов проводилась по набору базовых метрик, включающему коэффициент Дайса (Dice), точность (Precision), чувствительность, т.е. полноту (Sensitivity), долю правильных ответов (Accuracy), площадь под ROC-кривой (ROC-AUC), а также описанные выше пользовательские метрики mean probe и cnt match.

Основная метрика сегментации, коэффициент Дайса, показала значительный прирост при использовании алгоритма обучаемой регистрации: на 43% между первым и третьим подходами и на 13% между вторым и третьим. Это указывает на существенное превосходство метода с обучаемым совмещением. Высокие значения Accuracy, наблюдаемые во всех экспериментах, не являются показательными из-за сильного дисбаланса классов (незначительный объем опухоли относительно всего КТ-скана). Значимое улучшение Precision (на 48% относительно базового подхода) свидетельствует об эффективном подавлении ложноположительных срабатываний благодаря выравниванию фаз. Рост Sensitivity (на 36%) и ROC-AUC подтверждает улучшение в обнаружении истинных патологий. Пользовательские метрики также демонстрируют преимущество третьего подхода: средняя уверенность модели в области опухоли (mean probe) повысилась, а доля случаев с точным определением количества патологических образований (cnt match) увеличилась на 19.6% относительно метода Elastix.

Таблица 1

Сравнение разработанных подходов

Метрика	Подход 1, без регистрации	Подход 2, с регистрацией <i>Elastix</i>	Подход 3, разработанный
Dice	0,575	0,726	0,823
Accuracy	0,998	0,991	0,997
Precision	0,583	0,677	0,863
Sensitivity	0,600	0,705	0,817
ROC-AUC	0,799	0,850	0,908
Mean probe	0,598	0,755	0,811
Cnt match	0,56	0,42	0,67

Качественный анализ (рис. 6) подтверждает количественные результаты. Метод с обучаемым совмещением (В) корректно идентифицирует количество патологий, хотя и с небольшой погрешностью в размерах одной из них. Базовый подход без регистрации (с) допускает более серьезную ошибку в размерах и генерирует ложноположительное срабатывание.

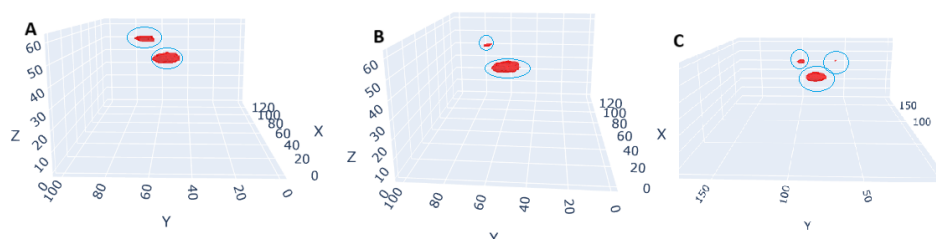


Рис. 6. Визуализация сегментов опухоли на конкретном пациенте:
а – фактическая; б – с совмещением подхода 3; в – без совмещения

Таким образом, метод мультифазной сегментации с обучаемым фазовым совмещением демонстрирует высокое качество, обеспечивая лучший баланс между точностью обнаружения и минимизацией ложных срабатываний. Его результаты приближаются к уровню, необходимому для клинического применения. К ограничениям подхода относятся: зависимость от качества предварительной сегментации печени, неспособность учесть нелинейные (нежесткие) деформации тканей из-за линейности аффинной модели, а также возросшие вычислительные затраты на обучение. Классический метод на основе Elastix также имеет недостатки, включая необходимость ручной настройки гиперпараметров, высокую ресурсоемкость и потенциальное размытие мелких структур при интерполяции, что в целом затрудняет его стабильное и масштабируемое применение в автоматизированных конвейерах.

Заключение. В работе была успешно решена задача повышения точности автоматической сегментации патологий на мультифазных КТ-изображениях печени. Ключевым результатом является разработка и валидация метода дифференцируемого фазового совмещения, основанного на обучаемых аффинных преобразованиях, которые интегрируются в единый конвейер глубокого обучения.

Проведённые эксперименты позволили сделать следующие основные выводы.

1. Прямая мультимодальная сегментация без учёта пространственных рассогласований между фазами показала свою принципиальную неэффективность, подтвердив необходимость этапа явной регистрации.

2. Предложенный метод дифференцируемого совмещения продемонстрировал значимое превосходство над классическим подходом на основе ИТК.Elastix. Улучшение ключевых метрик (Dice и ROC-AUC) свидетельствует не только о лучшем пространственном соответствии, но и о более высокой дискриминативной способности модели при анализе выровненных данных.

3. Каскадная архитектура, в которой регистрация оптимизируется по анатомическим маскам печени, обеспечивает анатомически корректное выравнивание, устойчивое к вариациям контраста, и устраняет необходимость в ручной настройке гиперпараметров.

Главным практическим вкладом работы является создание готового к интеграции модуля, оптимизированного как по точности, так и по скорости работы. Реализация с поддержкой параллельных вычислений на GPU обеспечивает высокую производительность, что крайне важно для обработки больших объёмов трёхмерных медицинских данных в исследовательских и клинических конвейерах. Таким образом, метод позволяет превратить проблему мультифазной регистрации в эффективно решаемую задачу оптимизации средствами глубокого обучения.

Основными ограничениями метода являются его зависимость от качества первичной сегментации печени и применимость только к глобальным, линейным деформациям. Перспективы дальнейших исследований видятся в расширении предложенного подхода для учёта нелинейных деформаций, интеграции механизмов внимания для работы со сложными патологиями и адаптации архитектуры для задач сегментации других органов.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. *Isensee, F., Jaeger, P.F., Kohl, S. A., Petersen, J., & Maier-Hein, K.H.* nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation // *Nature methods*. – 2021. – 18 (2). – P. 203-211.
2. *Palazzo S., Zambetta G., Calbi R.* An overview of segmentation techniques for CT and MRI images: Clinical implications and future directions in medical diagnostics // *Medical Imaging Process & Technology*. – 2024.
3. *Куликов А.А., Каширина И.Л., Савкина Е.Ф.* Сегментация объемных образований печени на мультифазных КТ-изображениях с использованием фреймворка nnU-Net // *Моделирование, оптимизация и информационные технологии*. – 2025. – Т. 13, № 1 (48).
4. *Zang L. et al.* Research on liver cancer segmentation method based on PCNN image processing and SE-ResUnet // *Scientific reports*. – 2023. – Vol. 13, No. 1. – P. 12779.
5. *Sun Q. et al.* UCR-Net: U-shaped context residual network for medical image segmentation // *Computers in Biology and Medicine*. – 2022. – Vol. 151. – P. 106203.
6. *Liu S. et al.* Segmenting nailfold capillaries using an improved U-net network // *Microvascular Research*. – 2020. – Vol. 130. – P. 104011.
7. *Ly P. et al.* An improved residual U-Net with morphological-based loss function for automatic liver segmentation in computed tomography // *Math. Biosci. Eng.* – 2022. – Vol. 19, No. 2. – P. 1426-1447.
8. *Черемискин А.В., Каширина И.Л.* Сегментация мультифазных КТ-изображений с использованием ансамбля моделей RESUNET // *Вестник Воронежского государственного университета. Серия: Системный анализ и информационные технологии*. – 2025. – № 3. – С. 140-152.
9. *Ермоленко С.В., Каширина И.Л., Старичкова Ю.В.* Постобработка результатов сегментации медицинских изображений // *Computational Nanotechnology*. – 2025. – Т. 12, № 2. – С. 109-118.
10. *Islam K.T., Wijewickrema S. & O'Leary S.* A deep learning based framework for the registration of three dimensional multi-modal medical images of the head // *Sci Rep*. – 2021. – 11, 1860. – <https://doi.org/10.1038/s41598-021-81044-7>.
11. *Uhm K.H. et al.* Deep learning for end-to-end kidney cancer diagnosis on multi-phase abdominal computed tomography // *NPJ precision oncology*. – 2021. – Vol. 5, No. 1. – P. 54.
12. *Pérez de Frutos J. et al.* Learning deep abdominal CT registration through adaptive loss weighting and synthetic data generation // *Plos one*. – 2023. – Vol. 18, No. 2. – P. e0282110.
13. *Balakrishnan G. et al.* Voxelmorph: a learning framework for deformable medical image registration // *IEEE transactions on medical imaging*. – 2019. – Vol. 38, No. 8. – P. 1788-1800.
14. *Wu L. et al.* Beyond radiologist-level liver lesion detection on multi-phase contrast-enhanced CT images by deep learning // *Iscience*. – 2023. – Vol. 26, No. 11.
15. *Nayantara P. V. et al.* Automatic liver segmentation from multiphase CT using modified SegNet and ASPP module // *SN Computer Science*. – 2024. – Vol. 5, No. 4. – P. 377.
16. *Bartnik K. et al.* WAW-TACE: a hepatocellular carcinoma multiphase CT dataset with segmentations, radiomics features, and clinical data // *Radiology: Artificial Intelligence*. – 2024. – Vol. 6, No. 6. – P. e240296.
17. *Wasserthal J. et al.* TotalSegmentator: robust segmentation of 104 anatomic structures in CT images // *Radiology: Artificial Intelligence*. – 2023. – Vol. 5, No. 5. – P. e230024.
18. *Isensee F. et al.* nnu-net revisited: A call for rigorous validation in 3d medical image segmentation // *International Conference on Medical Image Computing and Computer-Assisted Intervention*. – Cham: Springer Nature Switzerland, 2024. – P. 488-498.
19. *Klein S. et al.* Elastix: a toolbox for intensity-based medical image registration // *IEEE transactions on medical imaging*. – 2009. – Vol. 29, No. 1. – P. 196-205.
20. *Boussot V. et al.* IMPACT: A Generic Semantic Loss for Multimodal Medical Image Registration // *arXiv preprint arXiv:2503.24121*. – 2025.
21. *Paszke A. et al.* Pytorch: An imperative style, high-performance deep learning library // *Advances in neural information processing systems*. – 2019. – Vol. 32.
22. *Li Y., Zhang X.* Adaptive moment estimation optimization algorithm using projection gradient for deep learning // *Fifth International Conference on Applied Mathematics, Modelling, and Intelligent Computing (CAMMIC 2025)*. – SPIE, 2025. – Vol. 13644. – P. 501-511.

REFERENCES

1. *Isensee, F., Jaeger, P.F., Kohl, S. A., Petersen, J., & Maier-Hein, K.H.* nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation, *Nature methods*, 2021, 18 (2), pp. 203-211.
2. *Palazzo S., Zambetta G., Calbi R.* An overview of segmentation techniques for CT and MRI images: Clinical implications and future directions in medical diagnostics, *Medical Imaging Process & Technology*, 2024.
3. *Kulikov A.A., Kashirina I.L., Savkina E.F.* Segmentatsiya ob"emnykh obrazovaniy pecheni na mul'tifaznykh KT-izobrazheniyakh s ispol'zovaniem freymvorka nnU-Net [Segmentation of volumetric liver lesions on multiphase CT images using the nnU-Net framework], *Modelirovanie, optimizatsiya i informatsionnye tekhnologii* [Modeling, Optimization and Information Technology], 2025, Vol. 13, No. 1 (48).

4. Zang L. et al. Research on liver cancer segmentation method based on PCNN image processing and SE-ResUnet, *Scientific Reports*, 2023, Vol. 13, No. 1, pp. 12779.
5. Sun Q. et al. UCR-Net: U-shaped context residual network for medical image segmentation, *Computers in Biology and Medicine*, 2022, Vol. 151, pp. 106203.
6. Liu S. et al. Segmenting nailfold capillaries using an improved U-net network, *Microvascular Research*, 2020, Vol. 130, pp. 104011.
7. Lv P. et al. An improved residual U-Net with morphological-based loss function for automatic liver segmentation in computed tomography, *Math. Biosci. Eng.*, 2022, Vol. 19, No. 2, pp. 1426-1447.
8. Cheremiskin A.V., Kashirina I.L. Segmentatsiya mul'tifaznykh KT-izobrazheniy s ispol'zovaniem ansamblya modeley RESUNET [Segmentation of multiphase CT images using the RESUNET model ensemble], *Vestnik Voronezhskogo gosudarstvennogo universiteta. Seriya: Sistemnyy analiz i informatsionnye tekhnologii* [Bulletin of the Voronezh State University. Series: Systems Analysis and Information Technology], 2025, No. 3, pp. 140-152.
9. Ermolenko S.V., Kashirina I.L., Starichkova Yu.V. Postobrabotka rezul'tatov segmentatsii meditsinskikh izobrazheniy [Post-processing of medical image segmentation results], *Computational Nanotechnology*, 2025, Vol. 12, No. 2, pp. 109-118.
10. Islam K.T., Wijewickrema S. & O'Leary S. A deep learning based framework for the registration of three dimensional multi-modal medical images of the head, *Sci Rep.*, 2021, 11, 1860. Available at: <https://doi.org/10.1038/s41598-021-81044-7>.
11. Uhm K.H. et al. Deep learning for end-to-end kidney cancer diagnosis on multi-phase abdominal computed tomography, *NPJ precision oncology*, 2021, Vol. 5, No. 1, pp. 54.
12. Pérez de Frutos J. et al. Learning deep abdominal CT registration through adaptive loss weighting and synthetic data generation, *Plos one*, 2023, Vol. 18, No. 2, pp. e0282110.
13. Balakrishnan G. et al. Voxelmorph: a learning framework for deformable medical image registration, *IEEE transactions on medical imaging*, 2019, Vol. 38, No. 8, pp. 1788-1800.
14. Wu L. et al. Beyond radiologist-level liver lesion detection on multi-phase contrast-enhanced CT images by deep learning, *Iscience*, 2023, Vol. 26, No. 11.
15. Nayantara P.V. et al. Automatic liver segmentation from multiphase CT using modified SegNet and ASPP module, *SN Computer Science*, 2024, Vol. 5, No. 4, pp. 377.
16. Bartnik K. et al. WAW-TACE: a hepatocellular carcinoma multiphase CT dataset with segmentations, radiomics features, and clinical data, *Radiology: Artificial Intelligence*, 2024, Vol. 6, No. 6, pp. e240296.
17. Wasserthal J. et al. TotalSegmentator: robust segmentation of 104 anatomic structures in CT images, *Radiology: Artificial Intelligence*, 2023, Vol. 5, No. 5, pp. e230024.
18. Isensee F. et al. nnu-net revisited: A call for rigorous validation in 3d medical image segmentation, *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Cham: Springer Nature Switzerland, 2024, pp. 488-498.
19. Klein S. et al. Elastix: a toolbox for intensity-based medical image registration, *IEEE transactions on medical imaging*, 2009, Vol. 29, No. 1, pp. 196-205.
20. Boussof V. et al. IMPACT: A Generic Semantic Loss for Multimodal Medical Image Registration, *arXiv preprint arXiv:2503.24121*, 2025.
21. Paszke A. et al. Pytorch: An imperative style, high-performance deep learning library, *Advances in neural information processing systems*, 2019, Vol. 32.
22. Li Y., Zhang X. Adaptive moment estimation optimization algorithm using projection gradient for deep learning, *Fifth International Conference on Applied Mathematics, Modelling, and Intelligent Computing (CAMMIC 2025)*. SPIE, 2025, Vol. 13644, pp. 501-511.

Ермоленко Сергей Витальевич – Воронежский государственный университет; e-mail: ermolenko@math.vsu.ru; г. Воронеж, Россия; тел.: +79601087190; аспирант.

Каширина Ирина Леонидовна – МИРЭА – Российский технологический университет; e-mail: kashirina@mirea.ru; г. Москва, Россия; тел.: +79036539293; д.т.н.; профессор кафедры технологий искусственного интеллекта.

Ermolenko Sergey Vitalievich – Voronezh State University; e-mail: ermolenko@math.vsu.ru; Voronezh, Russia; phone: +79601087190; graduate student.

Kashirina Irina Leonidovna – MIREA – Russian Technological University; e-mail: kash.irina@mail.ru; Moscow, Russia; phone: +79036539293; dr. of eng. sc.; professor of the Department of Artificial Intelligence Technologies.

И.Е. Лысенко

**МАЛОРЕСУРСНАЯ АБСТРАКТИВНАЯ СУММАРИЗАЦИЯ ТЕКСТОВ
С ИСПОЛЬЗОВАНИЕМ КОНТРАСТИВНОГО ОБУЧЕНИЯ БЕЗ УЧИТЕЛЯ
И ROUGE-ОРИЕНТИРОВАННОЙ ФУНКЦИИ ПОТЕРЬ**

Актуальность данной работы определяется тем, что в специализированных предметных областях и малоресурсных языках получение достаточного числа примеров «текст-реферат» для эффективного решения задачи абстрактивной суммаризации текстов оказывается дорогостоящим и практически невыполнимым, тогда как современные модели глубокого обучения требуют крупных размеченных корпусов, а большие объемы неразмеченных текстов остаются неиспользуемыми. Цель исследования заключается в разработке нового метода обучения моделей абстрактивной суммаризации для малоресурсных (10-примерных и 100-примерных) режимов, обеспечивающего рост качества за счет дообучения модели на неразмеченных данных с использованием контрастивного обучения без учителя с аугментацией входа. Из задач исследования выделяются создание новой контрастивной функции потерь и проведение сравнения полученного подхода с существующими методами малоресурсной абстрактивной суммаризации. Методы и подходы включают новую функцию потерь для дообучения трансформера, включающую в себя контрастивную генеративную компоненту дифференцируемой аппроксимации метрики ROUGE-3. Предложены два варианта метода – последовательный «DiffROUGE-seq» и метод с частичным привлечением учителя «DiffROUGE-sim». В качестве базовой модели используется BART-large, а аугментации входа генерируются моделью FLAN-T5-large в 0-примерном режиме. Произведено сравнение на популярных выборках AESLC, Gigaword, XSum и Reddit. Предложенный метод показал заметное улучшение метрик ROUGE в успешных случаях в среднем на 2,69 для 10-примерного режима и в среднем на 2,13 для 100-примерного режима. Таким образом, новый предложенный метод малоресурсной абстрактивной суммаризации, использующий как размеченные, так и неразмеченные данные, заметно более эффективен по метрике ROUGE по сравнению с существующими подходами.

Обработка естественного языка; абстрактивная суммаризация; малоресурсный режим; глубокое обучение; трансформер; контрастивное обучение; дифференцируемая аппроксимация метрики.

I.E. Lysenko

**LOW-RESOURCE ABSTRACTIVE TEXT SUMMARIZATION VIA CONTRASTIVE
UNSUPERVISED LEARNING WITH A ROUGE-ORIENTED LOSS FUNCTION**

The relevance of this work is motivated by the fact that in specialized domains and low-resource languages obtaining a sufficient number of “document-summary” pairs for effective abstractive text summarization is expensive and often practically infeasible, whereas modern deep learning models require large labeled corpora and large volumes of text remain unused. The aim of this study is to develop a new training method for abstractive summarization models in low-resource (10-shot and 100-shot) settings that improves quality by fine-tuning the model on unlabeled data using unsupervised contrastive learning with input augmentation. The research tasks include designing a new contrastive loss function and comparing the proposed approach with existing methods of low-resource abstractive summarization. The methods and approaches comprise a new loss function for fine-tuning a transformer that includes a contrastive generative component based on a differentiable approximation of the ROUGE-3 metric. Two variants of the method are proposed – the sequential “DiffROUGE-seq” and the semi-supervised “DiffROUGE-sim”. BART-large is used as the base model, while input augmentations are generated by FLAN-T5-large in a zero-shot regime. Experiments are conducted on popular datasets AESLC, Gigaword, XSum, and Reddit. The proposed method achieves substantial improvements in ROUGE scores in successful cases, with average gains of 2.69 in the 10-shot setting and 2.13 in the 100-shot setting. In summary, the newly proposed low-resource abstractive summarization method that leverages both labeled and unlabeled data is significantly more effective in terms of ROUGE than existing approaches.

Natural language processing; abstractive summarization; low-resource setting; deep learning; transformer; contrastive learning; differentiable approximation of a metric.

Введение. Стремительный рост объемов текстовой информации в науке, промышленности и медиа усложняет ее анализ и усиливает потребность в надежных методах суммаризации, которые позволяют по входному тексту получить выжимку главной информации в нем. Это позволяет решить проблему информационной перегрузки читателя

и улучшает степень восприятия содержания текстов. Различается экстрактивная суммаризация, задача которой заключается в составлении реферата путем выбора наиболее информативных предложений текста, и абстрактивная суммаризация, задача которой заключается в перефразировании и новой формулировке исходного текста в более короткой форме. На данный момент, с ростом популярности больших языковых моделей, абстрактивная суммаризация становится все более распространенной, так как она позволяет получить не только информативный, но и связный текст реферата.

Абстрактивная суммаризация, так как она основана на больших языковых моделях, которые зачастую базируются на архитектуре «трансформер» [1], требует больших объемов тренировочных пар «текст-реферат», чтобы используемая модель научилась генерировать рефераты высокого качества. Однако в специализированных областях, таких как юриспруденция, финансы или медицина, получение достаточного числа таких пар является практически невыполнимым из-за стоимости разметки, требования к специализированным знаниям и т.п., но количество неразмеченных данных в данных областях зачастую большое. Таким образом, применение абстрактивной суммаризации к текстам из таких областей существенно ограничено и является открытой проблемой.

В данной работе мы рассматриваем случай, когда имеется мало размеченных примеров «текст-реферат» (10 или 100) и много неразмеченных примеров, то есть только текстов без сопутствующих им рефератов. В отличие от существующих методов обучения больших языковых моделей, таких как BART [2], PEGASUS [3] и T5 [4], предложенный нами подход позволяет воспользоваться информацией из неразмеченного набора данных для улучшения качества абстрактивной суммаризации. Мы предлагаем новый метод обучения, который заключается в максимизации предложенной нами контрастивной модификации дифференцируемой аппроксимации популярной метрики ROUGE-3 [5] между сгенерированными моделью текстами для исходного неразмеченного текста и его аугментации. Мы также показываем, что предложенный метод зачастую оказывается заметно лучше существующих в литературе методов по метрикам ROUGE [5].

1. Описание и постановка решаемой задачи. 1.1. Формулировка проблемы. Абстрактивная суммаризация – задача обработки естественного языка, которая заключается в том, чтобы по входному тексту $X = (x_1, x_2, \dots, x_{|X|})$ получить его краткий реферат $Y = (y_1, y_2, \dots, y_{|Y|})$, такой, что он является связным текстом, охватывающим всю важную информацию из исходного текста, и $|X| \gg |Y|$.

Задача абстрактивной суммаризации обычно оценивается метрикой ROUGE-N [5]. Формально, пусть f_θ – модель абстрактивной суммаризации с параметрами θ , которая по входному тексту X порождает реферат $Y = f_\theta(X)$, и пусть $D = \{(X^{(i)}, \hat{Y}^{(i)})\}_{i=1}^{|D|}$ – тестовая выборка пар «текст – эталонный реферат». В качестве технического параметра, являющегося критерием оценки качества абстрактивной суммаризации, используется метрика ROUGE-N [5], измеряющая долю совпадающих N-грамм между сгенерированным Y и эталонным $\hat{Y}$ рефератами. Обозначив через $C_T(g)$ число вхождений N-граммы g в текст T , определим:

$$C_Y^{\hat{Y}}(g) = \min(C_Y(g), C_{\hat{Y}}(g))$$

$$p_n(Y, \hat{Y}) = \frac{\sum_{g \in Y} C_Y^{\hat{Y}}(g)}{\sum_{g \in Y} C_Y(g)}; \quad r_n(Y, \hat{Y}) = \frac{\sum_{g \in Y} C_Y^{\hat{Y}}(g)}{\sum_{g \in \hat{Y}} C_{\hat{Y}}(g)}$$

$$ROUGE_N(Y, \hat{Y}) = 100 \cdot 2 \cdot \frac{p_n(Y, \hat{Y}) \cdot r_n(Y, \hat{Y})}{p_n(Y, \hat{Y}) + r_n(Y, \hat{Y})} \in [0, 100].$$

Критерием качества модели служит усредненное по тестовой выборке значение метрики:

$$Q(\theta) = \frac{1}{|D|} \sum_{i=1}^{|D|} ROUGE_N(Y^{(i)}, \hat{Y}^{(i)}).$$

Таким образом, задача формализуется как поиск параметров θ^* , максимизирующих критерий $Q(\theta)$:

$$\theta^* = \arg \max_{\theta} Q(\theta).$$

В данном исследовании рассматривается случай абстрактивной суммаризации, когда доступно 10 или 100 пар «текст-реферат» и много неразмеченных текстов без рефератов, так как этот случай описывает существующие области текстов, например, юриспруденцию, медицину или финансы, и существующие методы абстрактивной суммаризации плохо работают с текстами из этих областей ввиду недостаточного количества размеченных данных. Таким образом, проблема исследования заключается в том, что задача абстрактивной суммаризации текстов с малым числом размеченных данных плохо решается существующими популярными методами [2–4] ввиду того, что они требуют больших объемов пар «текст-реферат» для обучения.

1.2. Актуальность темы. Актуальность темы абстрактивной суммаризации обусловлена быстрым ростом объемов текстовой информации во многих областях, которую ввиду проблемы информационного перегруза сложно просматривать, воспринимать и анализировать, поэтому возникает потребность в автоматизированных средствах сжатия и обобщения данных. Системы абстрактивной суммаризации, способные формировать связные и информативные рефераты за счет генерации новых формулировок, имеют важное значение как для развития фундаментальных исследований в области обработки естественного языка (например, как одна из задач для оценки качества новых больших языковых моделей), так и для практических приложений – от новостных агрегаторов и интеллектуальных поисковых систем до поддержки принятия решений в таких областях, как медицина, юриспруденция или финансы.

Перспективы развития темы связаны с совершенствованием методов, устойчиво работающих в режимах дефицита данных. Ожидается, что развитие подходов к абстрактивной суммаризации, основанных на эффективном использовании неразмеченных корпусов и контрастивного обучения, позволит получать более точные и содержательные резюме для широкого круга прикладных задач и областей.

1.3. Обзор литературы. **1.3.1 Методы абстрактивной суммаризации.** Одними из самых популярных моделей для решения задачи абстрактивной суммаризации на данный момент являются BART [2], PEGASUS [3] и T5 [4], основанные на архитектуре «трансформер» [1].

BART [2] – модель кодировщик-декодировщик, которая училась без учителя на задачу восстановления текста после одного из пяти возможных методов зашумления: маскировки токенов, удаления токенов, заполнения текста, перестановки предложений и циклического сдвига документа.

PEGASUS [3] – модель кодировщик-декодировщик, которая была создана для решения задачи абстрактивной суммаризации. Данная модель, аналогично BART [2], обучалась без учителя на задачу восстановления текста после зашумления, однако при обучении данной модели для зашумления исходного текста используется метод генерации пропущенных предложений, предложенный авторами, чтобы метод зашумления более тесно соответствовал целевой задаче абстрактивной суммаризации и приводил к лучшим и более быстрым результатам дообучения. Данная модель показывает хорошие результаты для малопримерных режимов и зачастую показывает наилучшие результаты среди трех рассмотренных моделей, поэтому мы используем ее в качестве базовой для сравнения с предложенным нами подходом.

T5 [4] – модель кодировщик-декодировщик, которая сначала обучалась без учителя на задачу восстановления текста по искаженным фрагментам, а затем училась на отдельные задачи (абстрактивная суммаризация, машинный перевод и т.п.), используя префикс задачи («summarize: ...», «translate English to German: ...»). Таким образом, данная модель способна решать множество различных генеративных задач одновременно, используя разные префиксы задач.

Данные модели зачастую используются в качестве базовых, вокруг которых строятся существующие подходы к малопримерной абстрактивной суммаризации.

Активное обучение – метод оптимальной разметки данных с ограниченным бюджетом, который итеративно выбирает для разметки примеры, которые имеют максимальную оценку неопределенности для данной модели. Активное обучение часто используется в случае дорогой разметки, когда размеченных данных мало, а неразмеченных – много, в частности, в абстрактивной суммаризации. IDDS [6] – метод оценки неопределенности генеративной модели для использования в активном обучении, который каждый шаг приоритизирует объекты, которые находятся в среднем далеко от размеченных данных, но в среднем близко к неразмеченным. В данной статье авторы также предлагают метод самообучения, во время которого для разметки неразмеченных данных используется сама модель, обученная на малом количестве имеющихся размеченных данных. Для экспериментов в малоресурсных режимах авторы используют модель BART [2]. Во время сравнения мы используем название «Tsvigun et al.» для обращения к лучшим подходам данной статьи.

FactPEGASUS [7] – модель абстрактивной суммаризации, основанная на модели PEGASUS [3]. Авторы, во-первых, модифицируют метод генерации пропущенных предложений PEGASUS [3], чтобы выбирать более фактические предложения, а, во-вторых, предлагают использовать во время дообучения три модуля, улучшающих фактичность итоговой модели: соединитель, добавляющий токен маскировки в резюме для соответствия задаче предобучения, корректор, убирающий галлюцинации из резюме, и контрактор, который добавляет в функцию потерь контрастивную компоненту, позволяющую модели выбирать более фактические резюме. Используя данные нововведения, FactPEGASUS [7] показывает способность к генерации более фактических резюме в малоресурсных режимах.

MTL-ABS [8] – метод малоресурсной абстрактивной суммаризации, который использует как латентные знания большой языковой модели, так и внешние корпуса текстов, для улучшения качества решения данной задачи. Авторы предлагают добавить в трансформер [1] слой-адаптеры, чтобы не обучать модель полностью и облегчить перенос знаний из внешних корпусов текстов. Для экспериментов авторы используют трансформер [1] с 12 слоями кодировщика и 6 слоями декодировщика.

UniSumm [9] – метод обучения моделей суммаризации в малоресурсном режиме. Авторы сначала тренируют суммаризационную модель на различных выборках, помечая их специальными префиксами, включая универсальный префикс для дальнейшего эффективного малопримерного обучения. Далее авторы используют универсальный префикс для дообучения модели на конкретной малопримерной области. Авторы используют BART [2] в качестве базовой суммаризационной модели.

SummaFusion-large [10] – метод абстрактивной суммаризации, который заключается в комбинации нескольких рефератов-кандидатов для получения итогового реферата. Авторы предлагают передавать в декодировщик конкатенированное векторное представление исходного текста и всех рефератов-кандидатов. Авторы утверждают, что такой подход улучшает качество в случае, когда качество отдельных кандидатов низкое, например, в малоресурсной абстрактивной суммаризации. В качестве базовой модели авторы используют PEGASUS [3].

1.3.2. Выборки данных. Мы проводим эксперименты на четырех популярных выборках данных AESLC [11], Gigaword [12], XSum [13] и Reddit [14]. Аналогично существующим методам, мы симулируем малоресурсный режим, выбирая случайные 10 или 100 размеченных тренировочных примеров для обучения с учителем и используем остальные тренировочные данные для обучения без учителя.

AESLC [11] – выборка электронных писем компании Enron, суммаризационной задачей которой является генерация темы письма по его содержанию. И тексты, и резюме данной выборки достаточно короткие – в среднем 75 и 4 слова, соответственно. Выборка содержит 14 436 тренировочных, 1 960 валидационных и 1 906 тестовых примеров. Область выборки – деловые и персональные письма.

Gigaword [12] – выборка новостного характера, задачей для которой является составление заголовка новостной статьи по ее содержанию. И тексты, и резюме данной выборки также являются короткими – в среднем 31 и 8 слов, соответственно. Выборка содержит 3 803 957 тренировочных, 189 651 валидационных и 1 951 тестовых примера.

XSum [13] – выборка статей BBC из различных областей (новости, политика, спорт, погода и т.п.), суммаризационной задачей которой является составление резюме из одного предложения по содержанию статьи. Авторы использовали подзаголовок статьи в качестве верного резюме. Тексты данной выборки являются длинными, тогда как резюме являются короткими – в среднем 431 и 23 слова, соответственно. Выборка содержит 204 045 тренировочных, 11 332 валидационных и 11 334 тестовых примера.

Reddit [14] – выборка постов социальной сети Reddit, конкретно, форума r/TIFU. Авторы используют в качестве верного резюме поле «tl;dr», которое присутствует в каждом посте и описывает его краткое содержание. Тексты данной выборки являются длинными, а резюме короткими – в среднем 432,6 и 23 слова, соответственно. Выборка содержит 33 711 тренировочных, 4 214 валидационных и 4 214 тестовых примера.

1.3.3. Литература, связанная с предложенными методами. UDA [15] – негенеративный метод контрастного обучения моделей машинного обучения, который заключается в добавлении в функцию потерь слагаемого кросс-энтропии от выходного условного вероятностного распределения модели на входном примере и от выходного условного вероятностного распределения модели на аугментации входного примера, которая сохраняет семантический смысл исходного примера, но является другим входным примером. Данный метод позволяет использовать подход с частичным привлечением учителя, приближая с точки зрения модели семантически идентичные примеры. В нашем исследовании мы также используем аналогичное контрастивное слагаемое для нашего метода «DiffROUGE-sim», однако мы используем нашу контрастивную дифференцируемую аппроксимацию метрики ROUGE-3 [5] в качестве меры близости ответов модели.

P-BLEU [16] – дифференцируемая аппроксимация метрики BLEU [17], которая часто используется для измерения качества в задаче машинного перевода и измеряет пересечение N-грамм текста, аналогично ROUGE [5]. Авторы предлагают фреймворк составления дифференцируемых аппроксимаций метрик, основанных на пересечении N-грамм, однако он требует знания верного ответа для входного текста. В нашем исследовании мы модифицируем данный фреймворк для получения контрастивной дифференцируемой аппроксимации метрики ROUGE-3 [5].

FLAN-T5 [18] – улучшенная версия модели T5 [4], которая была обучена на 1 800 различных задачах, сформулированных в виде инструкций на естественном языке. Данная модель показывает высокое качество на многих задачах в 0-примерном режиме (единственный вход модели – инструкция, без примеров решения), поэтому мы используем ее в качестве генератора аугментаций текстов в 0-примерном режиме.

ROUGE [5] – популярная метрика качества абстрактивной суммаризации, которая основана на измерении пересечения N-грамм текстов и принимает значения от 0 (отсутствие пересечения) до 100 (полное пересечение). Хотя существуют более продвинутые метрики качества абстрактивной суммаризации, например, BERTScore [19] или BARTSCORE [20], которые, как показывают исследования, лучше коррелируют с человеческим мнением, мы используем ROUGE [5] для итогового сравнения моделей, так как, во-первых, модельные метрики выше также требуют дообучения для хорошего качества на малоресурсных областях, в отличие от ROUGE [5], а, во-вторых, мы используем BART [2] в качестве базовой модели для наших подходов, и в нашем случае использовать BARTSCORE [20] некорректно, так как данная метрика будет предвзята в пользу нашей модели.

1.3.4. Используемое программное обеспечение. В качестве библиотеки для загрузки, сохранения и обучения трансформеров [1] мы используем библиотеку transformers [21], так как она является наиболее популярной для данного типа моделей. Для автоматического дифференцирования и работы с тензорами мы используем популярную библиотеку PyTorch [22]. Для загрузки, обработки и сохранения текстовых выборок данных мы используем популярную библиотеку datasets [23], так как она хорошо интегрирована с используемой нами библиотекой transformers [21].

1.4. Цель исследования. Цель исследования заключается в разработке нового метода обучения моделей абстрактивной суммаризации для малоресурсных (10-примерных и 100-примерных) режимов, обеспечивающего рост качества за счет дообучения модели на неразмеченных данных с использованием контрастивного обучения без учителя с аугментацией входа.

1.5. Задачи исследования:

1. Проведение сравнительного анализа существующих методов решения задачи малоресурсной (10-примерной и 100-примерной) абстрактивной суммаризации.
2. Создание новой контрастивной функции потерь, основанной на дифференцируемой аппроксимации метрики ROUGE-3 [5], которая позволяет использовать неразмеченные данные в малопримерном режиме обучения задачи абстрактивной суммаризации.
3. Реализация и проведение экспериментов на популярных выборках данных с предложенным методом.
4. Сравнение предложенного метода с существующими, получение оценки эффективности подхода и формулирование рекомендаций по его применению.

1.6. Объект и предмет исследования. Объектом исследования являются процессы и методы абстрактивной суммаризации текстов на основе моделей глубокого обучения. Предметом исследования является метод малоресурсного (10-примерного и 100-примерного) обучения моделей абстрактивной суммаризации, основанный на контрастивном дообучении без учителя трансформера с использованием аугментации входных данных и ROUGE-ориентированной функции потерь.

2. Методика исследования. 2.1. Стратегия и подходы к исследованию. Выделяется несколько возможных подходов решения задачи малоресурсной абстрактивной суммаризации, в частности, классическое обучение с учителем, самообучение модели, использование специализированных промптов вкупе с большими языковыми моделями и контрастивное обучение. Мы используем именно подход с контрастивным обучением, поскольку он, в отличие от других подходов, позволяет эффективно использовать неразмеченные данные, избегает псевдо-разметки и формирует инвариантность и робастность к семантически идентичным входам. Остальные подходы также имеют свои недостатки:

- ◆ Классическое обучение с учителем вызывает быстрое переобучение и имеет проблему низкого качества и высокой чувствительности к инициализации и случайности выборки.

- ◆ Самообучение модели зачастую усиливает эффект ошибок разметки данных моделью и приводит к сильному снижению качества итоговой модели.

- ◆ Большие языковые модели, во-первых, вычислительно тяжелы и дороги в использовании и дообучении, а, во-вторых, их нельзя использовать в случае, когда данные не разрешено выкладывать в открытый доступ (например, в случае соглашения о неразглашении или ограничивающего согласия об обработке данных).

Мы используем метод эксперимента для эмпирического подтверждения разумности и эффективности нашего метода по сравнению с существующими, поскольку ввиду размера и сложности используемых моделей аналитический анализ метода затруднен.

Исследование подразделяется на несколько этапов:

1. Анализ существующих методов малоресурсной абстрактивной суммаризации – позволяет выявить, какие подходы уже предложены, какие результаты они показывают и какие ограничения остаются.
2. Формирование исследовательской гипотезы – задает четкую цель, к проверке которой направлены все последующие шаги.
3. Аугментация входных данных – создает основу в виде итоговых выборок данных для дальнейшего контрастивного обучения.
4. Программная реализация предложенных методов – переводит теоретическую идею в работающий прототип.
5. Проведение экспериментов с предложенными методами – дает количественную оценку того, как новый метод работает в конкретных условиях.

6. Проведение сравнения с существующими методами малоресурсной абстрактивной суммаризации в 10-примерном и 100-примерном режимах – позволяет ответить на ключевой вопрос: лучше ли предлагаемый метод по сравнению с существующими.

7. Анализ результатов и подведение итогов – обеспечивает интерпретацию полученных численных результатов.

2.1. Описание предлагаемых методов. В качестве контрастивной метрики близости сгенерированных моделью текстов мы используем метрику ROUGE-3 [5], которая принимает значения от 0 до 100 и измеряет долю пересечений 3-грамм входных текстов. Используя существующие обозначения из литературы [16], введем формальное описание метрики ROUGE-N [5]. Для сгенерированного моделью резюме $y \in Y$ по входу $x \in X$, верного ответа $\hat{y} \in Y$ и N-граммы g :

$$C_y(g) = \sum_{t=0}^{T-n} \prod_{i=1}^n 1\{g_i = y_{t+i}\}; \quad C_{\hat{y}}(g) = \min(C_y(g), C_{\hat{y}}(g))$$

$$p_n(y, \hat{y}) = \frac{\sum_{g \in y} C_{\hat{y}}(g)}{\sum_{g \in y} C_y(g)}; \quad r_n(y, \hat{y}) = \frac{\sum_{g \in y} C_{\hat{y}}(g)}{\sum_{g \in \hat{y}} C_{\hat{y}}(g)}$$

$$ROUGE_N(y, \hat{y}) = 100 \cdot 2 \cdot \frac{p_n(y, \hat{y}) \cdot r_n(y, \hat{y})}{p_n(y, \hat{y}) + r_n(y, \hat{y})} \in [0, 100].$$

Для получения дифференцируемой аппроксимации P-ROUGE для данной метрики, основанной на полученной от языковой модели условной вероятности следующего токена, авторы [16] вводят

$$\tilde{C}_y(g) = \sum_{t=0}^{T-n} \prod_{i=1}^n 1\{g_i = y_{t+i}\} P(y_{t+i} | y_{<t+i}, x, \theta); \quad \tilde{C}_{\hat{y}}(g) = \min(\tilde{C}_y(g), C_{\hat{y}}(g)).$$

Далее,

$$\tilde{p}_n(y, \hat{y}) = \frac{\sum_{g \in y} \tilde{C}_{\hat{y}}(g)}{\sum_{g \in y} \tilde{C}_y(g)}; \quad \tilde{r}_n(y, \hat{y}) = \frac{\sum_{g \in y} \tilde{C}_{\hat{y}}(g)}{\sum_{g \in \hat{y}} \tilde{C}_{\hat{y}}(g)}$$

$$P_{ROUGE_N(y, \hat{y})} = 100 \cdot 2 \cdot \frac{\tilde{p}_n(y, \hat{y}) \cdot \tilde{r}_n(y, \hat{y})}{\tilde{p}_n(y, \hat{y}) + \tilde{r}_n(y, \hat{y})}.$$

Однако в случае обучения без учителя верный ответ $\hat{y}$ недоступен, поэтому для нашего метода мы предлагаем модифицировать P-ROUGE, чтобы, во-первых, ориентироваться только на ответы модели в отсутствие размеченных верных ответов, а, во-вторых, улучшить дифференцируемость функции для возможности использования полной информации по градиенту (в частности, функция $\min$ не является дифференцируемой в классическом смысле).

Пусть существует некоторая функция аугментации $T: X \rightarrow X$, которая для входного текста $x \in X$ преобразует его в текст $x^T \in X$, который имеет такой же семантический смысл, как и x , но $x \neq x^T$. Вместо верного ответа $\hat{y} \in Y$ используем ответ модели $y^T \in Y$ по входу x^T и используем функцию $\text{LogSumExp}(x, y) = \log(\exp(x) + \exp(y))$ для получения дифференцируемой аппроксимации функции минимума [24]. Введем предложенную нами новую контрастивную метрику близости сгенерированных моделью текстов DiffROUGE_N :

$$M = \max(\tilde{C}_y(g), \tilde{C}_{y^T}(g))$$

$$\bar{C}_y^{y^T}(g) = -\frac{1}{\beta} \text{LogSumExp}(-\beta(\tilde{C}_y(g) - M), -\beta(\tilde{C}_{y^T}(g) - M)) + M, \beta \in \mathbb{R}.$$

Таким образом,

$$\bar{p}_n(y, y^T) = \frac{\sum_{g \in y} \bar{C}_y^{y^T}(g)}{\sum_{g \in y} \bar{C}_y(g)}; \quad \bar{r}_n(y, y^T) = \frac{\sum_{g \in y} \bar{C}_y^{y^T}(g)}{\sum_{g \in y^T} \bar{C}_y^{y^T}(g)}$$

$$\text{DiffROUGE}_N(y, y^T) = 2 \cdot \frac{\bar{p}_n(y, y^T) \cdot \bar{r}_n(y, y^T)}{\bar{p}_n(y, y^T) + \bar{r}_n(y, y^T)}.$$

Градиент не считается по M во время обучения методом градиентного спуска, так как этот параметр предназначен только для численной стабильности метода. В наших экспериментах параметр $\beta = 20$ и используется для того, чтобы предложенная аппроксимация функции минимума была ближе к самой функции минимума. В наших экспериментах $N = 3$. Мы также не умножаем результат на 100, так как при максимизации данной функции умножение на константу не повлияет на направление градиента.

Пусть $L = \{(x_i, \hat{y}_i) \mid i = 1 \dots |L|, x_i \in X, \hat{y}_i \in Y\}$ – размеченная выборка «текст-резюме», а $U = \{x_j \mid j = 1 \dots |U|, x_j \in X\}$ – неразмеченная выборка. Пусть для обучения модели $f_\theta(x) : X \rightarrow Y$ используется функция потерь $\ell(f_\theta(x), \hat{y}), (x, \hat{y}) \in L$, для трансформеров [1] это зачастую будет кросс-энтропийная функция потерь. Тогда назовем методом обучения «DiffROUGE-seq» последовательную минимизацию ℓ в рамках обучения с учителем на ограниченной размеченной выборке и $1 - \text{DiffROUGE}_3(f_\theta(x), f_\theta(T(x))), x \in U$ в рамках обучения без учителя на большой неразмеченной выборке. Методом обучения с частичным привлечением учителя «DiffROUGE-sim» назовем минимизацию следующей контрастивной функции потерь, построенную аналогично методу UDA [15]:

$$\tilde{\ell}(x_L, \hat{y}, x_U) = \ell(f_\theta(x_L), \hat{y}) - \lambda \frac{1}{|X_U|} \sum_{x_U \in X_U} \text{DiffROUGE}_3(f_\theta(x_U), f_\theta(T(x_U))), (x_L, \hat{y}) \in L, x_U \subseteq U, \lambda \in \mathbb{R}.$$

В наших экспериментах мы используем $\lambda = 3$, $|X_U| = 5$ и $|L| = 10$ или 100. Во время дообучения модели нашим методом мы используем скорость обучения $1e-6$ и линейное расписание изменения скорости обучения с долей разогрева 0,06. В качестве базовой модели f_θ для предложенных методов используется модель BART-large [2]. В качестве T мы используем модель FLAN-T5-large [18] в 0-примерном режиме без дообучения. Обучение и аугментация данных проходят на одной видеокарте Nvidia V100 32GB VRAM.

Важно заметить, что оптимизация DiffROUGE_N в данном случае не вызывает предвзятости модели к метрике ROUGE [5] во время итоговой оценки модели, так как DiffROUGE_N используется только в качестве контрастивной метрики близости между ответами модели для разных входов для формирования у модели робастности к семантически идентичным входам, и мы не используем ее во время обучения с учителем, когда имеются размеченные верные ответы. Таким образом, формат и цель использования DiffROUGE_N фундаментально отличается от случая итоговой оценки модели, когда сравнивается ответ модели и размеченный верный ответ для одного входа.

3. Результаты исследования. Мы проводим сравнения предложенных нами методов с существующими, используя стандартную метрику качества суммаризации ROUGE [5], так как она используется повсеместно для оценки качества суммаризации, и так как она не требует дообучения моделей машинного обучения, поскольку основана на правилах.

Ниже в табл. 1 и 2 предоставлены результаты экспериментов и сравнения метрик ROUGE [5] для существующих подходов и предложенных нами «DiffROUGE-seq» и «DiffROUGE-sim» на нескольких выборках данных. **Жирным** выделен лучший результат в столбце для режима, а *курсивом* выделено второе лучшее значение метрики для методов, предложенных нами, если они лучше существующих подходов. В каждой таблице мы опускаем методы, которые не предоставляют в оригинальных статьях результаты для выборок в таблице.

Табл. 1 показывает, что на выборках AESLC [11] и Gigaword [12], в которых и тексты, и рефераты короткие, предложенные нами методы оказываются лучше существующих по всем метрикам. «DiffROUGE-seq» оказывается лучшим на AESLC [11] в 10-примерном режиме и на Gigaword [12] в 100-примерном режиме, а «DiffROUGE-sim» оказывается лучше в остальных случаях.

Таблица 1

Сравнение метрик на выборках AESLC и Gigaword

	AESLC [11]			Gigaword [12]		
	ROUGE-1	ROUGE-2	ROUGE-L	ROUGE-1	ROUGE-2	ROUGE-L
10-примерный режим						
PEGASUS-large [3]	11,97	4,91	10,84	25,32	8,88	22,55
Tsvigun et al. [6]	17,78	8,69	17,61	24,33	8,85	22,43
MTL-ABS [8]	21,27	10,79	20,85	28,98	11,86	26,74
DiffROUGE-seq	25,04	12,59	22,79	31,25	13,19	28,27
DiffROUGE-sim	23,88	11,98	22,79	31,37	13,28	28,43
100-примерный режим						
PEGASUS-large [3]	16,05	7,20	15,32	29,71	12,44	27,30
Tsvigun et al. [6]	27,31	14,52	26,69	30,16	12,32	27,90
MTL-ABS [8]	23,88	12,06	23,18	30,03	12,70	27,71
DiffROUGE-seq	30,82	15,35	29,86	33,16	14,32	30,28
DiffROUGE-sim	31,34	15,50	30,41	32,11	13,82	29,56

В табл. 2, охватывающей результаты на выборках XSum [13] и Reddit [14], в которых длинные тексты, но короткие резюме, предложенные нами методы оказываются лучше остальных во всех случаях, кроме XSum [13] в 100-примерном режиме. В случаях, когда предложенные нами методы лучше существующих, преобладает метод «DiffROUGE-seq».

Таблица 2

Сравнение метрик на выборках XSum и Reddit

	XSum [13]			Reddit [14]		
	ROUGE-1	ROUGE-2	ROUGE-L	ROUGE-1	ROUGE-2	ROUGE-L
10-примерный режим						
PEGASUS-large [3]	19,39	3,45	14,02	15,36	2,91	10,76
FactPEGASUS [7]	-	-	23,44	-	-	-
MTL-ABS [8]	-	-	-	18,03	6,41	17,10
UniSumm [9]	26,10	7,20	19,92	22,88	5,60	17,02
SummaFusion-large [10]	30,41	9,92	22,93	20,79	4,77	14,58
DiffROUGE-seq	35,20	13,16	27,21	25,69	6,74	20,03
DiffROUGE-sim	34,64	12,79	26,57	23,39	6,48	17,91
100-примерный режим						
PEGASUS-large [3]	39,07	16,44	31,27	16,64	4,09	12,92
FactPEGASUS [7]	-	-	25,44	-	-	-
MTL-ABS [8]	-	-	-	20,14	7,71	19,38
UniSumm [9]	33,33	11,36	25,85	24,54	6,17	18,30
SummaFusion-large [10]	39,86	17,01	31,68	26,09	7,51	20,22
DiffROUGE-seq	37,11	14,93	29,40	27,22	8,34	21,88
DiffROUGE-sim	37,12	14,57	28,54	26,78	7,85	20,61

Таким образом, предложенные нами контрастивные методы обучения действительно приводят к улучшению качества малоресурсной абстрактивной суммаризации почти во всех случаях, кроме XSum [13] 100-примерного режима. Мы предполагаем, что это связано со спецификой XSum [13] как задачи экстремальной абстрактивной суммаризации: рефераты из одного предложения оказываются существенно более чувствительными к аугментациям входа. В этих условиях ROUGE-ориентированная функция потерь может приводить к излишней регуляризации и смещать модель от узкоспециализированных резюме, ожидаемых в XSum [13].

Выводы. В данном исследовании мы предложили новые методы обучения генеративных языковых моделей «DiffROUGE-seq» и «DiffROUGE-sim», научная новизна которых заключается в новой функции потерь, основанной на предложенной контрастивной версии дифференцируемой аппроксимации метрики ROUGE-3 [5], которая позволяет в режиме малоресурсной абстрактивной суммаризации воспользоваться большим корпусом неразмеченных данных для улучшения качества решения задачи абстрактивной суммаризации. Практическая значимость данного исследования заключается в том, что предложенный метод позволяет значительно снизить потенциальные затраты на разметку данных для абстрактивной суммаризации в областях, в которых мало размеченных данных для этой задачи, например, в юриспруденции, финансах или медицине. Системы абстрактивной суммаризации в этих областях позволяют снизить информационный перегруз читателя и дают читателю возможность сфокусироваться только на главной информации текста. Такие системы могут быть использованы, например, для сжатия научных медицинских публикаций, финансовых отчетностей или новостей, юридических документов.

Мы провели сравнение на четырех популярных выборках данных AESLC [11], Gigaword [12], XSum [13] и Reddit [14] в 10-примерном и 100-примерном режимах и показали, что предложенные методы улучшают метрики ROUGE [5] во всех случаях, кроме XSum [13] 100-примерного режима, вероятно, из-за специфики выборки как задачи экстремальной абстрактивной суммаризации. Так как метод «DiffROUGE-seq» оказывается наилучшим в большинстве случаев, мы рекомендуем отдавать предпочтение ему, однако, если в используемых данных короткие тексты, метод «DiffROUGE-sim» может оказаться лучше. Поставленные цели и задачи исследования были достигнуты полностью. Потенциальные направления дальнейших исследований включают полную оптимизацию гиперпараметров предложенных методов (N , $|X_U|$, λ , β и т.п.), использование другой базовой модели вместо BART-large [2], а также использование дифференцируемой аппроксимации метрики ROUGE-L [5] в функции потерь.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Vaswani A., Shazeer N., Parmar N., Uszkoreit J. Jones L., Gomez A. N., Kaiser Ł., Polosukhin I. Attention is all you need, *Proceedings of the 31st International Conference on Neural Information Processing Systems*, 2017, pp. 6000-6010. DOI: 10.5555/3295222.3295349.
2. Lewis M., Liu Y., Goyal N., Ghazvininejad M., Mohamed, A., Levy O., Stoyanov V., Zettlemoyer L. BART: Denoising Sequence-to-Sequence Pre-training for Natural Language Generation, Translation, and Comprehension, *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, 2020, pp. 7871-7880. DOI: 10.18653/v1/2020.acl-main.703.
3. Zhang J., Zhao Y., Saleh M., Liu P. J. PEGASUS: pre-training with extracted gap-sentences for abstractive summarization, *Proceedings of the 37th International Conference on Machine Learning*, 2020, pp. 7871-7880. DOI: 10.5555/3524938.3525989.
4. Raffel C., Shazeer N., Roberts A., Lee K., Narang S., Matena M., Zhou Y., Li W., Liu P. J. Exploring the limits of transfer learning with a unified text-to-text transformer, *The Journal of Machine Learning Research*, 2020, Vol. 21, No. 1, pp. 5485-5551. DOI: 10.5555/3455716.3455856.
5. Lin C.-Y. ROUGE: A Package for Automatic Evaluation of Summaries, *Text Summarization Branches Out*. Association for Computational Linguistics, 2004, pp. 74-81.
6. Tsvigun A., Lysenko I., Sedashov D., Lazichny I., Damirov E., Karlov V., Belousov A., Sanochkin L., Panov M., Panchenko A., Burtsev M., Shelmanov A. Active Learning for Abstractive Text Summarization, *Findings of the Association for Computational Linguistics: EMNLP 2022*. Association for Computational Linguistics, 2022, pp. 5128-5152. DOI: 10.18653/v1/2022.findings-emnlp.377.
7. Wan D., Bansal M. FactPEGASUS: Factuality-Aware Pre-training and Fine-tuning for Abstractive Summarization, *Proceedings of the 2022 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*. Association for Computational Linguistics, 2022, pp. 1010-1028. DOI: 10.18653/v1/2022.naacl-main.74.
8. Chen, Y.-S., Shuai, H.-H. Meta-Transfer Learning for Low-Resource Abstractive Summarization, *Proceedings of the AAAI Conference on Artificial Intelligence*, 2021, Vol. 35, No.14, pp. 12692-12700. DOI: 10.1609/aaai.v35i14.17503.
9. Chen Y., Liu Y., Xu R., Yang Z., Zhu C., Zeng M., Zhang Y. UniSumm and SummZoo: Unified Model and Diverse Benchmark for Few-Shot Summarization, *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Vol. 1: Long Papers)*. Association for Computational Linguistics, 2023, pp. 12833-12855. DOI: 10.18653/v1/2023.acl-long.718.
10. Ravaut M., Joty S., Chen N. Towards Summary Candidates Fusion, *Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing*. Association for Computational Linguistics, 2022, pp. 8488-8504. DOI: 10.18653/v1/2022.emnlp-main.581.

11. Zhang R., Tetreault J. This Email Could Save Your Life: Introducing the Task of Email Subject Line Generation, *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*. Association for Computational Linguistics, 2019, pp. 446-456. DOI: 10.18653/v1/P19-1043.
12. Napoles C., Gormley M., Durme B.V. Annotated Gigaword, *Proceedings of the Joint Workshop on Automatic Knowledge Base Construction and Web-scale Knowledge Extraction (AKBC-WEKEX)*. Association for Computational Linguistics, 2012, pp. 95-100.
13. Narayan S., Cohen S. B., Lapata M. Don't Give Me the Details, Just the Summary! Topic-Aware Convolutional Neural Networks for Extreme Summarization, *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*. Association for Computational Linguistics, 2018, pp. 1797-1807. DOI: 10.18653/v1/D18-1206.
14. Kim B., Kim H., Kim G. Abstractive Summarization of Reddit Posts with Multi-level Memory Networks, *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Vol. 1 (Long and Short Papers)*. Association for Computational Linguistics, 2019, pp. 2519-2531. DOI: 10.18653/v1/N19-1260.
15. Xie Q., Dai Z., Hovy E., Luong M.-T., Le Q. V. Unsupervised data augmentation for consistency training, *Proceedings of the 34th International Conference on Neural Information Processing Systems*, 2020. DOI: 10.5555/3495724.3496249.
16. Shao C., Chen X., Feng Y. Greedy Search with Probabilistic N-gram Matching for Neural Machine Translation, *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*. Association for Computational Linguistics, 2018, pp. 4778-4784. DOI: 10.18653/v1/D18-1510.
17. Papineni K., Roukos S., Ward T., Zhu W.-J. BLEU: a method for automatic evaluation of machine translation, *Proceedings of the 40th Annual Meeting on Association for Computational Linguistics*. Association for Computational Linguistics, 2002, pp. 311-318.
18. Chung H.W., Hou L., Longpre S., Zoph B., Tai Y., Fedus W., Li Y., Wang X., Dehghani M., Brahma S., Webson A., Gu S.S., Dai Z., Suzgun M., Chen X., Chowdhery A., Castro-Ros A., Pellat M., Robinson K., Valter D., Narang S., Mishra G., Yu A., Zhao V., Huang Y., Dai A., Yu H., Petrov S., Chi E.H., Dean J., Devlin J., Roberts A., Zhou D., Le Q.V., Wei J. Scaling instruction-finetuned language models, *The Journal of Machine Learning Research*, 2024, Vol. 25, No. 1, pp. 3381-3433. DOI: 10.5555/3722577.3722647.
19. Tianyi Z., Varsha K., Felix W., Kilian Q. W., Yoav A. BERTScore: Evaluating Text Generation with BERT, *Proceedings of the 40th Annual Meeting on Association for Computational Linguistics*. Association for Computational Linguistics, 2020, pp. 311-318.
20. Yuan W., Neubig G., Liu P. BARTSCORE: evaluating generated text as text generation, *Proceedings of the 35th International Conference on Neural Information Processing Systems*, 2021. DOI: 10.5555/3540261.3542349.
21. Wolf T., Debut L., Sanh V., Chaumond J., Delangue C., Moi A., Cistac P., Rault T., Louf R., Funtowicz M., Davison J., Shleifer S., von Platen P., Ma C., Jernite Y., Plu J., Xu C., Le Scao T., Gugger S., Drame M., Lhoest Q., Rush A. Transformers: State-of-the-Art Natural Language Processing, *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing: System Demonstrations*. Association for Computational Linguistics, 2020, pp. 38-45. DOI: 10.18653/v1/2020.emnlp-demos.6.
22. Ansel J., Yang E., He H., Gimelshein N., Jain A., Voznesensky M., Bao B., Bell P., Berard D., Burovski E., Chauhan G., Chourdia A., Constable W., Desmaison A., DeVito Z., Ellison E., Feng W., Gong J., Gschwind M., Hirsh B., Huang S., Kalambarkar K., Kirsch, L., Lazos M., Lezcano M., Liang Y., Liang J., Lu Y., Luk CK, Maher B., Pan Y., Puhersch C., Reso M., Saroufim M., Siraichi M. Y., Suk H., Suo M., Tillet P., Wang E., Wang X., Wen W., Zhang S., Zhao X., Zhou K., Zou R., Mathews A., Chanan G., Wu P., Chintala S. PyTorch 2: Faster Machine Learning Through Dynamic Python Bytecode Transformation and Graph Compilation, *29th ACM International Conference on Architectural Support for Programming Languages and Operating Systems, Vol. 2 (ASPLOS '24)*. ACM, 2024. DOI: 10.1145/3620665.3640366.
23. Lhoest Q., Villanova del Moral A., Jernite Y., Thakur A., von Platen P., Patil S., Chaumond J., Drame M., Plu J., Tunstall L., Davison J., Šaško M., Chhablani G., Malik B., Brandeis S., Le Scao T., Sanh V., Xu C., Patry N., McMillan-Major, J., Schmid P., Gugger S., Delangue C., Matussière T., Debut L., Bekman S., Cistac P., Goehringer T., Mustar V., Lagunas F., Rush A., Wolf T. Datasets: A Community Library for Natural Language Processing, *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing: System Demonstrations*. Association for Computational Linguistics, 2021, pp. 175-184. DOI: 10.18653/v1/2021.emnlp-demo.21.
24. Boyd S., Vandenberghe L. Convex optimization. Cambridge university press, 2004.

Лысенко Иван Егорович – Финансовый университет при Правительстве Российской Федерации; e-mail: 259996@edu.fa.ru; г. Москва, Россия; кафедра искусственного интеллекта; аспирант; <https://orcid.org/0009-0003-1990-021X>.

Lysenko Ivan Egorovich – Financial University under the Government of the Russian Federation; e-mail: 259996@edu.fa.ru; Moscow, Russia; Artificial Intelligence department; postgraduate student; <https://orcid.org/0009-0003-1990-021X>.

С.И. Носков, А.П. Медведев, И.Д. Кириллов

**ПРИМЕНЕНИЕ МЕТОДА СМЕШАННОГО ОЦЕНИВАНИЯ ПАРАМЕТРОВ
ПРИ ПОСТРОЕНИИ ОДНОРОДНОЙ ВЛОЖЕННОЙ КУСОЧНО-ЛИНЕЙНОЙ
РЕГРЕССИИ ПЕРВОГО ТИПА**

Статья посвящена разработке алгоритма идентификации параметров однородной вложенной кусочно-линейной регрессии первого типа – модели, востребованной при анализе сложных систем, поведение которых не удаётся адекватно описать гладкими функциями. В обзорной части работы систематизированы современные публикации, иллюстрирующие применение кусочно-линейных форм в самых разных предметных областях: от моделирования энергопотребления и нелинейных систем управления до задач обработки изображений, восстановления генетических сетей и фильтрации геофизических данных. Новизна исследования заключается в том, что впервые для данного класса моделей предложен алгоритм идентификации, основанный на методе смешанного оценивания (МСО), позволяющий гибко комбинировать два различных критерия качества. Показано, что благодаря введению дополнительных булевых переменных и вспомогательных ограничений исходная оптимизационная задача сводится к стандартной задаче линейно-булева программирования, что позволяет применять для её решения доступные численные методы. Эффективность разработанного алгоритма продемонстрирована на реальных данных горно-металлургической компании «Норникель» за 2010–2024 годы. В качестве зависимой переменной используется выручка, в качестве предикторов – объёмы добычи никеля, палладия, меди и платины. Построены две альтернативные модели – классическим методом наименьших модулей и предлагаемым методом смешанного оценивания. Сравнительный анализ показывает, что вторая модель имеет несколько более высокую среднюю процентную ошибку, однако существенно выигрывает в величине максимальной ошибки на контрольной подвыборке, что делает её предпочтительной в условиях, где критичны выбросы. Полученные результаты подтверждают практическую ценность МСО для построения интерпретируемых и устойчивых регрессионных зависимостей.

Идентификация параметров; однородная вложенная кусочно-линейная регрессия; методы наименьших модулей и смешанного оценивания; задача линейно-булева программирования; горно-металлургическая компания «Норникель».

S.I. Noskov, A.P. Medvedev, I.D. Kirillov

**APPLICATION OF THE MIXED PARAMETER ESTIMATION METHOD
IN THE CONSTRUCTION OF A HOMOGENEOUS NESTED PIECEWISE LINEAR
REGRESSION OF THE FIRST TYPE**

The paper is devoted to the development of an algorithm for identifying the parameters of a homogeneous nested piecewise linear regression of the first type – a model that is in demand when analyzing complex systems whose behavior cannot be adequately described by smooth functions. The review part of the work systematizes modern publications illustrating the use of piecewise linear forms in various subject areas: from modeling energy consumption and nonlinear control systems to image processing, reconstruction of genetic networks, and filtering of geophysical data. The novelty of the study lies in the fact that for the first time for this class of models an identification algorithm based on the mixed estimation method (MEM) is proposed, which allows flexible combination of two different quality criteria. It is shown that by introducing additional Boolean variables and auxiliary constraints, the original optimization problem is reduced to a standard linear Boolean programming problem, which makes it possible to use available numerical methods for its solution. The effectiveness of the developed algorithm is demonstrated on real data of the mining and metallurgical company Norilsk Nickel for 2010–2024. The dependent variable is revenue, and the predictors are the production volumes of nickel, palladium, copper and platinum. Two alternative models are constructed – using the classical least absolute deviations method and the proposed mixed estimation method. A comparative analysis shows that the second model has a slightly higher average percentage error, but significantly outperforms in the magnitude of the maximum error on the control subsample, which makes it preferable in conditions where outliers are critical. The results obtained confirm the practical value of MEM for constructing interpretable and robust regression dependencies.

Identification of parameters; homogeneous nested piecewise linear regression; methods of least modules and mixed estimation; linear Boolean programming problem; mining and metallurgical company Norilsk Nickel.

Введение. Характер функционирования исследуемого объекта часто предполагает при его анализе методами математического моделирования возможность использования наряду с гладкими различными кусочно-линейными модельными форм. Так, в работе [1] исследуется, как меняется потребление электроэнергии в жилых помещениях по мере увеличения доходов в Китае. При этом вместо обычно применяемых линейных или квадратичных двойных логарифмических моделей с наложением предварительных ограничений на нечеткую взаимосвязь между потреблением электроэнергии и доходом используется уменьшенная кусочно-линейная модель, которая является самоадаптируемой, очень гибкой и обходит эту проблему. В исследовании [2] рассматривается новая задача комбинаторной оптимизации, связанная с линейными системами (MIN PFS), которая заключается в разбиении системы на минимальное число допустимых подсистем. MIN PFS позволяет формализовать фундаментальную задачу оценки кусочно-линейной модели, что является привлекательной альтернативой при моделировании широкого спектра нелинейных явлений. В [3] представлена нелинейная модель прогнозирующего управления, основанная на модели Винера с кусочно-линейным коэффициентом усиления. Этот подход сохраняет все свойства классической линейной модели прогнозирующего управления и упрощает выполнение вычислений благодаря канонической структуре этого нелинейного коэффициента. Обсуждаются некоторые рекомендации по определению номинальной модели, а также границ ее неопределенности. В статье [4] представлен основанный на использовании кусочно-линейной модели новый метод повышения резкости изображений с помехами. При этом повышается резкость по краям изображения и уменьшается шум. В работе [5] рассматривается задача подгонки выпуклой кусочно-линейной функции к заданным многомерным данным. Описан метод ее решения, представляющий собой вариацию алгоритма кластеризации с использованием k -средних.

В статье [6] разрабатывается несколько формулировок задач нелинейного программирования со смешанными целыми числами, использующих многомерные кусочно-линейные модели для аппроксимации нелинейных функций областями, объединенными в гиперкубы и симплексы. В исследовании [7] представлено несколько альтернативных вариантов выбора и комбинирования моделей, которые позволяют получить непрерывную кусочно-линейную регрессионную модель из крупномасштабных наборов данных с использованием алгоритма обучения, основанного на так называемых линейных шарнирах. Работа [8] посвящена описанию нового алгоритма построения автоматической кусочно-линейной регрессии, который предполагает повышение интерпретируемости модели, а также учитывает нелинейные взаимосвязи и возможные условия взаимодействия предикторов. В [9] моделируется активность генов с помощью кусочно-линейных дифференциальных уравнений, формируемых путем встраивания логической коммутационной сети в само уравнение. При этом генерируются непрерывные нелинейные уравнения с заменой разрывов ступенчатой функции в кусочно-линейных уравнениях сигмоидальными управляющими функциями. В [10] показано, что кусочно-дифференцируемые функции могут быть локально аппроксимированы кусочно-линейной моделью, основанной на паре точек выборки данных. Доказано, что расхождение между функцией и моделью в любой точке имеет билинейный порядок. При реализации процедуры кусочной линеаризации использован обобщенный метод Ньютона.

Следует также отметить следующие интересные работы, касающиеся применения кусочно-линейных модельных форм при исследовании сложных систем: [11] (кусочно-линейная аппроксимация пространства параметров модели для восстановления диаграммы динамической системы), [12] (алгоритм фильтрации шумов, основанный на двумерных скользящих локальных кусочно-линейных моделях), [13] (применение кусочно-линейного стохастического отображения при моделировании электрической активности уединенного нейрона), [14] (алгоритмы построения кусочно-линейных аппроксимационных моделей для обработки геомагнитных наблюдений), [15] (применение кусочно-линейной функции при оценке технической готовности транспортного средства), [16] (кусочно-линейные модели зависимости качества воды от физико-химических параметров источника водоснабжения).

Целью настоящей работы является разработка алгоритма идентификации параметров однородной вложенной кусочно-линейной регрессии (ОВКЛР) первого типа с помощью метода смешанного оценивания (МСО).

Разработка однородной вложенной кусочно-линейной регрессионной модели на основе использования метода смешанного оценивания. В работах [17, 18] наряду с некоторыми другими вложенными регрессионными формами введена однородная вложенная кусочно-линейная регрессия первого типа, имеющая вид:

$$y_k = \min\{\min_{i \in I^1}\{a_i^1 x_{ki}\}, \dots, \min_{i \in I^G}\{a_i^G x_{ki}\}\} + \varepsilon_k, k = \overline{1, n}. \quad (1)$$

Здесь k – номер наблюдения, y – зависимая переменная, $x_i, i = \overline{1, m}$ – независимые переменные, $a_i^j, j = \overline{1, G}$ – идентифицируемые параметры, $\varepsilon_k, k = \overline{1, n}$ – ошибки приближения, n – длина выборки данных. Множества $I^i, i = \overline{1, G}$ являются подмножествами множества индексов $\{1, 2, \dots, m\}$ и могут иметь непустые пересечения. Все переменные и ошибки в модели (1) детерминированы.

Определим вектор $a = (a_i^j), j = \overline{1, G}, i \in I^j$.

В работе [19], в случае использования при оценивании параметров ОВКЛР (1) метода наименьших модулей (МНМ) с функцией потерь

$$J_1(a) = \sum_{k=1}^n |\varepsilon_k| \quad (2)$$

предложено сводить задачу ее минимизации к задаче линейно-булева программирования (ЛБП) следующим образом.

Введем новые переменные:

$$v_{kj} = \min_{i \in I^j}\{a_i^j x_{ki}\}, k = \overline{1, n}, j = \overline{1, G},$$

$$w_k = \min_{j=\overline{1, G}} v_{kj}, k = \overline{1, n}.$$

Тогда задача минимизации функции потерь (2) сводится к следующей задаче ЛБП [19]:

$$a_i^j x_{ki} - v_{kj} \leq (1 - s_{kij})M, k = \overline{1, n}, i \in I^j, j = \overline{1, G}, \quad (3)$$

$$\sum_{i \in I^j} s_{kij} = 1, k = \overline{1, n}, j = \overline{1, G}, \quad (4)$$

$$w_k \leq v_{kj}, k = \overline{1, n}, j = \overline{1, G}, \quad (5)$$

$$v_{kj} - w_k \leq (1 - r_{kj})M, k = \overline{1, n}, j = \overline{1, G}, \quad (6)$$

$$\sum_{j=1}^G r_{kj} = 1, k = \overline{1, n}, \quad (7)$$

$$w_k + u_k - b_k = y_k, k = \overline{1, n}, \quad (8)$$

$$u_k \geq 0, b_k \geq 0, k = \overline{1, n}, \quad (9)$$

$$v_{kj} \geq 0, k = \overline{1, n}, j = \overline{1, G}, \quad (10)$$

$$s_{kij} \in \{0, 1\}, k = \overline{1, n}, i \in I^j, j = \overline{1, G}, \quad (11)$$

$$r_{kj} \in \{0, 1\}, k = \overline{1, n}, j = \overline{1, G}, \quad (12)$$

$$\delta \sum_{j=1}^G \sum_{i \in I^j} a_i^j + \sum_{k=1}^n (u_k + b_k) \rightarrow \min. \quad (13)$$

Здесь δ и M – заранее заданные малое и большое положительные числа соответственно.

Поставим задачу идентификации параметров ОВКЛР (1) несколько по-иному. Определим, в соответствии с классической монографией [20], функцию потерь общего вида:

$$J_v(a) = \sum_{k=1}^n |\varepsilon_k|^v, v \geq 1.$$

Разобьем всю выборку данных на p непересекающихся подвыборок с номерами наблюдений из индексных множеств $P_s \subset \{1, 2, \dots, n\}, s = \overline{1, l}$:

$$\bigcup_{s=1}^l P_s = \{1, 2, \dots, n\}, P_i \cap P_j = \emptyset, i \neq j.$$

Преобразуем функцию потерь (2) в следующую:

$$J(\alpha) = \sum_{s=1}^l \sum_{k \in P_s} |\varepsilon_k|^{v_s}, \quad (14)$$

где числа $v_s \in [1, \infty]$ заданы заранее. При этом имеет место известное равенство:

$$J_\infty(\alpha) = \lim_{v \rightarrow \infty} \sum_{k=1}^n |\varepsilon_k|^v = \max_{k=\overline{1,n}} |\varepsilon_k|. \quad (15)$$

Функция потерь (14) соответствует методу смешанного оценивания (МСО) параметров в регрессионном анализе (см. в частности, [21, 22]), а (15) – методу антиробастного оценивания (МАО) (см., например, [23]). Заметим, что в работе [24] описан алгоритм построения простой вложенной кусочно-линейной регрессии с помощью МСО.

Зададим значения $l = 2$, $v_1 = 1$, $v_2 = \infty$. Тогда, как и в [24], задача оценивания параметров ОВКЛР (1) с применением МСО также может быть сведена к задаче ЛБП. Для этого следует дополнить систему ограничений задачи ЛБП (3)–(13) условием

$$u_k + b_k \leq \tau, k \in P_2, \quad (16)$$

а целевую функцию (3) заменить на следующую

$$\sum_{k \in P_1} (u_k + b_k) / s + \tau + \eta \sum_{k \in P_2} (u_k + b_k) + \delta \sum_{j=1}^G \sum_{i \in I_j} a_i^j \rightarrow \min, \quad (17)$$

где s – число элементов в множестве P_1 , η – заранее заданная малая положительная константа.

Однородная вложенная кусочно-линейная модель горно-металлургической компании «Норникель». Используем описанный алгоритм для построения модели выручки горно-металлургической компании «Норникель». В качестве информационной базы моделирования привлечем статистические данные за 2010–2024 гг. (табл. 1) [25, 26]. Введем обозначения: y – выручка, млрд. руб.; x_1 – объем добычи никеля, тыс. т.; x_2 – объем добычи палладия, т.; x_3 – объем добычи меди, тыс. т.; x_4 – объем добычи платины, т.

Таблица 1

Статистические данные

Год	x_1	x_2	x_3	x_4	y
2010	295.0	87.3	378.0	21.6	388.0
2011	300.0	84.9	364.0	21.2	415.1
2012	285.0	82.8	371.0	20.2	384.5
2013	266.0	83.6	369.0	20.4	366.2
2014	274.0	85.6	368.0	20.6	456.0
2015	266.4	83.6	369.4	20.4	506.1
2016	235.7	81.4	360.2	20.0	548.8
2017	217.1	86.5	401.1	20.8	536.8
2018	218.8	84.9	473.7	20.3	728.9
2019	228.7	90.9	499.1	21.8	877.8
2020	235.7	87.9	487.2	21.6	1117.0
2021	193.0	81.4	406.8	19.9	1317.0
2022	219.0	86.8	433.0	20.2	1185.0
2023	208.6	83.7	425.4	20.7	1172.0
2024	205.0	83.7	433.0	20.7	1166.0

Примем $G = 2$, $I^1 = \{1, 4\}$, $I^2 = \{2, 3\}$ и зададим составы индексных множеств P_1 и P_2 :

$$P_1 = \{1, 2, 3, 4, 5, 6, 7\}, P_2 = \{8, 9, 10, 11, 12, 13, 14, 15\}.$$

Вначале построим ОВКЛР с помощью МНМ путем решения задачи ЛБП (3)–(13):

$$y_k = \min\{\min\{3.38x_{k1}, 35.9x_{k4}\}, \min\{8.58x_{k2}, 1.6x_{k3}\}\} + \varepsilon_k, k = \overline{1,15}, \quad (18)$$

$$E = 8.7\%, \tau(P_2) = 686.9.$$

Здесь E – средняя процентная ошибка:

$$E = 100\% \sum_{k=1}^{15} |\varepsilon_k| / \sum_{k=1}^{15} y_k,$$

$\tau(P_2)$ – максимальная по модулю ошибка на множестве P_2 :

$$\tau(P_2) = \max_{k \in P_2} |\varepsilon_k|.$$

Теперь построим модель (1) с помощью МСО, решив задачу ЛБП (3)–(12), (16), (17):

$$y_k = \min\{\min\{4.2x_{k1}, 60.1x_{k4}\}, \min\{14.9x_{k2}, 1.3x_{k3}\}\} + \varepsilon_k, k = \overline{1,15}, \quad (19)$$

$$E = 10.3\%, \tau(P_2) = 104.5.$$

Отметим существенное расхождение в оценках параметров моделей (18) и (19), в частности, усиление значимости объема добычи меди (x_3) при уменьшении влияния остальных предикторов. Модель (19) несколько уступает по средней процентной ошибке модели (18), но значительно ее превосходит по максимальной ошибке на второй половине выборки. Последнее обстоятельство может стать решающим при окончательном выборе модели.

Заключение. В работе предложен алгоритмический способ идентификации параметров однородной вложенной кусочно-линейной регрессии первого типа на основе использования метода смешанного оценивания. Этот способ предполагает сведение поставленной оптимизационной задачи к задаче линейно-булева программирования. Построены два альтернативных варианта модели данного класса для горно-металлургической компании «Норникель» – методами наименьших модулей и смешанного оценивания, для каждого из которых рассчитаны значения двух критериев адекватности.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Liu Y., Gao Y., Hao X., Liao H. The Relationship between Residential Electricity Consumption and Income: A Piecewise Linear Model with Panel Data // *Energies*. – 2016. – Vol. 9, No. 10. – P. 831. – DOI: 10.3390/en9100831.
2. Amaldi E., Mattavelli M. The MIN PFS problem and piecewise linear model estimation // *Discrete Applied Mathematics*. – 2002. – Vol. 118, Issues 1–2. – P. 115–143. – DOI: 10.1016/S0166-218X(01)00260-8.
3. Cervantes A.L., Agamennoni O.E., Figueroa J.L. A nonlinear model predictive control system based on Wiener piecewise linear models // *Journal of Process Control*. – 2003. – Vol. 13, No. 7. – P. 655–666. – DOI: 10.1016/S0959-1524(02)00121-X.
4. Russo F. Piecewise Linear Model-Based Image Enhancement // *EURASIP Journal on Advances in Signal Processing*. – 2004. – Article No. 983173. – DOI: 10.1155/S1110865704404041.
5. Magnani A., Boyd S.P. Convex piecewise-linear fitting // *Optimization and Engineering*. – 2009. – Vol. 10. – P. 1–17. – DOI: 10.1007/s11081-008-9045-3.
6. Silva T.L., Camponogara E.A. Computational analysis of multidimensional piecewise-linear models with applications to oil production optimization // *European Journal of Operational Research*. – 2014. – Vol. 232, No. 3. – P. 630–642. – DOI: 10.1016/j.ejor.2013.07.040.
7. Moreno-Carbonell S., Sánchez-Úbeda F. Piecewise Linear Regression Model Ensemble for Large-Scale Curve Fitting // *Algorithms*. – 2024. – Vol. 17, No. 4. – 147 p.
8. von Ottenbreit M., De Bin R. Automatic piecewise linear regression // *Computational Statistics*. – 2024. – Vol. 39. – P. 1867–1907.
9. Shahrear P., Glass L., Wilds R., Edwards R. Dynamics in piecewise linear and continuous models of complex switching networks // *Mathematics and Computers in Simulation*. – 2015. – Vol. 110. – P. 33–39.
10. Griewank A., Streubel T., Lehmann L., Radons M., Hasenfelder R. Piecewise linear secant approximation via algorithmic piecewise differentiation // *Optimization Methods and Software*. – 2018. – Vol. 33, Issues 4–6.
11. Бозатиков Е.В., Битюцкая Л.А., Бормонтов Е.Н., Хухрянский М.Ю. Реконструкция бифуркационной диаграммы нелинейной динамической системы по стационарным и нестационарным временным рядам // *Вестник Воронежского государственного университета. Серия: Физика. Математика*. – 2006. – № 2. – С. 24–35.

12. Гетманов В.Г., Гвишиани А.Д., Сидоров и др. Фильтрация наблюдений угловых распределений мюонных потоков от годоскопа «Ураган» // Ядерная физика и инжиниринг. – 2017. – Т. 8, № 6. – С. 506-512.
13. Андреев К.В., Красичков Л.В. Моделирование электрической активности нейрона с помощью кусочно-непрерывных отображений // Письма в Журнал технической физики. – 2003. – Т. 29, № 3. – С. 46-52.
14. Гетманов В.Г. Локальные и сплайновые аппроксимации в цифровой обработке геомагнитных наблюдений // Чебышевский сборник. – 2018. – Т. 19, № 4 (68). – С. 26-42.
15. Егоров С.Я., Салих Х.С., Затонский А.В. Имитационная модель технической готовности крупного автопарка // Вестник Южно-Уральского государственного университета. Серия: Компьютерные технологии, управление, радиоэлектроника. – 2020. – Т. 20, № 2. – С. 14-25.
16. Бубыр Д.С. Система раннего предупреждения о нарушении показателей качества питьевой воды // Программные продукты и системы. – 2015. – № 2. – С. 119-123.
17. Носков С.И. Подход к формализации вложенной кусочно-линейной регрессии // Международный журнал гуманитарных и естественных наук. – 2023. – № 1-2 (76). – С. 218-220.
18. Носков С.И. Некоторые формы вложенной кусочно-линейной регрессии // Известия Тульского государственного университета. Технические науки. – 2023. – № 3. – С. 467-469.
19. Носков С.И., Белинская С.И. Вычисление оценок параметров однородной вложенной кусочно-линейной регрессии // Вестник Дагестанского государственного технического университета. Технические науки. – 2023. – Т. 50, № 4. – С. 115-120.
20. Демиденко Е.З. Линейная и нелинейная регрессии. – М.: Финансы и статистика, 1981. – 302 с.
21. Носков С.И. О методе смешанного оценивания параметров линейной регрессии // Информационные технологии и математическое моделирование в управлении сложными системами. – 2019. – № 1 (2). – С. 41-45.
22. Носков С.И. Метод смешанного оценивания параметров линейной регрессии: особенности применения // Вестник Воронежского государственного университета. Серия: Системный анализ и информационные технологии. – 2021. – № 1. – С. 126-132.
23. Носков С.И. Метод антиробастного оценивания параметров линейной регрессии: число максимальных по модулю ошибок аппроксимации // Южно-Сибирский научный вестник. – 2020. – № 1 (29). – С. 51-54.
24. Носков С.И., Знайдюк А.Н. Вычисление параметров простой формы вложенной кусочно-линейной регрессии методом смешанного оценивания // Вестник кибернетики. – 2024. – Т. 23, № 1. – С. 54-59.
25. Норникель: финансовая отчетность // Официальный сайт ПАО «ГМК «Норильский никель». – Режим доступа: <https://nornickel.ru/investors/disclosure/financials/#accordion-2014> (дата обращения: 15.11.2025).
26. ГМК Норникель (GMKN): годовая финансовая отчетность МСФО // Информационный портал Smart-Lab. – Режим доступа: <https://smart-lab.ru/q/GMKN/f/y> (дата обращения: 15.11.2025).

REFERENCES

1. Liu Y., Gao Y., Hao X., Liao H. The Relationship between Residential Electricity Consumption and Income: A Piecewise Linear Model with Panel Data, *Energies*, 2016, Vol. 9, No. 10, pp. 831. DOI: 10.3390/en9100831.
2. Amaldi E., Mattavelli M. The MIN PFS problem and piecewise linear model estimation, *Discrete Applied Mathematics*, 2002, Vol. 118, Issues 1–2, pp. 115-143. DOI: 10.1016/S0166-218X(01)00260-8.
3. Cervantes A.L., Agamennoni O.E., Figueroa J.L. A nonlinear model predictive control system based on Wiener piecewise linear models, *Journal of Process Control*, 2003, Vol. 13, No. 7, pp. 655-666. DOI: 10.1016/S0959-1524(02)00121-X.
4. Russo F. Piecewise Linear Model-Based Image Enhancement, *EURASIP Journal on Advances in Signal Processing*, 2004, Article No. 983173. DOI: 10.1155/S1110865704404041.
5. Magnani A., Boyd S.P. Convex piecewise-linear fitting, *Optimization and Engineering*, 2009, Vol. 10, pp. 1-17. DOI: 10.1007/s11081-008-9045-3.
6. Silva T.L., Camponogara E.A. Computational analysis of multidimensional piecewise-linear models with applications to oil production optimization, *European Journal of Operational Research*, 2014, Vol. 232, No. 3, pp. 630-642. DOI: 10.1016/j.ejor.2013.07.040.
7. Moreno-Carbonell S., Sánchez-Úbeda F. Piecewise Linear Regression Model Ensemble for Large-Scale Curve Fitting, *Algorithms*, 2024, Vol. 17, No. 4, 147 p.
8. von Ottenbreit M., De Bin R. Automatic piecewise linear regression, *Computational Statistics*, 2024, Vol. 39, pp. 1867-1907.
9. Shahrear P., Glass L., Wilds R., Edwards R. Dynamics in piecewise linear and continuous models of complex switching networks, *Mathematics and Computers in Simulation*, 2015, Vol. 110, pp. 33-39.

10. Griewank A., Streubel T., Lehmann L., Radons M., Hasenfelder R. Piecewise linear secant approximation via algorithmic piecewise differentiation, *Optimization Methods and Software*, 2018, Vol. 33, Issues 4–6.
11. Bogatkov E.V., Bityutskaya L.A., Bormontov E.N., Khukhryansky M.Yu. Rekonstruktsiya bifurkatsionnoy diagrammy nelineinoy dinamicheskoy sistemy po statsionarnym i nestatsionarnym vremennym ryadam [Reconstruction of the bifurcation diagram of a nonlinear dynamical system from stationary and non-stationary time series], *Vestnik Voronezhskogo gosudarstvennogo universiteta. Seriya: Fizika. Matematika* [Bulletin of Voronezh State University. Series: Physics. Mathematics], 2006, No. 2, pp. 24-35.
12. Getmanov V.G., Gvishiani A.D., Sidorov i dr. Fil'tratsiya nablyudeniy uglovykh raspredeleniy myuonnykh potokov ot godoskopa «Uragan» [Filtration of observations of angular distributions of muon fluxes from the "Uragan" hodoscope], *Yadernaya fizika i inzhiniring* [Nuclear Physics and Engineering], 2017, Vol. 8, No. 6, pp. 506-512.
13. Andreev K.V., Krasichkov L.V. Modelirovanie elektricheskoy aktivnosti neyrona s pomoshch'yu kusochno-nepreryvnykh otobrazheniy [Modeling the electrical activity of a neuron using piecewise continuous mappings], *Pis'ma v Zhurnal tekhnicheskoy fiziki* [Technical Physics Letters], 2003, Vol. 29, No. 3, pp. 46-52.
14. Getmanov V.G. Lokal'nye i splaynovye approksimatsii v tsifrovoy obrabotke geomagnitnykh nablyudeniy [Local and Spline Approximations in Digital Processing of Geomagnetic Observations], *Chebyshevskiy sbornik* [Chebyshevskii Collection], 2018, Vol. 19, No. 4 (68), pp. 26-42.
15. Egorov S.Ya., Salikh Kh.S., Zatonskiy A.V. Imitatsionnaya model' tekhnicheskoy gotovnosti krupnogo avtoparka [Simulation model of the technical readiness of a large vehicle fleet], *Vestnik Yuzhno-Ural'skogo gosudarstvennogo universiteta. Seriya: Komp'yuternye tekhnologii, upravlenie, radioelektronika* [Bulletin of South Ural State University. Series: Computer Technologies, Control, Radio Electronics], 2020, Vol. 20, No. 2, pp. 14-25.
16. Bubyry' D.S. Sistema rannego preduprezhdeniya o narushenii pokazateley kachestva pit'evoy vody [Early warning system for drinking water quality indicators violation], *Programmnye produkty i sistemy* [Software & Systems], 2015, No. 2, pp. 119-123.
17. Noskov S.I. Podkhod k formalizatsii vlozhennoy kusochno-lineynoy regressii [An approach to formalizing nested piecewise linear regression], *Mezhdunarodnyy zhurnal gumanitarnykh i estestvennykh nauk* [International Journal of Humanities and Natural Sciences], 2023, No. 1–2 (76), pp. 218-220.
18. Noskov S.I. Nekotorye formy vlozhennoy kusochno-lineynoy regressii [Some Forms of Nested Piecewise Linear Regression], *Izvestiya Tul'skogo gosudarstvennogo universiteta. Tekhnicheskie nauki* [News of the Tula State University. Technical Sciences], 2023, No. 3, pp. 467-469.
19. Noskov S.I., Belinskaya S.I. Vychislenie otsenok parametrov odnorodnoy vlozhennoy kusochno-lineynoy regressii [Calculation of parameter estimates for homogeneous nested piecewise linear regression], *Vestnik Dagestanskogo gosudarstvennogo tekhnicheskogo universiteta. Tekhnicheskie nauki* [Bulletin of the Dagestan State Technical University. Technical Sciences], 2023, Vol. 50, No. 4, pp. 115-120.
20. Demidenko E.Z. Lineynaya i nelineynaya regressii [Linear and Nonlinear Regression]. Moscow: Finansy i statistika, 1981, 302 p.
21. Noskov S.I. O metode smeshannogo otsenivaniya parametrov lineynoy regressii [On the method of mixed estimation of linear regression parameters], *Informatsionnye tekhnologii i matematicheskoe modelirovanie v upravlenii slozhnyimi sistemami* [Information Technologies and Mathematical Modeling in Management of Complex Systems], 2019, No. 1 (2), pp. 41-45.
22. Noskov S.I. Metod smeshannogo otsenivaniya parametrov lineynoy regressii: osobennosti primeneniya [Method of Mixed Estimation of Linear Regression Parameters: Application Features], *Vestnik Voronezhskogo gosudarstvennogo universiteta. Seriya: Sistemnyy analiz i informatsionnye tekhnologii* [Bulletin of Voronezh State University. Series: Systems Analysis and Information Technologies], 2021, No. 1, pp. 126-132.
23. Noskov S.I. Metod antirobastnogo otsenivaniya parametrov lineynoy regressii: chislo maksimal'nykh po modulyu oshibok approksimatsii [Method of Anti-Robust Estimation of Linear Regression Parameters: Number of Maximum Absolute Approximation Errors], *Yuzhno-Sibirskiy nauchnyy vestnik* [South-Siberian Scientific Bulletin], 2020, No. 1 (29), pp. 51-54.
24. Noskov S.I., Znaydyuk A.N. Vychislenie parametrov prostoy formy vlozhennoy kusochno-lineynoy regressii metodom smeshannogo otsenivaniya [Calculation of Parameters for a Simple Form of Nested Piecewise Linear Regression by the Mixed Estimation Method], *Vestnik kibernetiki* [Cybernetic Bulletin], 2024, Vol. 23, No. 1, pp. 54-59.
25. Normikel': finansovaya otchetnost' [Normikel: Financial Statements], *Ofitsial'nyy sayt PAO «GMK «Noril'skiy nikel'»* [Official website of PJSC MMC Norilsk Nickel]. Available at: <https://normikel.ru/investors/disclosure/financials/#accordion-2014> (accessed 15 November 2025).
26. GMK Normikel' (GMKN): godovaya finansovaya otchetnost' MSFO [Norilsk Nickel (GMKN): Annual IFRS Financial Statements], *Informatsionnyy portal Smart-Lab* [Smart-Lab Information Portal]. Available at: <https://smart-lab.ru/q/GMKN/f/y> (accessed 15 November 2025).

Носков Сергей Иванович – Иркутский государственный университет путей и сообщения; e-mail: sergey.noskov.57@mail.ru; г. Иркутск, Россия; кафедра информационных систем и защиты информации; д.т.н.; профессор.

Медведев Александр Петрович – Иркутский государственный университет путей и сообщения; e-mail: medvedeff.a.p@yandex.ru; г. Иркутск, Россия; кафедра информационных систем и защиты информации; ассистент.

Кириллов Илья Денисович – Иркутский государственный университет путей и сообщения; e-mail: iliakirillov07@gmail.com; г. Иркутск, Россия; тел.: +79247192344; студент.

Noskov Sergey Ivanovich – Irkutsk State Transport University; e-mail: sergey.noskov.57@mail.ru; Irkutsk, Russia; the Department of Information Systems and Information Security; dr. of eng. sc.; professor.

Medvedev Alexander Petrovich – Irkutsk State Transport University; e-mail: medvedeff.a.p@yandex.ru; Irkutsk, Russia; the Department of Information Systems and Informations Security; assistant.

Kirillov Ilya Denisovich – Irkutsk State Transport University; e-mail: iliakirillov07@gmail.com; Irkutsk, Russia; phone: +79247192344; student.

УДК 001.891.57, 519.876.5

DOI 10.18522/2311-3103-2026-4-109-119

Т.Р. Гамберов, Р.Н. Сафин, Э.В. Чеботарева, Т.Г. Цой, Е.А. Магид

ОЦЕНКА АЛГОРИТМА СЛЕДОВАНИЯ МОБИЛЬНОГО РОБОТА TIAGO BASE ЗА ПОЛЬЗОВАТЕЛЕМ С ПРИМЕНЕНИЕМ ФУНКЦИОНАЛЬНЫХ ЦИФРОВЫХ МОДЕЛЕЙ ЧЕЛОВЕКА В СИМУЛЯТОРЕ GAZEBO

Цифровые модели человека (ЦМЧ) становятся важным инструментом в области робототехники, позволяя осуществлять воспроизводимую и контролируемую проверку алгоритмов компьютерного зрения. К числу таких алгоритмов относятся обнаружение, отслеживание и следование за человеком, которые играют ключевую роль в задачах автономной навигации, планирования траектории и безопасного взаимодействия робота с людьми в общих пространствах. При этом стандартные симуляторы, как правило, ограничены по степени реалистичности виртуальных акторов, разнообразию их внешнего вида и вариативности поведения. Такие ограничения снижают достоверность получаемых результатов и затрудняют перенос алгоритмов на реальных роботов. В данной работе представлен набор настраиваемых ЦМЧ для симулятора Gazebo. Разработанные модели человека отличаются вариативностью одежды, антропометрическими параметрами и внешним обликом, а также включают библиотеку анимаций ходьбы, имитирующих различные сценарии движения человека, включая изменение скорости и направления. Для проверки эффективности предложенного подхода были проведены вычислительные эксперименты в Gazebo в реалистичной офисной среде с использованием мобильного робота TIAGo Base с дифференциальным приводом, оснащённого двухмерным лазерным сканером и монокулярной камерой. Эффективность работы алгоритма следования за пользователем с применением новых ЦМЧ оценивалась по средней пройденной дистанции и количеству случаев потери цели (количеству ложных срабатываний). Результаты показали, что предложенные ЦМЧ позволяют воспроизводить сложные условия восприятия, включая окклюзии и динамические препятствия, а также обеспечивают масштабируемую и систематическую проверку алгоритмов. Данные, полученные в результате вычислительных экспериментов, подтверждают значимость разработанных ЦМЧ для тестирования робототехнических систем, ориентированных на взаимодействие с человеком в условиях, приближённых к реальным.

Цифровые модели человека; компьютерное моделирование; компьютерное зрение; ROS; Gazebo.

T.R. Gamberov, R.N. Safin, E.V. Chebotareva, T.G. Tsoy, E.A. Magid

EVALUATING THE TIAGO BASE MOBILE ROBOT'S PERSON-FOLLOWING ALGORITHM USING REALISTIC DIGITAL HUMAN MODELS IN THE GAZEBO SIMULATOR

Digital Human Models (DHMs) are becoming an important tool in robotics, enabling reproducible and controlled validation of computer vision algorithms. Such algorithms include human detection, tracking, and following, which play a key role in autonomous navigation, trajectory planning, and safe robot-

human interaction in shared spaces. However, standard simulators are typically limited in the realism of virtual actors, the diversity of their appearance, and the variability of their behavior. These limitations reduce the reliability of the results and make it harder to transfer algorithms to real-world robots. This paper presents a set of customizable DHMs for the Gazebo simulator. The developed models feature variability in clothing, anthropometric parameters, and appearance, and include a library of walking animations simulating various human movement scenarios, including changes in speed and direction. To evaluate the effectiveness of the proposed approach, virtual experiments were conducted in a realistic office environment using the TIAGo Base mobile robot with differential drive, equipped with a 2D lidar and a monocular camera. The efficiency of the user-following algorithm was evaluated in terms of average distance traveled and the number of target loss events (the number of false positives). The results showed that the proposed DHMs make it possible to reproduce complex perception conditions, including occlusions and dynamic obstacles, while providing scalable and systematic validation of algorithms. The obtained data confirm the significance of the developed DHMs for testing robotic systems oriented toward human interaction in conditions close to real-world settings.

Digital human models; computer modeling; computer vision; ROS; Gazebo.

Введение. Компьютерное зрение играет ключевую роль в робототехнике и автоматизации, поскольку именно оно позволяет системам обнаруживать, отслеживать и интерпретировать присутствие и поведение человека [1, 2]. Сфера применения охватывает широкий круг задач: от автономной навигации в пространствах с высокой плотностью людей [3, 4] до взаимодействия человека и робота в условиях промышленного производства [5, 6]. Так, в задачах одновременной локализации и построения карты (SLAM) [7, 8] отслеживание человека помогает роботам различать динамические и статические объекты [9, 10], что повышает точность построения карт и снижает вероятность ошибок локализации. В медицине мобильные сервисные роботы с системами компьютерного зрения применяются для доставки лекарств и расходных материалов [11], сопровождения пациентов [12] и внутрибольничных транспортных операций, снижая нагрузку на персонал [13, 14]. Стандартизированные сценарии для коридоров, дверных проёмов и зон ожидания позволяют сравнивать алгоритмы в типичных клинических условиях [15]. В складской логистике алгоритмы обнаружения и прогнозирования траекторий движения сотрудников применяются для организации безопасного передвижения роботов [16, 17], что не только минимизирует риск столкновений, но и повышает эффективность рабочих процессов [18, 19]. В сфере услуг роботизированные ассистенты применяют аналогичные методы при навигации в гостиницах, аэропортах и торговых центрах [20]. При этом возникает ряд проблем [21]: необходимо учитывать высокую плотность потока людей [22, 23] и их непредсказуемое поведение [24]. В сфере сельского хозяйства применение машинного обучения и компьютерного зрения позволяет автоматизировать задачи контроля численности грызунов [25] и сбора урожая [26].

Тем не менее получение реальных данных для обучения и тестирования алгоритмов компьютерного зрения требует значительных временных и финансовых ресурсов и нередко сопряжено с этическими ограничениями [27]. Испытания в реальных условиях осложняются целым рядом факторов. Прежде всего, это риски для безопасности человека и оборудования [28], особенно когда робот работает в узких помещениях или рядом с персоналом. Существенную роль играет сложность самой среды: динамические препятствия, изменения освещённости, помехи датчикам и другие факторы создают дополнительные трудности для систем восприятия [29]. Все перечисленные обстоятельства подчёркивают необходимость создания контролируемой и воспроизводимой экспериментальной среды [30], которая позволит моделировать реальные сценарии без угрозы для безопасности и обеспечит возможность масштабирования исследований.

Цифровые модели человека (ЦМЧ) являются одним из наиболее перспективных способов решения обозначенной проблемы [31, 32]. Они позволяют создавать реалистичные виртуальные среды для проверки алгоритмов [33, 34], а симуляторы, такие как Gazebo, обеспечивают удобные и доступные платформы для их интеграции. Применение ЦМЧ даёт возможность систематически оценивать работу алгоритмов компьютерного зрения в различных условиях: от анализа позы [35, 36] и распознавания действий [2, 37] до обработки ситуаций с окклюзиями [9] и реализации задач отслеживания и следования

за человеком [38]. Вместе с тем существующие модели нередко ограничены по разнообразию и степени реалистичности, что не позволяет в полной мере проверить надёжность алгоритмов.

Перед исследованием была поставлена следующая задача: в связи с низкой эффективностью (единственной) стандартной ЦМЧ симулятора Gazebo для оценивания алгоритмов компьютерного зрения необходимо разработать новый набор ЦМЧ. Новые ЦМЧ должны иметь широкую вариативность внешнего облика и поведения. Практическую применимость новых ЦМЧ для задач компьютерного зрения необходимо проверить в вычислительных экспериментах. В качестве бенчмарка для тестирования используется алгоритм отслеживания и следования за человеком в виртуальной офисной среде [39]. Эффективность работы алгоритма с применением новых ЦМЧ оценивалась по двум критериям: способность устойчиво удерживать цель на протяжении всего маршрута (измерялась как средняя пройденная дистанция) и отсутствие ошибок распознавания (измерялось как количество случаев потери цели или количество ложных срабатываний). Более подробно критерии описаны в разделе «Методика экспериментов».

В данной работе представлен набор детализированных функциональных настраиваемых ЦМЧ, созданных для тестирования алгоритмов отслеживания и следования за человеком-пользователем в симуляторе Gazebo. Разработанные модели отличаются вариативностью: предусмотрены разные типы одежды (в том числе брюки, юбка, длинное платье) и несколько характерных сценариев движения (прямолинейное перемещение, зигзагообразные траектории, повороты на 90°). Это позволяет более полно оценивать работу алгоритмов в разнообразных условиях.

Для проверки предложенного подхода использовался алгоритм отслеживания и следования за человеком мобильной платформой TIAGo Base с дифференциальным приводом, оснащённой лидаром и камерой. Проведённые эксперименты показали, что применение ЦМЧ позволяет воспроизводить характерные для реальных сценариев сложности, включая окклюзии и появление динамических препятствий, и тем самым обеспечивая более надёжную валидацию алгоритмов.

Моделирование цифровых моделей человека. В разделе представлена разработка и интеграция ЦМЧ в Gazebo для последующей оценки алгоритмов компьютерного зрения. При проектировании основное внимание уделялось созданию моделей, которые не только визуально напоминают реальных людей, но и обладают характеристиками, необходимыми для корректной работы алгоритмов восприятия. Разработанные ЦМЧ включают: 1) визуальные текстуры и материалы, позволяющие оценивать алгоритмы, работающие с визуальными данными; 2) коллизионную геометрию, обеспечивающую корректное моделирование столкновений и окклюзий; 3) вариативность поведения, реализованную через библиотеку анимаций движения.

Для интеграции моделей использовался формат SDF. Для приближения условий эксперимента к реальной среде был создан набор ЦМЧ с реалистичными пропорциями тела, вариациями одежды и анимациями движения, ориентированными на задачи отслеживания. При построении моделей применялись программные пакеты Blender [40] и MakeHuman [41] (рис. 1). Для повышения реалистичности поведения человека в виртуальной среде была разработана библиотека анимаций, включающая как базовые, так и более сложные движения. Использование скелетной анимации в среде Blender и заимствование отдельных движений из открытых библиотек позволили сформировать набор сценариев, которые имитируют естественное поведение человека. Библиотека включает следующие компоненты:

1. Статические позы: стояние, сидение, жестикуляция.
2. Локомоция: ходьба с различными скоростями, повороты на месте и движение по зигзагообразным траекториям.
3. Взаимодействие с объектами: движения, связанные с достижением определённых целей в окружающей среде: открывание дверей, взаимодействие с мебелью, наклоны за предметами.



Рис. 1. Примеры ЦМЧ с вариациями одежды: брюки (А), юбка (Б) и длинное платье (В)

Методика экспериментов. Цель экспериментов состояла в том, чтобы проверить устойчивость и точность работы алгоритма в различных условиях, включая разные траектории движения человека и варианты одежды. В качестве основы экспериментов был использован алгоритм, предложенный Чеботаревой и соавторами [39]. Он объединяет данные лидара и камеры, что позволяет сгладить ограничения каждого из датчиков и повысить надёжность восприятия в динамических условиях. В предыдущих исследованиях данный алгоритм показал высокую эффективность при работе с базовыми моделями акторов в симуляторе Gazebo. DL-подходы к сопровождению человека в помещениях подтверждают применимость подобных схем [42].

Эксперименты выполнялись с использованием виртуальной модели мобильного робота TIAGo Base компании PAL Robotics в среде Robot Operating System (ROS) и симуляторе Gazebo. Платформа оснащена дифференциальным приводом, что делает её подходящей для задач навигации и отслеживания целей в помещениях, двухмерным лазерным сканером (лидаром) с углом обзора 270° (установлен на высоте 10 см от основания платформы) и монокулярной камерой с углом обзора 58° и разрешением 640×480 пикселей (закреплена на высоте 40 см). Виртуальная сцена воспроизводит типичное офисное пространство. Сцена включает центральный коридор длиной около 72 м и шириной 4 м, который использовался в качестве основной экспериментальной зоны. Коридор соединяется с помещением, где размещены офисные объекты (рис. 2).

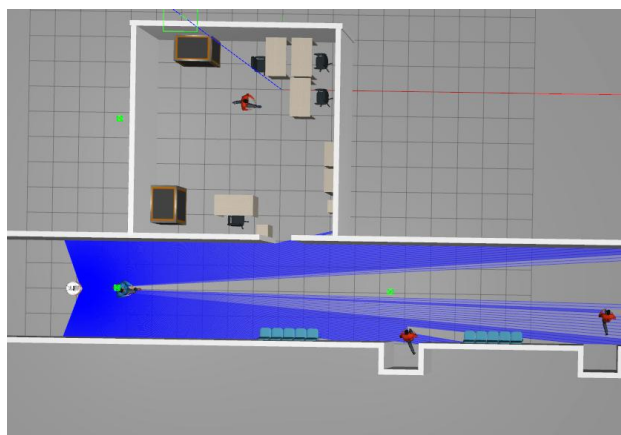


Рис. 2. Структура виртуальной офисной среды: центральный коридор и помещение с офисной мебелью

Целевая модель человека была одета в зеленую одежду и двигалась по заранее заданным траекториям. Средняя скорость движения составляла 0,6 м/с, что соответствует типичному шагу человека. Для проверки алгоритма использовались три варианта одежды: брюки, юбка и длинное платье (см. рис. 1). Помимо целевой модели в сцену добавлялись и другие модели, выполнявшие роль динамических препятствий [43] (см. рис. 2).

Процедура эксперимента состояла из четырех этапов:

1. Инициализация. Робот и целевая модель размещались в заданных начальных позициях виртуальной сцены. Модель устанавливали перед роботом, чтобы датчики могли корректно зафиксировать её положение.
2. Исполнение. После запуска сценария модель начинала движение по заранее определенной траектории, а робот выполнял задачу отслеживания и следования.
3. Сбор данных. В ходе каждой попытки фиксировались координаты робота, траектория движения цели и случаи ложных срабатываний алгоритма.
4. Завершение. Эксперимент завершался, когда модель достигала конечной точки маршрута либо при ошибке отслеживания (ложном срабатывании).

Вычислительные эксперименты были направлены на проверку алгоритма отслеживания и следования за человеком в различных условиях: тип одежды целевой модели (брюки, юбка и длинное платье), тип траектории движения (прямолинейный, зигзагообразный и повороты на 90°). Для каждой комбинации факторов выполнялось 50 повторов. В сумме было проведено 450 испытаний.

Для оценки работы алгоритма были выбраны два показателя: средняя пройденная дистанция и ложные срабатывания. Средняя пройденная дистанция определяется как среднее расстояние, которое робот проходит за целевой моделью до завершения эксперимента. Показатель отражает способность алгоритма устойчиво удерживать цель на протяжении всего маршрута. Ложные срабатывания подсчитываются как количество случаев, когда робот ошибочно принимает другой объект за целевую модель. При возникновении такой ошибки эксперимент прерывался, а результатом считалась дистанция, пройденная до момента сбоя.

Для интегральной оценки использовался показатель точности следования, определяемый по формуле:

$$M = \frac{1}{2} (1 + (\sum_i d_i / \sum_i D_i) - K_f / K), \quad (1)$$

где K – общее количество экспериментов, d_i – дистанция, пройденная роботом в i -ой попытке, D_i – длина заданного маршрута, K_f – число потерь цели (ложных срабатываний) за все попытки. При идеальном следовании без ошибок $M = 1$, а при систематических сбоях и отсутствии продвижения $M = 0$. Таким образом, чем ближе значение показателя к единице, тем устойчивее и надежнее работа алгоритма.

Результаты экспериментов. Результаты работы алгоритма отслеживания и следования за пользователем представлены в табл. 1. В ней приведены данные для различных комбинаций траекторий движения и вариантов одежды целевой модели:

1. Тип одежды: брюки (Б), юбка (Ю), длинное платье (ДП).
2. Тип траектории: прямолинейная (ПЛ), зигзагообразная (ЗЗ), поворот на 90° (П90).
3. Средняя пройденная дистанция (в метрах), отражающая, какое расстояние в среднем преодолевал робот, двигаясь за целью.
4. Количество сбоев отслеживания (из 50 попыток), где значение 50 означает полную неспособность алгоритма пройти данный сценарий.
5. Показатель точности M – показатель качества работы алгоритма, учитывающий долю пройденного пути и число сбоев.

Прямолинейные траектории. Во всех сценариях робот проходил в среднем от 58% до 62% маршрута (29-31 м из 50 м). Это демонстрирует, что алгоритм сохраняет относительную устойчивость при простых сценариях движения без резких изменений направления движения пользователя.

Зигзагообразные траектории. Этот сценарий оказался наиболее сложным для алгоритма: во всех экспериментах фиксировались сбои отслеживания, а средняя дистанция составляла лишь 22-35% маршрута (11-17 м из 50 м). Такие результаты подчеркивают ограниченную способность алгоритма адаптироваться к нелинейным траекториям движения.

Повороты на 90°. Результаты носили смешанный характер. В отдельных случаях (например, при использовании ЦМЧ в брюках) робот успешно преодолевал до 84% маршрута (12,5 м из 15 м), что говорит об относительной устойчивости алгоритма. Однако в других условиях (например, для пользователя в длинном платье) частота сбоев оставалась высокой, а эффективность снижалась. Это свидетельствует о том, что алгоритм способен справляться с резкими изменениями направления, но его надёжность в значительной степени зависит от внешнего вида целевой ЦМЧ.

Брюки. На прямолинейной траектории робот проходил в среднем 29,51 м при 42 сбоях, что соответствовало точности 0,375. При зигзагообразном движении средняя дистанция снизилась до 11,83 м, и все 50 попыток завершились сбоем (точность 0,118). На повороте на 90° робот преодолевал в среднем 12,55 м при 38 сбоях, а метрика точности достигала 0,538 благодаря тому, что значительная часть короткого маршрута проходила успешно.

Таблица 1

**Результаты работы алгоритма отслеживания и следования за пользователем
для различных комбинаций траекторий и вариантов одежды**

Тип одежды	Тип траектории	Средняя пройденная дистанция (м)	Количество сбоев (из 50)	Метрика точности
Брюки	ПЛ	29,51	42	0,375
	ЗЗ	11,83	50	0,118
	П90	12,55	38	0,538
Юбка	ПЛ	30,8	38	0,428
	ЗЗ	17,29	50	0,173
	П90	6,22	40	0,307
Длинное платье	ПЛ	29,09	45	0,341
	ЗЗ	11,12	50	0,111
	П90	11,54	50	0,385

Юбка. На прямолинейной траектории алгоритм обрабатывал более успешно, достигнув средней длины пути 30,8 м при 38 сбоях (точность 0,428). При зигзагообразном движении средняя дистанция выросла по сравнению со сценарием с брюками (17,29 м), но все попытки завершились сбоем (точность 0,173). На повороте на 90° средняя дистанция составила 6,22 м при 40 сбоях, а метрика точности равнялась 0,307.

Длинное платье. Прямолинейная траектория также показала наилучший результат: средняя длина пути 29,09 м при 45 сбоях (точность 0,341). При зигзагообразном движении средняя дистанция составила 11,12 м, и все 50 попыток закончились сбоем (точность 0,111). На повороте на 90° робот проходил в среднем 11,54 м, также со 100% сбоев, но метрика достигала 0,385 за счет небольшой длины маршрута.

В целом алгоритм демонстрирует умеренную устойчивость при простых условиях, но его эффективность существенно снижается в ситуациях с нелинейным движением и вариативностью внешнего вида пользователя. Наиболее стабильные результаты достигнуты при использовании модели в брюках. В сценариях с юбкой и длинным платьем алгоритм показывал меньшее качество отслеживания, особенно на сложных траекториях. Это объясняется тем, что свободная одежда изменяет силуэт человека и может снижать надёжность детекции. Эти результаты подчеркивают необходимость дальнейшего совершенствования алгоритмов, особенно в направлении повышения их надёжности к изменениям траектории и визуального представления цели.

Несмотря на полученные положительные результаты, использование разработанных функциональных ЦМЧ имеет ряд ограничений, которые следует учитывать. Во-первых, уровень графической реалистичности цифровых моделей в симуляторе Gazebo остается ограниченным по сравнению с натурными объектами [44]. Система визуализации не полностью воспроизводит реальные условия освещения, что может снижать точность алгоритмов компьютерного зрения, особенно тех, которые чувствительны к изменению яркости, динамике теней и цветовых характеристик.

Во-вторых, используемое окружение представляет собой упрощенную модель офисного пространства. Хотя в нем присутствуют препятствия и динамические объекты, уровень сложности ниже, чем в реальных условиях, где возможны непредсказуемые сценарии и широкий спектр визуальных помех. В то же время стоит отметить, что новые версии симулятора Gazebo интегрируют современные движки визуализации (например, OGRE следующего поколения) и расширенные плагины освещения [45], что открывает возможности для повышения реалистичности последующих экспериментов с разработанными ЦМЧ.

Заключение. В работе проведена оценка алгоритма отслеживания и следования за человеком с использованием ЦМЧ в симуляторе Gazebo. В исследовании сделан акцент на практическом применении новых детализированных функциональных настраиваемых ЦМЧ, которые заменили использованные ранее стандартные модели человека в симуляторе Gazebo. Вычислительные эксперименты охватывали разные варианты одежды пользователя и траектории его/её движения, что позволило выявить влияние внешнего вида и характера движения на устойчивость работы алгоритма. Результаты показали, что наилучшие показатели достигаются при прямолинейных траекториях, тогда как зигзагообразные маршруты приводят к частым сбоям. В сценариях с поворотами на 90° эффективность алгоритма зависит как от траектории, так и от одежды ЦМЧ. Наиболее устойчивое отслеживание обеспечивали модели в брюках, в то время как юбка и длинное платье снижали надёжность алгоритма. Использование новых ЦМЧ подтвердило свою ценность как инструмента для воспроизводимой оценки алгоритмов компьютерного зрения. Разработанный подход может служить основой для дальнейших исследований, направленных на повышение надежности и безопасности робототехнических систем, работающих в человеко-ориентированных средах.

Исследование выполнено за счет гранта Российского научного Фонда и Академии наук Республики Татарстан по проекту № 26-11-20030.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Bonci A., Cen Cheng P.D., Indri M., Nabissi G., Sibona F. Human-Robot Perception in Industrial Environments: A Survey // *Sensors*. – 2021. – Vol. 21, No. 5. – P. 1571.
2. Chen K., Zhang D., Yao L., Guo B., Yu Z., Liu Y. Deep learning for sensor-based human activity recognition: Overview, challenges, and opportunities // *ACM Computing Surveys*. – 2021. – Vol. 54, No. 4. – P. 1-40.
3. Chen C., Liu Y., Kreiss S., Alahi A. Crowd-Robot Interaction: Crowd-Aware Robot Navigation with Attention-Based Deep Reinforcement Learning // *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA 2019)*. – 2019. – P. 6015-6022.
4. Liu L., Dugas D., Cesari G., Siegwart R., Dub R. Robot navigation in crowded environments using deep reinforcement learning // *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. – 2020. – P. 5671-5677.
5. Martinetti A., Chemweno P.K., Nizamis K., Fosch-Villaronga E. Redefining safety in light of human-robot interaction: A critical review of current standards and regulations // *Frontiers in Chemical Engineering*. – 2021. – Vol. 3. – P. 666237.
6. Sanchez J.A.S., Lopez-Gonzalez E., Magid E., Martinez-Garcia E.A. Double spiraliform path planning and tracking for agricultural mobile robotics: A modeling and simulation study // *Computers and Electronics in Agriculture*. – 2025. – Vol. 237. – P. 110715.
7. Bavle H., Sanchez-Lopez J.L., Cimarrelli C., Tourani A., Voos H. From SLAM to situational awareness: Challenges and survey // *Sensors*. – 2023. – Vol. 23, No. 10. – P. 4849.
8. Campos C., Elvira R., Gomez-Rodriguez J.J., Montiel J.M.M., Tardos J.D. ORB-SLAM3: An Accurate Open-Source Library for Visual, Visual-Inertial and Multi-Map SLAM // *IEEE Transactions on Robotics*. – 2021. – Vol. 37, No. 6. – P. 1874-1890.
9. Bescos B., Facil J.M., Civera J., Neira J. DynaSLAM: Tracking, Mapping and Inpainting in Dynamic Scenes // *IEEE Robotics and Automation Letters*. – 2018. – Vol. 3, No. 4. – P. 4076-4083.
10. Mingachev E., Lavrenov R., Tsoy T., Matsuno F., Svinin M., Suthacorn J., Magid E. Comparison of ROS-based monocular visual SLAM methods: DSO, LDSO, ORB-SLAM2 & DynaSLAM // *Lecture Notes in Computer Science*. – 2020. – Vol. 12336. – P. 222-233.
11. Zachariae A., Plahl F., Tang Y., Mamaev I., Hein B., Wurll C. Human-robot interactions in autonomous hospital transports // *Robotics and Autonomous Systems*. – 2024. – Vol. 179. – P. 104755.

12. Yoshimitsu K., Shiigi Y., Kobayashi E., Kuwana K., Kusuda K., Hoshide K., Masamune K. Exploring the potential of service robots in guiding patients and assisting mobility in medical facilities // *Advanced Robotics*. – 2024. – Vol. 38, No. 24. – P. 1758-1769.
13. Valner R., Masnavi H., Rybalskii I., Poluaar R., Koiv E., Aabloo A., Kruusamae K., Singh A.K. Scalable and heterogenous mobile robot fleet-based task automation in crowded hospital environments – a field test // *Frontiers in Robotics and AI*. – 2022. – Vol. 9. – P. 922835.
14. Yoo H.J., Kim E.H., Lee H. Mobile robots for isolation-room hospital settings: A scenario-based preliminary study // *Computational and Structural Biotechnology Journal*. – 2024. – Vol. 24. – P. 237-246.
15. Rondoni C., Scotto Di Luizo F., Tamantini C., Tagliamonte N.L., Chiurazzi M., Ciuti G., Zollo L. Navigation benchmarking for autonomous mobile robots in hospital environment // *Scientific Reports*. – 2024. – Vol. 14. – P. 18334.
16. Chebotareva E., Mustafin M., Safin R., Tsoy T., Martinez-Garcia E.A., Magid E. Camera-based safety system for collaborative assembly // *Journal of Intelligent Manufacturing*. – 2024.
17. Custodio L., Machado R. Flexible automated warehouse: A literature review and an innovative framework // *International Journal of Advanced Manufacturing Technology*. – 2020. – Vol. 106. – P. 533-558.
18. Lackner T., Hermann J., Kuhn C., Palm D. Review of autonomous mobile robots in intralogistics: state-of-the-art, limitations and research gaps // *Procedia CIRP*. – 2024. – Vol. 130. – P. 930-935.
19. Pasparakis A., De Vries J., De Koster R. Assessing the Impact of Human–Robot Collaborative Order Picking Systems on Warehouse Workers // *International Journal of Production Research*. – 2023. – Vol. 61, No. 22. – P. 7776-7790.
20. Gao Y., Huang C.M. Evaluation of Socially-Aware Robot Navigation // *Frontiers in Robotics and AI*. – 2022. – Vol. 8. – P. 721317.
21. Mavrogiannis C., Baldini F., Wang A., Zhao D., Trautman P., Steinfeld A., Oh J. Core Challenges of Social Robot Navigation: A Survey // *ACM Transactions on Human-Robot Interaction*. – 2023. – Vol. 12, No. 3. – P. 1-39.
22. Chen C., Liu Y., Kreiss S., Alahi A. Crowd-Robot Interaction: Crowd-Aware Robot Navigation with Attention-Based Deep Reinforcement Learning // *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA 2019)*. – 2019. – P. 6015-6022.
23. Singamaneni P.T., Bachiller-Burgos P., Manso L.J., Garrell A., Sanfeliu A., Spalanzani A., Alami R. A Survey on Socially Aware Robot Navigation: Taxonomy and Future Challenges // *The International Journal of Robotics Research*. – 2024. – Vol. 43, No. 10. – P. 1523-1572.
24. Guillen-RUIZ S., Bandera J.P., Hidalgo-Paniagua A., Bandera A. Evolution of Socially-Aware Robot Navigation // *Electronics*. – 2023. – Vol. 12, No. 7. – P. 1570.
25. Астапова М.А. Интеллектуальное управление робототехническими средствами в задаче сегментации нор грызунов с использованием глубоких сверточных архитектур // *Известия ЮФУ. Технические науки*. – 2025. – № 2. – С. 19-30.
26. Филиппов Р.А. Робототехнические средства для сбора ягод земляники // *Мехатроника, автоматика и робототехника*. – 2017. – № 1. – С. 26-27.
27. Paullada A., Raji I.D., Bender E.M., Denton E., Hanna A. Data and its (dis)contents: A survey of dataset development and use in machine learning research // *Patterns*. – 2021. – Vol. 2, No. 11. – P. 100336.
28. Martinetti A., Chemweno P.K., Nizamis K., Fosch-Villaronga E. Redefining safety in light of human-robot interaction: A critical review of current standards and regulations // *Frontiers in Chemical Engineering*. – 2021. – Vol. 3. – P. 666237.
29. Shi X., Li D., Zhao P., Tian Q., Tian Y., Long Q., Zhu C., Song J., Qiao F., Song L., Guo Y., Wang Z., Zhang Y., Qin B., Yang W., Wang F., Chan R.H.M., She Q. Are We Ready for Service Robots? The OpenLORIS-Scene Datasets for Lifelong SLAM // *2020 IEEE International Conference on Robotics and Automation (ICRA 2020)*. – 2020. – P. 3139-3145.
30. Линн Х.Х., Лулин С.А., Линн Ч.Н.3. и др. Натурное и имитационное моделирование централизованной системы управления транспортными роботами // *International Journal of Open Information Technologies*. – 2020. – Т. 8, № 4. – С. 17-24.
31. Demirel H.O., Ahmed S., Duffy V.G. Digital human modeling: A review and reappraisal of origins, present, and expected future methods for representing humans computationally // *International Journal of Human-Computer Interaction*. – 2021. – Vol. 38, No. 10. – P. 897-937.
32. Yin M.Y., Li J.G. A systematic review on digital human models in assembly process planning // *International Journal of Advanced Manufacturing Technology*. – 2023. – Vol. 125. – P. 1037-1059.
33. Abbyasov B., Gamberov T., Zhukova V., Tsoy T., Martinez-Garcia E.A., Magid E. A tutorial on modeling a real office environment in Gazebo simulator // *Journal of Robotics, Networking and Artificial Life*. – 2023. – Vol. 10, No. 2. – P. 166-169.
34. Tukhtamanov N., Lavrenov R., Chebotareva E., Svinin., Tsoy T., Magid E. Open-source library of human models for Gazebo simulator // *2022 International Siberian Conference on Control and Communications (SIBCON)*. – 2022. – P. 1-5.

35. Fortini L., Leonori M., Gandarias J.M., De Momi E., Ajoudani A. Open-VICO: An open-source Gazebo toolkit for vision-based skeleton tracking in human-robot collaboration // *Proceedings of the 31st IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*. – 2022. – P. 511-517.
36. Wang J., Li X., Gao L., Liu J., Xu C., Wang Y. Deep 3D human pose estimation: A review // *Computer Vision and Image Understanding*. – 2021. – Vol. 210. – P. 103225.
37. Diamantis D.E., Kalozoumis P.G., Iakovidis D.K. Digital twin for simulation and evaluation of assistive navigation systems // *Digital Twins for Digital Transformation: Innovation in Industry*. – Cham: Springer, 2022. – Vol. 423. – P. 147-170.
38. Eirale A., Martini M., Chiaberge M. Human-Centered Navigation and Person-Following with Omnidirectional Robot for Indoor Assistance and Monitoring // *Robotics*. – 2022. – Vol. 11, No. 5. – P. 108.
39. Chebotareva E., Hsia K.H., Yakovlev K., Magid E. Laser rangefinder and monocular camera data fusion for human-following algorithm by PMB-2 mobile robot in simulated Gazebo environment // *Smart Innovation, Systems and Technologies*. – Singapore: Springer, 2021. – 2021. – Vol. 187. – P. 357-369.
40. Paul G., Scataglini S. Open-source software to create a kinematic model in digital human modeling // *DHM and Posturography*. – Cambridge, MA: Academic Press, 2019. – 2019. – P. 201-213.
41. Briceno L., Paul G. MakeHuman: A review of the modelling framework // *Proceedings of the 20th Congress of the International Ergonomics Association (IEA 2018)*. – Cham: Springer, 2019. – 2019. – Vol. 822. – P. 224-232.
42. Algabri R., Choi M.-T. Deep-Learning-Based Indoor Human Following of Mobile Robot Using Color Feature // *Sensors*. – 2020. – Vol. 20, No. 9. – P. 2699.
43. Xie Z., Xin P., Dames P. Towards safe navigation through crowded dynamic environments // *2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. – 2021. – P. 4934-4940.
44. Nguyen D.M., Nazeri M., Payandeh A., Datar A., Xiao X. Toward Human-Like Social Robot Navigation: A Large-Scale, Multi-Modal, Social Human Navigation Dataset // *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2023)*. – 2023. – P. 7442-7447.
45. Sultanov R., Safin R., Martinez-Garcia E.A., Magid E. Modeling sunlight in Gazebo for vision-based applications under varying light conditions // *Proceedings of the 21st International Conference on Informatics in Control, Automation and Robotics (ICINCO)*. – 2024. – Vol. 1. – P. 519-526.

REFERENCES

1. Bonci A., Cen Cheng P.D., Indri M., Nabissi G., Sibona F. Human-Robot Perception in Industrial Environments: A Survey, *Sensors*, 2021, Vol. 21, No. 5, pp. 1571.
2. Chen K., Zhang D., Yao L., Guo B., Yu Z., Liu Y. Deep learning for sensor-based human activity recognition: Overview, challenges, and opportunities, *ACM Computing Surveys*, 2021, Vol. 54, No. 4, pp. 1-40.
3. Chen C., Liu Y., Kreiss S., Alahi A. Crowd-Robot Interaction: Crowd-Aware Robot Navigation with Attention-Based Deep Reinforcement Learning, *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA 2019)*, 2019, pp. 6015-6022.
4. Liu L., Dugas D., Cesari G., Siegwart R., Dub R. Robot navigation in crowded environments using deep reinforcement learning, *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2020, pp. 5671-5677.
5. Martinetti A., Chemweno P.K., Nizamis K., Fosch-Villaronga E. Redefining safety in light of human-robot interaction: A critical review of current standards and regulations, *Frontiers in Chemical Engineering*, 2021, Vol. 3, pp. 666237.
6. Sanchez J.A.S., Lopez-Gonzalez E., Magid E., Martinez-Garcia E.A. Double spiraliform path planning and tracking for agricultural mobile robotics: A modeling and simulation study, *Computers and Electronics in Agriculture*, 2025, Vol. 237, pp. 110715.
7. Bavle H., Sanchez-Lopez J.L., Cimarelli C., Tourani A., Voos H. From SLAM to situational awareness: Challenges and survey, *Sensors*, 2023, Vol. 23, No. 10, pp. 4849.
8. Campos C., Elvira R., Gomez-Rodriguez J.J., Montiel J.M.M., Tardos J.D. ORB-SLAM3: An Accurate Open-Source Library for Visual, Visual-Inertial and Multi-Map SLAM, *IEEE Transactions on Robotics*, 2021, Vol. 37, No. 6, pp. 1874-1890.
9. Bescos B., Facil J.M., Civera J., Neira J. DynaSLAM: Tracking, Mapping and Inpainting in Dynamic Scenes, *IEEE Robotics and Automation Letters*, 2018, Vol. 3, No. 4, pp. 4076-4083.
10. Mingachev E., Lavrenov R., Tsoy T., Matsuno F., Svinin M., Suthacorn J., Magid E. Comparison of ROS-based monocular visual SLAM methods: DSO, LDSO, ORB-SLAM2 & DynaSLAM, *Lecture Notes in Computer Science*, 2020, Vol. 12336, pp. 222-233.
11. Zachariae A., Plahl F., Tang Y., Mamaev I., Hein B., Wurll C. Human-robot interactions in autonomous hospital transports, *Robotics and Autonomous Systems*, 2024, Vol. 179, pp. 104755.

12. Yoshimitsu K., Shiigi Y., Kobayashi E., Kuwana K., Kusuda K., Hoshide K., Masamune K. Exploring the potential of service robots in guiding patients and assisting mobility in medical facilities, *Advanced Robotics*, 2024, Vol. 38, No. 24, pp. 1758-1769.
13. Valner R., Masnavi H., Rybalskii I., Poluaar R., Koiv E., Aabloo A., Kruusamae K., Singh A.K. Scalable and heterogenous mobile robot fleet-based task automation in crowded hospital environments – a field test, *Frontiers in Robotics and AI*, 2022, Vol. 9, pp. 922835.
14. Yoo H.J., Kim E.H., Lee H. Mobile robots for isolation-room hospital settings: A scenario-based preliminary study, *Computational and Structural Biotechnology Journal*, 2024, Vol. 24, pp. 237-246.
15. Rondoni C., Scotto Di Luizo F., Tamantini C., Tagliamonte N.L., Chiurazzi M., Ciuti G., Zollo L. Navigation benchmarking for autonomous mobile robots in hospital environment, *Scientific Reports*, 2024, Vol. 14, pp. 18334.
16. Chebotareva E., Mustafin M., Safin R., Tsoy T., Martinez-Garcia E.A., Magid E. Camera-based safety system for collaborative assembly, *Journal of Intelligent Manufacturing*, 2024.
17. Custodio L., Machado R. Flexible automated warehouse: A literature review and an innovative framework, *International Journal of Advanced Manufacturing Technology*, 2020, Vol. 106, pp. 533-558.
18. Lackner T., Hermann J., Kuhn C., Palm D. Review of autonomous mobile robots in intralogistics: state-of-the-art, limitations and research gaps, *Procedia CIRP*, 2024, Vol. 130, pp. 930-935.
19. Pasparakis A., De Vries J., De Koster R. Assessing the Impact of Human–Robot Collaborative Order Picking Systems on Warehouse Workers, *International Journal of Production Research*, 2023, Vol. 61, No. 22, pp. 7776-7790.
20. Gao Y., Huang C.M. Evaluation of Socially-Aware Robot Navigation, *Frontiers in Robotics and AI*, 2022, Vol. 8, pp. 721317.
21. Mavrogiannis C., Baldini F., Wang A., Zhao D., Trautman P., Steinfeld A., Oh J. Core Challenges of Social Robot Navigation: A Survey, *ACM Transactions on Human-Robot Interaction*, 2023, Vol. 12, No. 3, pp. 1-39.
22. Chen C., Liu Y., Kreiss S., Alahi A. Crowd-Robot Interaction: Crowd-Aware Robot Navigation with Attention-Based Deep Reinforcement Learning, *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA 2019)*, 2019, pp. 6015-6022.
23. Singamaneni P.T., Bachiller-Burgos P., Manso L.J., Garrell A., Sanfeliu A., Spalanzani A., Alami R. A Survey on Socially Aware Robot Navigation: Taxonomy and Future Challenges, *The International Journal of Robotics Research*, 2024, Vol. 43, No. 10, pp. 1523-1572.
24. Guillen-RUIZ S., Bandera J.P., Hidalgo-Paniagua A., Bandera A. Evolution of Socially-Aware Robot Navigation, *Electronics*, 2023, Vol. 12, No. 7, pp. 1570.
25. Astapova M.A. Intel'ektual'noe upravlenie robototekhnicheskimi sredstvami v zadache segmentatsii nor gryzunov s ispol'zovaniem glubokikh svertochnykh arkhitektur [Intelligent control of robotic means in the problem of segmentation of rodent burrows using deep convolutional architectures], *Izvestiya YuFU. Tekhnicheskie nauki* [Izvestiya SFedU. Engineering Sciences], 2025, No. 2, pp. 19-30.
26. Filippov R.A. Robototekhnicheskie sredstva dlya sbora yagod zemlyaniki [Robotic means for picking strawberries], *Mekhatronika, avtomatika i robototekhnika* [Mechatronics, Automation and Robotics], 2017, No. 1, pp. 26-27.
27. Paullada A., Raji I.D., Bender E.M., Denton E., Hanna A. Data and its (dis)contents: A survey of dataset development and use in machine learning research, *Patterns*, 2021, Vol. 2, No. 11, pp. 100336.
28. Martinetti A., Chemweno P.K., Nizamis K., Fosch-Villaronga E. Redefining safety in light of human-robot interaction: A critical review of current standards and regulations, *Frontiers in Chemical Engineering*, 2021, Vol. 3, pp. 666237.
29. Shi X., Li D., Zhao P., Tian Q., Tian Y., Long Q., Zhu C., Song J., Qiao F., Song L., Guo Y., Wang Z., Zhang Y., Qin B., Yang W., Wang F., Chan R.H.M., She Q. Are We Ready for Service Robots? The OpenLORIS-Scene Datasets for Lifelong SLAM, *2020 IEEE International Conference on Robotics and Automation (ICRA 2020)*, 2020, pp. 3139-3145.
30. Linn Kh.Kh., Lupin S.A., Linn Ch.N.Z. i dr. Natural'noe i imitatsionnoe modelirovanie tsentralizovannoy sistemy upravleniya transportnymi robotami [Natural and simulation modeling of a centralized control system for transport robots], *International Journal of Open Information Technologies*, 2020, Vol. 8, No. 4, pp. 17-24.
31. Demirel H.O., Ahmed S., Duffy V.G. Digital human modeling: A review and reappraisal of origins, present, and expected future methods for representing humans computationally, *International Journal of Human-Computer Interaction*, 2021, Vol. 38, No. 10, pp. 897-937.
32. Yin M.Y., Li J.G. A systematic review on digital human models in assembly process planning, *International Journal of Advanced Manufacturing Technology*, 2023, Vol. 125, pp. 1037-1059.
33. Abbyasov B., Gamberov T., Zhukova V., Tsoy T., Martinez-Garcia E.A., Magid E. A tutorial on modeling a real office environment in Gazebo simulator, *Journal of Robotics, Networking and Artificial Life*, 2023, Vol. 10, No. 2, pp. 166-169.

34. Tukhtamanov N., Lavrenov R., Chebotareva E., Svinin., Tsoy T., Magid E. Open-source library of human models for Gazebo simulator, *2022 International Siberian Conference on Control and Communications (SIBCON)*, 2022, pp. 1-5.
35. Fortini L., Leonori M., Gandarias J.M., De Momi E., Ajoudani A. Open-VICO: An open-source Gazebo toolkit for vision-based skeleton tracking in human-robot collaboration, *Proceedings of the 31st IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*, 2022, pp. 511-517.
36. Wang J., Li X., Gao L., Liu J., Xu C., Wang Y. Deep 3D human pose estimation: A review, *Computer Vision and Image Understanding*, 2021, Vol. 210, pp. 103225.
37. Diamantis D.E., Kalozoumis P.G., Iakovidis D.K. Digital twin for simulation and evaluation of assistive navigation systems, *Digital Twins for Digital Transformation: Innovation in Industry*. Cham: Springer, 2022, Vol. 423, pp. 147-170.
38. Eirale A., Martini M., Chiaberge M. Human-Centered Navigation and Person-Following with Omnidirectional Robot for Indoor Assistance and Monitoring, *Robotics*, 2022, Vol. 11, No. 5, pp. 108.
39. Chebotareva E., Hsia K.H., Yakovlev K., Magid E. Laser rangefinder and monocular camera data fusion for human-following algorithm by PMB-2 mobile robot in simulated Gazebo environment, *Smart Innovation, Systems and Technologies*. Singapore: Springer, 2021, Vol. 187, pp. 357-369.
40. Paul G., Scataglini S. Open-source software to create a kinematic model in digital human modeling, *DHM and Posturography*. Cambridge, MA: Academic Press, 2019, pp. 201-213.
41. Briceno L., Paul G. MakeHuman: A review of the modelling framework, *Proceedings of the 20th Congress of the International Ergonomics Association (IEA 2018)*. Cham: Springer, 2019, Vol. 822, pp. 224-232.
42. Algabri R., Choi M.-T. Deep-Learning-Based Indoor Human Following of Mobile Robot Using Color Feature, *Sensors*, 2020, Vol. 20, No. 9, pp. 2699.
43. Xie Z., Xin P., Dames P. Towards safe navigation through crowded dynamic environments, *2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2021, pp. 4934-4940.
44. Nguyen D.M., Nazeri M., Payandeh A., Datar A., Xiao X. Toward Human-Like Social Robot Navigation: A Large-Scale, Multi-Modal, Social Human Navigation Dataset, *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2023)*, 2023, pp. 7442-7447.
45. Sultanov R., Safin R., Martinez-Garcia E.A., Magid E. Modeling sunlight in Gazebo for vision-based applications under varying light conditions, *Proceedings of the 21st International Conference on Informatics in Control, Automation and Robotics (ICINCO)*, 2024, Vol. 1, pp. 519-526.

Гамберов Тимур Рустамович – Казанский (Приволжский) федеральный университет; e-mail: timycho2@it.kfu.ru; г. Казань, Россия; кафедра интеллектуальной робототехники; лаборант.

Сафин Рамиль Наibuлович – Казанский (Приволжский) федеральный университет; e-mail: safin.ramil@it.kfu.ru; г. Казань, Россия; кафедра интеллектуальной робототехники; старший преподаватель.

Чеботарева Эльвира Валерьевна – Казанский (Приволжский) федеральный университет; e-mail: elvira.chebotareva@it.kfu.ru; г. Казань, Россия; кафедра интеллектуальной робототехники; к.ф.-м.н.; доцент.

Цой Татьяна Григорьевна – Казанский (Приволжский) федеральный университет; e-mail: tt@it.kfu.ru; г. Казань, Россия; кафедра интеллектуальной робототехники; к.т.н.; доцент.

Магид Евгений Аркадьевич – Казанский (Приволжский) федеральный университет; e-mail: magid@it.kfu.ru; г. Казань, Россия; кафедра интеллектуальной робототехники; PhD; зав. кафедрой.

Gamberov Timur Rustamovich – Kazan Federal University; e-mail: timycho2@it.kfu.ru; Kazan, Russia; the Department of Intelligent Robotics; research assistant.

Safin Ramil Naibulovich – Kazan Federal University; e-mail: safin.ramil@it.kfu.ru; Kazan, Russia; the Department of Intelligent Robotics; research associate.

Chebotareva Elvira Valer'evna – Kazan Federal University, e-mail: elvira.chebotareva@it.kfu.ru; Kazan, Russia; the Department of Intelligent Robotics; cand. of phys. and math. sc.; associate professor.

Tsoy Tatyana Grigor'evna – Kazan Federal University; e-mail: tt@it.kfu.ru; Kazan, Russia; the Department of Intelligent Robotics; cand. of eng. sc.; associate professor.

Magid Evgeni Arkad'evich – Kazan Federal University; e-mail: magid@it.kfu.ru; Kazan, Russia; the Department of Intelligent Robotics; PhD; head of department.

А.Д. Скорик

АЛГОРИТМ ВЫБОРА ОПТИМАЛЬНОЙ РАБОЧЕЙ ЧАСТОТЫ И ПАРАМЕТРОВ ПЕРЕДАЧИ В ДКМ-КАНАЛЕ НА ОСНОВЕ ПРОГНОЗИРОВАНИЯ ПО ДАННЫМ ГНСС-МОНИТОРИНГА

Актуальность. Современные требования к автоматизированным системам управления (АСУ) специального назначения и достижения в области техники связи возрождают интерес к декаметровому (ДКМ) диапазону как к перспективному средству управления робототехническими комплексами (РТК) различного базирования, прежде всего морскими и воздушными. Существующие алгоритмы адаптации ДКМ-систем, регламентированные национальными и международными стандартами являются реактивными и не учитывают частотную зависимость параметров ионосферы, что ограничивает их эффективность для решения задач надежного управления РТК. **Цель работы** – разработка алгоритма выбора оптимальной рабочей частоты и параметров передачи (полоса, скорость, мощность, глубина перемежения) на основе прогнозирования частотно-зависимых характеристик диффузного ДКМ-канала по данным пассивного ГНСС-мониторинга. **Методы.** Используются структурно-многолучевая модель для интервалов корреляции, методики ITU-R для среднего отношения сигнал/помеха, аналитические выражения для параметра Накагами и допустимого отношения сигнал/помеха с учётом разнесённого приёма. **Результаты.** Разработан алгоритм, включающий оценку параметров ионосферы по результатам ГНСС-мониторинга, расчёт частотно-зависимых характеристик канала и выбор оптимального режима работы, максимизирующего энергетический запас и надёжность связи. Численное моделирование для типовой трассы подтвердило его работоспособность. **Выводы.** Предложенный алгоритм обеспечивает упреждающий выбор параметров связи с учётом реального состояния ионосферы, повышая надёжность и скорость передачи, что критически важно для задач управления РТК. Он является шагом к созданию когнитивных КВ-систем.

Декаметровая связь; робототехнические комплексы; оптимальная рабочая частота; прогнозирование; ионосфера; ГНСС; параметр Накагами; когнитивное радио.

A.D. Skorik

ALGORITHM FOR SELECTING THE OPTIMAL OPERATING FREQUENCY AND TRANSMISSION PARAMETERS IN THE HF CHANNEL BASED ON PREDICTION USING GNSS MONITORING DATA

Relevance. Modern requirements for specialized automated control systems and advances in communication technology are reviving interest in the HF band as a promising means of controlling ground, sea, and air-based robotic complexes (RCs). Existing adaptation algorithms for HF communication systems, by national and international standards are reactive and do not take into account the frequency dependence of ionospheric parameters, which limits their effectiveness for reliable RC control. **Objective.** To develop an algorithm for selecting the optimal operating frequency and transmission parameters (bandwidth, data rate, power, interleaving depth) based on prediction of frequency-dependent characteristics of a diffuse HF channel using passive GNSS monitoring data. **Methods.** A structural multipath model for correlation intervals, the ITU-R method for the mean signal-to-interference ratio, and analytical expressions for the Nakagami parameter and the permissible signal-to-interference ratio taking into account diversity reception are used. **Results.** An algorithm has been developed that includes estimating ionospheric parameters from GNSS data, calculating the frequency-dependent characteristics of the channel, and selecting the optimal operating mode that maximizes the energy margin and communication reliability. Numerical simulation for a typical path confirmed its efficiency. **Conclusions.** The proposed algorithm provides proactive selection of communication parameters taking into account the real state of the ionosphere, increasing reliability and data rate, which is critically important for RC control tasks. It is a step towards the creation of cognitive HF systems.

HF communication; robotic complexes; optimal operating frequency; prediction; ionosphere; GNSS; Nakagami parameter; cognitive radio.

Введение. В последние годы наблюдается устойчивый рост требований к автоматизированным системам управления (АСУ) специального назначения, связанный с необходимостью обеспечения надежного и устойчивого канала связи с подвижными объектами [1, 2], действующими в условиях отсутствия или подавления спутниковых систем связи. Особую актуальность эта проблема приобретает для управления робототехническими комплексами (РТК) различного базирования – наземными, морскими и воздушными. Для морских и воздушных РТК, действующих на значительном удалении от пунктов управления, традиционные средства связи УКВ-диапазона оказываются неэффективными из-за ограниченной дальности прямой видимости, а спутниковые каналы – недоступными в силу различных факторов (высокая эксплуатационная стоимость и стоимость терминала, ограниченность применения, например, из-за низких углов закрытия и пр.)

В этом контексте декаметровый (ДКМ) диапазон (3–30 МГц), обладающий уникальным свойством сверхдальнего распространения радиоволн за счет отражения от ионосферы, вновь привлекает внимание разработчиков. Современные достижения в области цифровой обработки сигналов, помехоустойчивого кодирования и адаптивных алгоритмов позволяют преодолеть традиционные недостатки ДКМ-канала – его нестационарность и подверженность многолучевым замираниям. Это открывает перспективы использования КВ-каналов не только как резервных, но и как основных для передачи команд телеуправления, телеметрии и даже потокового видео с борта РТК в реальном времени.

Важно отметить, что техническая база для практической реализации таких решений уже существует. Разработаны и серийно выпускаются относительно компактные КВ-радиосредства с малыми массогабаритными характеристиками, пригодные для установки на различные типы РТК. В качестве возможного примера можно привести авиационную бортовую радиостанцию «Прима-КВ» (диапазон частот 2–30 МГц, масса передатчика не более 13 кг), которая поддерживает современные режимы адаптивной связи ALE и стандарты передачи данных, включая STANAG 4538 и MIL-STD-188-110B [3]. Наличие таких аппаратных средств в сочетании с предложенным алгоритмом когнитивного управления позволяет рассчитывать на практическую реализацию высоконадежных каналов управления РТК в ДКМ-диапазоне.

Однако существующие алгоритмы адаптации ДКМ-систем, регламентированные действующими стандартами [4–6], являются преимущественно реактивными: они изменяют параметры передачи (частоту, мощность, скорость) только после обнаружения факта ухудшения качества связи, что может приводить к потере управления или задержкам. Ключевой недостаток таких алгоритмов – отсутствие в контуре управления физической модели среды распространения, позволяющей прогнозировать состояние канала и его частотно-зависимые характеристики на основе независимых данных.

Целью настоящей работы является разработка и обоснование алгоритма, упреждающего (когнитивного) выбора оптимальной рабочей частоты и параметров передачи в ДКМ-канале, ориентированного на задачи управления РТК. Алгоритм основан на прогнозировании частотно-зависимых характеристик диффузной ионосферы по данным пассивного ГНСС-мониторинга.

1. Формальная постановка задачи. Рассматривается одномодовый ДКМ-канал связи, образованный однократным отражением волны от слоя F ионосферы. Канал характеризуется диффузной многолучевостью, вызванной рассеянием радиоволн на мелкомасштабных неоднородностях электронной концентрации с интенсивностью β_n и поперечным масштабом l_s [7, 8]. В районе трассы функционирует двухчастотный ГНСС-приёмник, обеспечивающий непрерывное измерение:

- ♦ наклонного полного электронного содержания (ПЭС) $I = \int N \, dl$;
- ♦ дисперсии флуктуаций фазы $\sigma_{\phi, GPS}^2$ (или индекса сцинтилляций S_4);
- ♦ скорости дрейфа неоднородностей $V_{др}$ (по доплеровским измерениям).

Цель работы – разработать алгоритм, который для текущего состояния ионосферы позволяет решить следующие подзадачи:

1. По данным ГНСС восстановить ключевые параметры ионосферы: β_n , l_s , критическую частоту $f_{кр}$, максимально применимую частоту (МПЧ) f_m и наименьшую применимую частоту (НПЧ).

2. Для всего диапазона рабочих частот f_0 от НПЧ до МПЧ вычислить прогнозируемые значения:

- ♦ СКО фазовых флуктуаций $\sigma_\varphi(f_0)$ с учётом масштабирования по частоте [7];
- ♦ параметр Накагами $m(f_0)$ по формуле, связывающей его с $\sigma_\varphi^2(f_0)$ [10];
- ♦ интервала частотной корреляции (полосы когерентности) $\Delta F_k(f_0)$ [7];
- ♦ интервала пространственной корреляции $\Delta \rho_k(f_0)$ [9];
- ♦ интервала временной корреляции $\Delta t_k(f_0) = \Delta \rho_k(f_0)/V_{др}$;
- ♦ среднего отношения сигнал/помеха $\bar{Z}(f_0)$ согласно рекомендаций ITU-R [11, 12];
- ♦ допустимого отношения сигнал/помеха $Z_{доп}(f_0, n)$ для заданной вероятности ошибки и числа ветвей разнесения n [13].

3. Определить частоту $f_0^{opt(n)}$, максимизирующую энергетический запас

$$\Delta Z(f_0, n) = \bar{Z}(f_0) - Z_{доп}(f_0, n).$$

4. Для найденной оптимальной частоты выбрать сигнально-кодую конструкцию (СКК) с максимальной скоростью, удовлетворяющую условию $\Delta F_c \leq \Delta F_k(f_0^{opt})$, также при необходимости обосновать разнос антенн для пространственного разнесения по приему ($\Delta r \geq \Delta \rho_k(f_0^{opt})$).

5. Определить требуемый запас мощности и установить глубину перемерения $Tил \geq 5\Delta t_k(f_0^{opt})$ для борьбы с временными замираниями.

Таким образом, в результате выполнения алгоритма будут определены все параметры, необходимые для адаптации к текущим условиям распространения: рабочая частота, полоса сигнала, глубина перемерения, а также обоснована возможность использования пространственного разнесения и необходимый энергетический запас по мощности.

2. Анализ существующих алгоритмов адаптации. Действующие алгоритмы адаптации ДКМ-радионаправлений, как отечественные [4], так и зарубежные [5, 6], построены по реактивной схеме. Они включают этапы мониторинга (активное зондирование или пассивный анализ), ранжирования частот, установления связи и адаптации в сеансе при превышении порога ошибок. Основные этапы и их отличия от предлагаемого прогностического подхода (ГНСС + модель) приведены в табл. 1.

Таблица 1

Сравнение этапов существующего и предлагаемого алгоритмов

Этап работы алгоритма адаптации	Алгоритм адаптации	
	По ГОСТ Р 55711-2013 [4]	Предлагаемый прогностический подход
1. Мониторинг среды	Активное зондирование тестовыми посылками, оценка $K_{ош}$	Пассивное определение $\beta_{и}$, I_s , $f_{кр}$, $V_{др}$ по данным ГНСС-мониторинга
2. Экстраполяция/прогноз	Отсутствует, работа «по факту»	Расчёт частотно зависимых $\sigma_\varphi^2(f_0)$, $m(f_0)$, $\Delta F_k(f_0)$, $\bar{Z}(f_0)$, $Z_{доп}(f_0, n)$ для всего интервала частот (от НПЧ до МПЧ)
3. Выбор частоты	Ранжирование по результатам зондирования	Поиск $f_0^{opt(n)}$ максимизирующую $\Delta Z(f_0, n)$
4. Выбор СКК	Жёсткая иерархия (снижение скорости, смена частоты)	Фильтр по полосе $\Delta F_c \leq \Delta F_k(f_0^{opt})$
5. Настройка интерливера	Фиксированная структура кадра	Выбор глубины: $Tил \geq 5\Delta t_k$ [13, 14]
6. Расчёт запаса	Фиксированные пороги по мощности	Динамический запас $Z_{доп}(f_0^{opt}, n)$
7. Реконфигурация	Через разрыв сеанса (квитанции)	Упреждающая смена параметров

Принципиальное ограничение существующих стандартов – отсутствие в контуре адаптации физической модели среды, связывающей параметры ионосферы с корреляционными характеристиками канала. Это не позволяет системе адаптироваться к быстрым изменениям диффузной многолучевости, что критически важно для высоких требований по надежности, предъявляемых системами управления РТК.

3. Предложения по реализации (Физическая модель и алгоритм). 3.1. Восстановление параметров ионосферы по данным ГНСС. По измеренному ПЭС и известной геометрии «спутник – приёмник» с использованием модели IRI [15] восстанавливаются критическая частота $f_{кр}$ и высота максимума h_m слоя F. Для заданной дальности связи R рассчитываются угол падения волны на ионосферу θ_0 и эквивалентный путь в ионосфере $S's(d)$ [16].

Ключевым этапом является оценка интенсивности мелкомасштабных неоднородностей β_i . Используя данные ГНСС ($\sigma_{\phi, GPS}^2(f_{GPS}) f_{GPS} \approx 1575$ МГц) и соотношения структурно-многолучевой модели [7], интенсивность неоднородностей может быть выражена через СКО флуктуаций электронной концентрации $\sigma_{\Delta N}(h)$ [17]:

$$\beta_i = \frac{\sigma_{\Delta N}(h)}{\bar{N}(h)}, \quad (1)$$

3.2. Прогнозирование частотно-зависимых характеристик ДКМ-канала. На основе восстановленных параметров для всего диапазона рабочих частот f_0 от НПЧ до МПЧ (с шагом Δf) вычисляются следующие функции.

Среднеквадратическое отклонение флуктуаций фазы на выходе ионосферы [7, 9]:

$$\sigma_{\phi}(f_0) = \pi \beta_i f_0 \sqrt{2L_0 L_3} / c K_s^2 \sec^2 \varphi_0, \quad (2)$$

где L_3 – эквивалентный однородный путь распространения волны в отражающем слое F ионосферы [м], который соответствует длине воображаемой криволинейной траектории распространения волны в «трубке» от входа до выхода отражающего слоя F ионосферы с неизменной (однородной) по высоте (h) средней электронной концентрацией $\bar{N}(h) = const$, соответствующей ее наибольшему значению на высоте $h=h_{от}$ отражения волны $\bar{N}(h_{от})$ [м⁻³], l_0 – наибольший масштаб неоднородностей, K_s – коэффициент сферичности Земли и ионосферы, φ_0 – угол падения волны на ионосферу, c – скорость света.

Интервал частотной корреляции (полоса когерентности) [7, 18]:

$$\Delta F_k(f_0) = \frac{f_0}{\sigma_{\phi}(f_0) \sqrt{2+d_1^2(f_0)}}, \quad (3)$$

где d_1 – дифракционный коэффициент, характеризующийся частотой и геометрией трассы [7, 18]. Полоса сигнала ΔF_c выбираемой СКК должна удовлетворять условию $\Delta F_c \leq \Delta F_k(f_0)$ для исключения частотно-селективных замираний.

Интервал пространственной корреляции [7, 9]:

В общем случае он зависит от $\sigma_{\phi}^2(f_0)$ и l_s . Уточнённое выражение, полученное в [9], имеет вид:

$$\Delta \rho_k(f_0) = l_s * \Phi(\sigma_{\phi}(f_0)), \quad (4)$$

где функция $\Phi(\sigma_{\phi}(f_0))$ определяется численно из условия $K_n(\Delta \rho_k) = e^{-1}$ для корреляционной функции [17]. Для инженерных расчётов можно использовать аппроксимацию:

$$\Phi(\sigma_{\phi}(f_0)) = \begin{cases} 1, & \sigma_{\phi}(f_0) \leq 1 \\ 1 - 0,8(\sigma_{\phi} - 1), & 1 < \sigma_{\phi} \leq 1,25 \\ 1/\sigma_{\phi}(f_0), & \sigma_{\phi}(f_0) > 1,25 \end{cases}.$$

Для обеспечения независимости ветвей разнесённого приёма необходимо выполнение условия разнесения по приему $\Delta r \geq \Delta \rho_k(f_0)$, где Δr – разнос антенн.

Интервал временной корреляции на основе гипотезы Тейлора о «замороженной» турбулентности [9, 19, 20]:

$$\Delta t_k(f_0) = \frac{\Delta \rho_k(f_0)}{V_{др}}. \quad (5)$$

Глубина перемежения выбирается из условия [14]

$$T_{ил} \geq 5\Delta t_k(f_0). \quad (6)$$

Параметр Накагами, характеризующий глубину замираний [10]:

$$m(f_0) = \frac{1}{1 - e^{-2\sigma_\phi^2(f_0)}}. \quad (7)$$

При $\sigma_\phi^2 \rightarrow 0$ (регулярный канал) $m \rightarrow \infty$; при $\sigma_\phi^2 \rightarrow \infty$ (рэлеевский канал) $m \rightarrow 1$.

Среднее отношение сигнал/помеха $\bar{Z}(f_0)$ определяется разностью медианных уровней сигнала и помех [11, 12]

$$\bar{Z}(f_0) = \bar{E}_c(f_0) - \bar{E}_n(f_0) \text{ (дБ)}. \quad (8)$$

Допустимое отношение сигнал/помеха для обеспечения заданной вероятности ошибки $P_{ош, доп}$ с учётом разнесённого приёма (n ветвей) [13]:

$$Z_{доп}(f_{0,n}) = 10 \lg \left[2m(f_0)(2P_{ош, доп})^{-1/(nm(f_0))} - 1 \right]. \quad (9)$$

3.3. Выбор оптимального режима работы. Для каждой частоты f_0 вычисляется энергетический запас

$$\Delta Z(f_0, n) = \bar{Z}_{f_0} - Z_{доп}(f_0, n). \quad (10)$$

Оптимальной считается частота $f_0^{opt}(n)$, на которой этот запас максимален. Соответствующая надёжность связи оценивается как

$$D_{св}(f_{0,n}) = F \left[\frac{\Delta Z(f_0, n)}{\sigma_z} \right], \quad (11)$$

где $F(\cdot)$ – функция Лапласа, $\sigma_z \approx 14$ дБ – среднеквадратическое отклонение превышения С/П [13].

Алгоритм реализует иерархию: выбор частоты $\rightarrow$ выбор СКК по условию $\Delta F_c \leq \Delta F_k \rightarrow$ определение глубины перемежения $T_{ил} \rightarrow$ обоснование разнесения $\Delta r \geq \Delta r_k \rightarrow$ корректировка мощности.

Предложенная иерархия полностью обоснована физическими зависимостями: частота первична, затем следуют параметры сигнала, адаптируемые под полосу когерентности, затем временная и пространственная корреляция, и наконец энергетика.

Блок-схема предлагаемого алгоритма представлена на рис. 1.

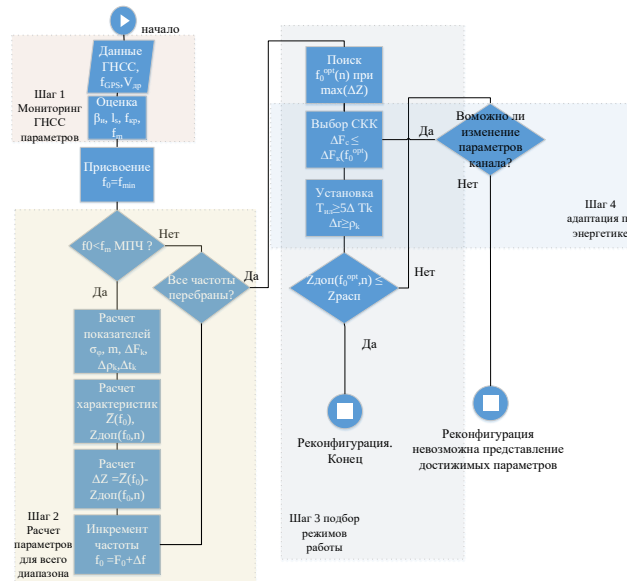


Рис. 1. Блок-схема алгоритма выбора оптимальных параметров передачи

4. Результаты численного моделирования и их обсуждение. Для верификации предложенного подхода проведено численное моделирование для типовой среднеширотной радиолинии протяжённостью $R = 2000$ км в дневных условиях. Исходные данные (по [7, 13]): $f_{кр} = 7$ МГц, $f_m = 12,1$ МГц, $h_0 = 250$ км, $l_s = 200$ м, $K_s = 1,1$, $V_{др} = 100$ м/с, $P_{ош, доп} = 3 \cdot 10^{-3}$, $\sigma_z = 14$ дБ. Рассматривались два уровня диффузности ионосферы: слабая ($\beta_n = 10^{-3}$) и сильная ($\beta_n = 10^{-1}$).

Значения $\sigma_{\phi, GPS}$ для этих уровней выбирались из условия, что на частоте $f_0 = 10$ МГц СКО флуктуации фазы составляют 0,5 рад (слабая диффузность) и 3,0 рад (сильная диффузность). На основе этих данных по формулам (2)-(11) рассчитаны соответствующие параметры. В табл. 2 и 4 приведены зависимости основных параметров канала от частоты для слабой и сильной диффузностей соответственно. В табл. 3 и 5 представлены энергетический запас и надёжность связи с учетом организации разнесенного приема. Сводные результаты на оптимальных частотах представлены в табл. 6.

Таблица 2

Расчётные значения параметров канала для случая $\beta_n = 10^{-3}$ (слабая диффузность)

f_0 , МГц	σ_{ϕ}	m	ΔF_k , кГц	$\bar{Z}$, дБ	$\Delta \rho_k$, м	Δt_k , с	$T_{ил}$, с
2	0,1	50	29,6	45	200	2	10
4	0,2	12,5	27,4	43	200	2	10
6	0,3	5,56	24,9	41	200	2	10
8	0,4	3,16	22,3	39	200	2	10
10	0,5	2,16	20	37	200	2	10
10,5	0,525	1,99	19,2	36	200	2	10
12	0,6	1,64	17,2	35	200	2	10

Видно, что с ростом частоты σ_{ϕ} растёт, параметр Накагами m уменьшается (замирания становятся глубже), полоса когерентности ΔF_k сужается – всё это соответствует физике и базовым работам [7, 18]. Значения ΔF_k лежат в типичном диапазоне 15–30 кГц.

Таблица 3

Энергетический запас и надёжность связи с учетом организации разнесенного приема для случая $\beta_n = 10^{-3}$ (слабая диффузность)

f_0 , МГц	$Z_{доп} (n=1)$, дБ	$\Delta Z (n=1)$	$D_{св} (n=1)$	$Z_{доп} (n=2)$, дБ	$\Delta Z (n=2)$	$D_{св} (n=2)$
2	20,4	14,6	0,851	20,2	14,8	0,855
4	15,6	20,4	0,927	14,7	21,3	0,936
6	14,3	22,7	0,947	12,2	24,8	0,962
8	14,9	24,1	0,957	11,2	27,8	0,976
10	16,6	24,4	0,96	11,2	29,8	0,983
10,5	17,1	25,9	0,968	11,3	31,7	0,988
12	18,7	26,4	0,97	11,6	33,4	0,991

По критерию максимизации энергетического запаса ΔZ оптимальная частота лежит в районе МПЧ (12 МГц 26,4 дБ соответственно), обеспечивая надёжность связи на уровне – 0,97. Разнесенный двоякный прием повышает и энергетический запас (33,4 дБ) и надёжность связи до 0,991, однако полоса когерентности в этом случае составит 17,2 кГц, что в 1,68 раза меньше максимально достижимого в диапазоне частот значения (29,6 кГц на НПЧ=2 МГц), время интерливера во всем диапазоне (задержка перемежения) составит 10 секунд.

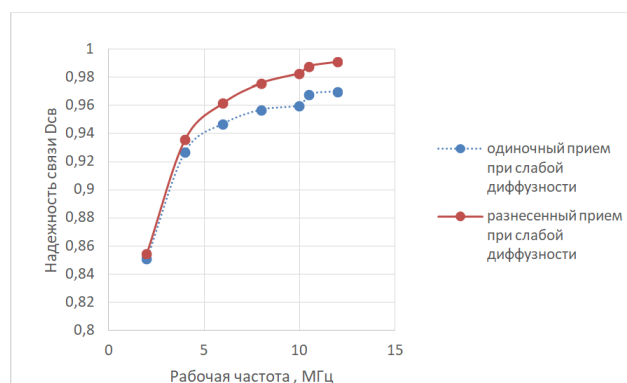


Рис. 2. Характер зависимости надежности связи в расчетном примере для одиночного и двойного разнесенного приема при слабой диффузности

Для ситуаций с сильной диффузностью ситуация несколько меняется.

Таблица 4

Расчётные значения параметров канала для случая $\beta n = 10^{-1}$ (сильная диффузность)

f_0 , МГц	σ_ϕ	m	ΔF_k , кГц	$\bar{Z}$, дБ	$\Delta \rho_k$, м	Δt_k , с	$T_{пл}$, с
2	0,6	1,64	4,9	45	200	2	10
4	1,2	1,07	4,2	43	167	1,67	8,35
6	1,8	1,01	3,6	41	111	1,11	5,55
8	2,4	1	3	39	83	0,83	4,15
10	3	1	2,5	37	67	0,67	3,35
10,5	3,15	1	2,3	36	63	0,63	3,15
12	3,6	1	2	35	56	0,56	2,8

Здесь σ_ϕ велико, поэтому m близко к 1, замирания практически релеевские. Полоса когерентности ΔF_k значительно уже, чем при слабой диффузности, и также убывает с частотой. Интервал пространственной корреляции $\Delta \rho_k$ вычислен по формуле (4) с учётом $\sigma_\phi > 1,25$: $\Delta \rho_k = l_s / \sigma_\phi$, что приводит к уменьшению требуемого разнosa антенн на высоких частотах – полученные значения подтверждаются соответствием базовым физическим процессам в ДКМ радиосвязи.

Таблица 5

Энергетический запас и надежность связи с учетом организации разнесенного приема для случая $\beta n = 10^{-1}$ (сильная диффузность)

f_0 , МГц	$Z_{доп} (n=1)$, дБ	$\Delta Z (n=1)$	$D_{св} (n=1)$	$Z_{доп}(n=2)$, дБ	$\Delta Z (n=2)$	$D_{св} (n=2)$
2	18,7	16,3	0,879	11,6	23,4	0,952
4	24,1	11,9	0,803	13,5	22,5	0,946
6	25	12	0,804	13,9	23,1	0,951
8	25,2	13,8	0,837	14	25	0,963
10	25,2	15,8	0,87	14	27	0,973
10,5	25,2	17,8	0,898	14	29	0,981
12	25,2	19,8	0,921	14	31	0,987

Из-за сильной диффузности почти пропадает маневр энергетического запаса за счет $Z_{\text{доп}}$ – допустимого отношения сигнал/помеха (оно теперь почти на всех частотах максимальное и составляет 25 дБ). Энергетический выигрыш по-прежнему фиксируется на МПЧ, но в абсолютных значениях этот запас (19,8 дБ) меньше энергетического выигрыша, достигаемого в условиях слабой диффузности (26,4 дБ). Соответственно и достижимые значения для надежности связи несколько ниже, чем для ситуации со слабой диффузностью (рис. 3). Фактически без разнесенного приема не удастся обеспечивать надежность связи выше 0,9.

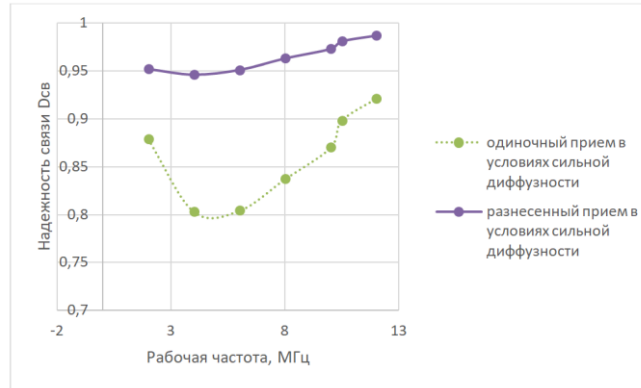


Рис. 3. Характер зависимости надежности связи в расчетном примере для одиночного и двойного разнесенного приема при сильной диффузности

Таблица 6

Оптимальные режимы работы

$\beta_{\text{и}}$	$f_{\text{опт}}$, МГц	n	$Z_{\text{доп}}$, дБ	ΔZ , дБ	$D_{\text{св}}$	Тил, с
10^{-3}	10,5	1	15,4	20,6	0,929	10
10^{-3}	10,5	2	11,1	24,9	0,962	10
10^{-1}	12	1	25,2	9,8	0,758	4
10^{-1}	12	2	14	21	0,933	4

Обсуждение результатов. Во всех рассмотренных случаях максимум энергетического запаса ΔZ достигается на максимально применимой частоте (12 МГц). Это обусловлено ростом $\bar{Z}(f_0)$ с частотой, что соответствует типичной физической картине (уменьшение поглощения при увеличении частоты). При слабой диффузности разнесённый приём ($n=2$) повышает надёжность с 0,970 до 0,991 и может быть использован для перехода к работе на более низких частотах с целью использования СКК с более широкой полосой сигнала или скоростью; при сильной диффузности разнесенный прием позволяет повысить надежность с 0,921 до 0,987 и в принципе без разнесённого приема отсутствует возможность обеспечить надежность связи выше 0,95. Полоса когерентности при сильной диффузности резко сужается (до 2 кГц), что ограничивает максимальную скорость передачи (рис. 4). Поэтому алгоритм автоматически исключит широкополосные СКК на этапе выбора. Пространственный разнос антенн для сильной диффузности требуется всего 56 м против 200 м при слабой, что упрощает реализацию разнесённого приёма в сложных условиях.

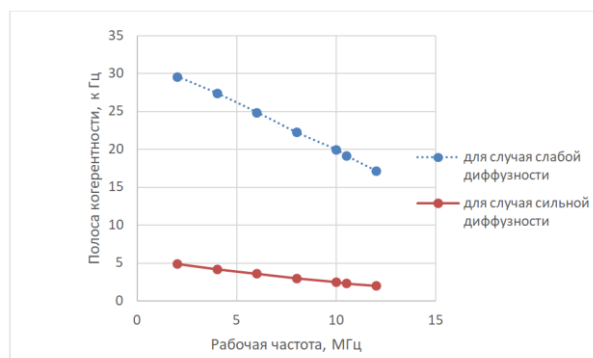


Рис. 4. Зависимость интервала частотной корреляции от рабочей частоты при различных уровнях диффузности.

Выводы:

1. Разработан алгоритм упреждающего выбора оптимальной рабочей частоты и параметров передачи в декаметровом канале связи, ориентированный на задачи управления робототехническими комплексами. В отличие от существующих реактивных методов [4–6], алгоритм использует данные пассивного ГНСС-мониторинга для прогнозирования частотно-зависимых характеристик диффузной ионосферы – среднеквадратического отклонения флуктуаций фазы σ_ϕ , интервалов частотной, пространственной и временной корреляции (ΔF_k , $\Delta \rho_k$, Δt_k), параметра Накагами m и допустимого отношения сигнал/помеха $Z_{\text{доп}}$. Это позволяет адаптироваться к текущему состоянию среды без активного зондирования и задержек на установление связи, что критически важно для обеспечения требуемой надёжности и скорости передачи при управлении подвижными объектами.

2. Новизна подхода заключается в комплексном использовании физической модели среды распространения, объединяющей структурно-многолучевое описание [7, 9] с уточнёнными зависимостями для интервалов корреляции [17, 18]. Благодаря корректному учёту частотной зависимости $\sigma_\phi(f_0)$, полученной из фундаментальных соотношений [7], все рассчитываемые параметры демонстрируют правильное физическое поведение: с ростом частоты глубина замираний увеличивается (m уменьшается), полоса когерентности ΔF_k сужается, а интервал пространственной корреляции $\Delta \rho_k$ сокращается, что согласуется с результатами базовых работ [7, 18] и экспериментальными данными.

3. Численное моделирование для типовой среднеширотной трассы подтвердило работоспособность алгоритма. Показано, что при слабой диффузности ионосферы максимум энергетического запаса ΔZ и, соответственно, оптимальная рабочая частота лежат вблизи традиционной оптимальной рабочей частоты (0,85–0,9 fm), а при сильной диффузности оптимум смещается к максимально применимой частоте. Использование разнесённого приёма ($n=2$) обеспечивает существенное повышение надёжности связи (с 0,921 до 0,987 в условиях сильной диффузности). При этом автоматически учитываются ограничения, накладываемые полосой когерентности и временем перемежения, что гарантирует выбор сигнально-кодовых конструкций, исключающих частотно-селективные замирания.

4. Практическая значимость работы состоит в том, что разработанный алгоритм может быть реализован в программном обеспечении современных ДКМ-радиостанций и модемов, включая уже существующие компактные образцы типа «Прима-КВ» [3]. Это позволит повысить качество и устойчивость каналов управления робототехническими комплексами морского и воздушного базирования. Дальнейшие исследования предполагают экспериментальную проверку алгоритма на действующих трассах, его адаптацию для многолучевых каналов, а также уточнение модели частотной корреляции путём введения частотно-зависимого дифракционного параметра $d1(f_0)$ согласно методике [18].

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Будко П.А., Винограденко А.М., Жуков Г.А., Литвинов А.И. Комплексное использование разнородных каналов связи для управления робототехническими комплексами на базе единой системы радиомониторинга // Научные технологии в космических исследованиях Земли. – 2017. – Т. 9, № 1. – С. 18-41.
2. Будко П.А., Винограденко А.М., Литвинов А.И. Реконфигурация каналов связи при управлении смешанными группировками робототехнических комплексов // Известия ЮФУ. Технические науки. – 2017. – № 1. – С. 266-278. – DOI: 10.18522/2311-3103-2017-1-266278.
3. Прима-KB (с режимом ALE). НПП «ПРИМА». – URL: <https://www.prima-systems.ru/catalog/main/product/50/> (дата обращения: 18.03.2026).
4. ГОСТ Р 55711-2013. Комплекс технических средств автоматизированной адаптивной ВЧ (KB) дуплексной радиосвязи. Алгоритмы работы. – М.: Стандартинформ, 2014.
5. STANAG 4538. Technical Standards for an Automatic Radio Control System (ARCS) for HF Communication. – NATO, 2014.
6. MIL-STD-188-141B. Interoperability and Performance Standards for Medium and High Frequency Radio Systems. – Department of Defense, USA, 1999.
7. Пашинцев В.П., Коваль С.А., Цимбал В.А., Тошкин В.Е., Сенокосов М.А., Скорик А.Д. Структурно-многолучевой подход к разработке пространственно-временной модели одномодового декаметрового канала связи с диффузной многолучевостью // Журнал радиоэлектроники. – 2022. – № 6. – DOI 10.30898/1684-1719.2022.6.5.
8. Blaunstein N., Plohotniuc E. Ionosphere and Applied Aspects of Radio Communication and Radar. – New York: Taylor & Francis, 2008. – 577 p.
9. Пашинцев В.П., Коваль С.А., Потягов Д.А., Скорик А.Д., Сенокосов М.А. Уточненный метод определения интервала пространственной корреляции замираний в однолучевой декаметровой радиолнии // Журнал радиоэлектроники. – 2021. – № 2. – DOI: 10.30898/1684-1719.2021.2.5.
10. Nakagami M. The m-distribution – A general formula of intensity distribution of rapid fading // Statistical Methods in Radio Wave Propagation. – Oxford: Pergamon Press, 1960. – P. 3-36.
11. Recommendation ITU-R P.533-13. Method for the prediction of the performance of HF circuits. – ITU, 2015.
12. Recommendation ITU-R P.372-12. Radio noise. – ITU, 2015.
13. Пашинцев В.П., Скорик А.Д., Коваль С.А., Киселев Д.П., Сенокосов М.А. Зависимость надежности связи в декаметровой радиолнии от выбора рабочей частоты с учетом сигнально-помеховой обстановки и диффузности ионосферы // Системы управления, связи и безопасности. – 2019. – № 4. – С. 300-322. – DOI: 10.24411/2410-9916-2019-10412.
14. Кловский Д.Д. Передача дискретных сообщений по радиоканалам. – М.: Радио и связь, 1982. – 304 с.
15. Bilitza D., Altadill D., Truhlik V., Shubin V., Galkin I., Reinisch B., Huang X. International Reference Ionosphere 2016: From ionospheric climate to real-time weather predictions // Space Weather. – 2017. – Vol. 15. – P. 418-429. – DOI: 10.1002/2016SW001593.
16. Пашинцев В.П., Тишкин С.А., Смирнов А.А., Боровлев И.И. Эквивалентный путь распространения декаметровой волны в сферически-слоистой ионосфере // Журнал радиоэлектроники. – 2001. – № 8. – URL: <http://jre.cplire.ru/jre/aug01/1/text.html> (дата обращения: 10.03.2026).
17. Стейн С., Джонс Дж. Принципы современной теории связи и их применение к передаче дискретных сообщений. – М.: Связь, 1971. – 376 с.
18. Пашинцев В.П., Скорик А.Д., Коваль С.А., Алексеев Д.В., Сенокосов М.А. Алгоритм расчета интервала частотной корреляции коротковолновой радиолнии с учетом сферичности и мелко-масштабных неоднородностей ионосферы // Системы управления, связи и безопасности. – 2020. – № 2. – С. 49-72. – DOI 10.24411/2410-9916-2020-10203.
19. Taylor G.I. The spectrum of turbulence // Proceedings of the Royal Society of London. Series A. – 1938. – Vol. 164, No. 919. – P. 476-490. – DOI: 10.1098/rspa.1938.0032.
20. Pashintsev V.P., Peskov M.V., Kiselev N.V., Mikhailov D.A., Dukhovnyi D.V. Analysis of noise immunity of satellite communications under small-scale. - ionospheric disturbances and time-selective fading of received signals // The Egyptian Journal of Remote Sensing and Space Sciences. – 2023. – 26. – P. 1036-1045. – DOI: 10.1016/j.ejrs.2023.11.002.

REFERENCES

1. Budko P.A., Vinogradenko A.M., Zhukov G.A., Litvinov A.I. Kompleksnoe ispol'zovanie raznorodnykh kanalov svyazi dlya upravleniya robototekhnicheskimi kompleksami na baze edinoi sistemy radiomonitoringa [Integrated use of heterogeneous communication channels for control of robotic complexes based on a single radio monitoring system], *Nauchnye tekhnologii v kosmicheskikh issledovaniyakh Zemli* [High-Tech Technologies in Earth Space Research], 2017, Vol. 9, No. 1, pp. 18-41.
2. Budko P.A., Vinogradenko A.M., Litvinov A.I. Rekonfiguratsiya kanalov svyazi pri upravlenii smeshannymi gruppировками robototekhnicheskikh kompleksov [Reconfiguration of communication channels when controlling mixed groups of robotic systems], *Izvestiya YuFU. Tekhnicheskie nauki* [Izvestiya SFedU. Engineering Sciences], 2017, No. 1, pp. 266-278. DOI: 10.18522/2311-3103-2017-1-266278.

3. Prima-KV (s rezhimom ALE) [Prima-KV (with ALE mode)], NPP «PRIMA». Available at: <https://www.prima-systems.ru/catalog/main/product/50/> (accessed 18 March 2026).
4. GOST R 55711-2013. Kompleks tekhnicheskikh sredstv avtomatizirovannoy adaptivnoy VCH (KV) dupleksnoy radiosvyazi. Algoritmy raboty [GOST R 55711-2013. Complex of technical means of automated adaptive HF (HF) duplex radio communication. Operating algorithms]. Moscow: Standartinform, 2014.
5. STANAG 4538. Technical Standards for an Automatic Radio Control System (ARCS) for HF Communication. NATO, 2014.
6. MIL-STD-188-141B. Interoperability and Performance Standards for Medium and High Frequency Radio Systems. Department of Defense, USA, 1999.
7. Pashintsev V.P., Koval' S.A., Tsimbal V.A., Toshkin V.E., Senokosov M.A., Skorik A.D. Strukturno-mnogoluchevoy podkhod k razrabotke prostranstvenno-vremennoy modeli odnomodovogo deka-metrovogo kanala svyazi s diffuznoy mnogoluchevost'yu [Structural-multipath approach to the development of a space-time model of a single-mode decameter communication channel with diffuse multipath], *Zhurnal radioelektroniki* [Journal of Radio Electronics], 2022, No. 6. DOI 10.30898/1684-1719.2022.6.5.
8. Blaunstein N., Plohotniuc E. Ionosphere and Applied Aspects of Radio Communication and Radar. New York: Taylor & Frances, 2008, 577 p.
9. Pashintsev V.P., Koval' S.A., Potyagov D.A., Skorik A.D., Senokosov M.A. Utochnennyy metod opredeleniya intervala prostranstvennoy korrelyatsii zamiraniy v odnoluchevoy dekametrovoy radiolinii [Refined method for determining the spatial correlation interval of fading in a single-hop decameter radio link], *Zhurnal radioelektroniki* [Journal of Radio Electronics], 2021, No. 2. DOI: 10.30898/1684-1719.2021.2.5.
10. Nakagami M. The m-distribution – A general formula of intensity distribution of rapid fading, *Statistical Methods in Radio Wave Propagation*. Oxford: Pergamon Press, 1960, pp. 3-36.
11. Recommendation ITU-R P.533-13. Method for the prediction of the performance of HF circuits. ITU, 2015.
12. Recommendation ITU-R P.372-12. Radio noise. ITU, 2015.
13. Pashintsev V.P., Skorik A.D., Koval' S.A., Kiselev D.P., Senokosov M.A. Zavisimost' nadezhnosti svyazi v dekametrovoy radiolinii ot vybora rabochey chastoty s uchetom signal'no-pomekhovoy obstanovki i diffuznosti ionosfery [Dependence of communication reliability in a decameter radio line on the choice of operating frequency, taking into account the signal-to-interference environment and ionospheric diffusivity], *Sistemy upravleniya, svyazi i bezopasnosti* [Systems of Control, Communication and Security], 2019, No. 4, pp. 300-322. DOI: 10.24411/2410-9916-2019-10412.
14. Klovskiy D.D. Peredacha diskretnykh soobshcheniy po radiokanalam [Transmission of discrete messages over radio channels]. Moscow: Radio i svyaz', 1982, 304 p.
15. Bilitza D., Altadill D., Truhlik V., Shubin V., Galkin I., Reinisch B., Huang X. International Reference Ionosphere 2016: From ionospheric climate to real-time weather predictions, *Space Weather*, 2017, Vol. 15, pp. 418-429. DOI: 10.1002/2016SW001593.
16. Pashintsev V.P., Tishkin S.A., Smirnov A.A., Borovlev I.I. Ekvivalentnyy put' rasprostraneniya dekametrovoy volny v sfericheski-sloistoy ionosfere [Equivalent propagation path of a decameter wave in a spherically layered ionosphere], *Zhurnal radioelektroniki* [Journal of Radio Electronics], 2001, No. 8. – Available at: <http://jre.cplire.ru/jre/aug01/1/text.html> (Accessed 10 March 2026).
17. Steyn S., Dzhons Dzh. Printsipy sovremennoy teorii svyazi i ikh primeneniye k peredache diskretnykh soobshcheniy [Principles of modern communication theory and their application to the transmission of discrete messages]. Moscow: Svyaz', 1971, 376 p.
18. Pashintsev V.P., Skorik A.D., Koval' S.A., Alekseev D.V., Senokosov M.A. Algoritm rascheta intervala chastotnoy korrelyatsii korotkovolnovoy radiolinii s uchetom sferichnosti i melko-masshtabnykh neodnorodnostey ionosfery [Algorithm for calculating the frequency correlation interval of a short-wave radio link taking into account the sphericity and small-scale inhomogeneities of the ionosphere], *Sistemy upravleniya, svyazi i bezopasnosti* [Systems of Control, Communication and Security], 2020, No. 2, pp. 49-72. DOI 10.24411/2410-9916-2020-10203.
19. Taylor G.I. The spectrum of turbulence, *Proceedings of the Royal Society of London. Series A*, 1938, Vol. 164, No. 919, pp. 476-490. DOI: 10.1098/rspa.1938.0032.
20. Pashintsev V.P., Peskov M.V., Kiselev N.V., Mikhailov D.A., Dukhovnyi D.V. Analysis of noise immunity of satellite communications under small-scale. - ionospheric disturbances and time-selective fading of received signals, *The Egyptian Journal of Remote Sensing and Space Sciences*, 2023, 26, pp. 1036-1045. DOI: 10.1016/j.ejrs.2023.11.002.

Скорик Александр Дмитриевич – АО «Концерн ВКО «Алмаз-Антей»; e-mail: a.skorik@almaz-antey.ru; г. Москва, Россия; тел.: +74952762901; советник генерального конструктора.

Skorik Alexander Dmitrievich – Almaz-Antey Concern; e-mail: a.skorik@almaz-antey.ru; Moscow, Russia; phone: +74952762901; adviser to the general designer.

М.М. Сорокин, Л.А. Гладков, Н.В. Гладкова**СОВРЕМЕННЫЕ ПОДХОДЫ К РЕШЕНИЮ ЗАДАЧИ 3D УПАКОВКИ**

Статья посвящена рассмотрению современных тенденций и подходов к решению актуальной оптимизационной задачи трехмерной упаковки в контейнеры. Важность построения эффективных методов решения данной задачи обусловлена бурным ростом направления электронной коммерции, когда достижение даже инкрементных улучшений плотности заполнения контейнеров грузами может привести к значительным сокращениям затрат на перевозку и хранение грузов. В статье приведен анализ различных типов задач и предложена классификация задач упаковки по различным критериям, в том числе: офлайн- и онлайн-упаковка, по размерности, по видам и количеству контейнеров и грузов. Приведена постановка классической оптимизационной задачи о рюкзаке и рассмотрены различные варианты ограничений, обусловленные особенностями решаемых задач. Приведен краткий обзор основных подходов к решению задачи. Выполнен анализ и обобщение характерных особенностей применения метаэвристических подходов, основанных на использовании эволюционных и биоинспирированных алгоритмов и методов машинного обучения, отмечены их достоинства и недостатки. В силу сложности рассматриваемой задачи предлагается активно использовать известные и разрабатывать новые модификации метаэвристических алгоритмов, которые позволяют находить квазиоптимальные решения за полиномиальное время. Приведен анализ известных методов машинного обучения, биоинспирированных алгоритмов, описаны принципы их работы, выделены их основные особенности, преимущества и недостатки, отмечена перспективность их разработки и применения для решения NP-полных задач комбинаторной оптимизации. Приведен обобщенный принцип работы метаэвристических алгоритмов. Сравнительный анализ применения различных методов оптимизации показал эффективность использования метаэвристических методов для решения задачи трехмерной упаковки.

Методы оптимизации; задача трехмерной упаковки; метаэвристические алгоритмы; биоинспирированные методы; машинное обучение.

M.M. Sorokin, L.A. Gladkov, N.V. Gladkova**MODERN APPROACHES TO SOLVING THE 3D BIN PACKING PROBLEM**

The article is devoted to the consideration of current trends and approaches to solving the urgent optimization problem of three-dimensional bin packing problem. The importance of building effective methods for solving this problem is due to the rapid growth of e-commerce, where achieving even incremental improvements in container filling density can lead to significant reductions in freight transportation and storage costs. The article provides an analysis of various types of problems and suggests a classification of bin packing problems according to various criteria, including: offline and online packing, by dimension, by type and quantity of containers and cargo. The formulation of the classical optimization knapsack problem is given and various options for constraints due to the specifics of the tasks being solved are considered. A brief overview of the main approaches to solving the problem is given. The analysis and generalization of the characteristic features of the application of metaheuristic approaches based on the use of evolutionary and bioinspired algorithms and machine learning methods is carried out, their advantages and disadvantages are noted. Due to the complexity of the problem under consideration, it is proposed to actively use known and develop new modifications of metaheuristic algorithms that make it possible to find quasi-optimal solutions in polynomial time. The analysis of known machine learning methods and bioinspired algorithms is given, the principles of their operation are described, their main features, advantages and disadvantages are highlighted, and the prospects for their development and application to solve NP-complete combinatorial optimization problems are noted. A generalized principle of operation of metaheuristic algorithms is given. A comparative analysis of the application of various optimization methods has shown the effectiveness of using metaheuristic methods to solve the problem of three-dimensional packaging.

Optimization methods; three-dimensional bin packing problem; metaheuristic algorithms; bioinspired methods; machine learning.

Введение. Актуальность разработки новых методов решения задачи упаковки, позволяющих повысить плотность и возможную загрузку контейнеров, обусловлена практической ценностью и возможностью применения в самых различных прикладных областях. Область практического использования данной задачи необычайно широка: заполнение и погрузка контейнеров в различных видах транспортных перевозок (ж/д, авиа- и морские перевозки), логистика, задачи оптимизации перемещения и складирования товаров, распределение товаров по складам и партиям. Необходимость интеграции экономик различных регионов и создание единого внутрироссийского экономического пространства определяют требования к пропускной способности транспортных терминалов и складов временного хранения. Соответственно, скорость и стоимость доставки товаров между производителями и потребителем напрямую зависят от эффективности работы информационных процессов на складах, моделирующих процессы разгрузки контейнеров, их временное хранение на складе и их отправку. Актуальность задачи и ее экономическая значимость растут по мере активного развития электронной коммерции, которая трансформирует логистические процессы. Так, в 2023 году только в Китае службы доставки обработали 37.1 млрд отправок, продемонстрировав рост на 17% по сравнению с предыдущим годом [1].

Кроме того, специфика электронной коммерции (а именно, преобладание высокочастотных заказов малого объема с широким ассортиментом товарных позиций, необходимость работы с разнотипной тарой и необходимость принятия решений в режиме реального времени) принципиально усложняет логистику и значимость эффективности методов упаковки [1, 2]. Это превращает классическую задачу трехмерной упаковки в острую операционную проблему, решение которой имеет не только экономическое, но и экологическое значение. Нерациональная упаковка приводит к увеличению объемов отходов и представляет серьезную угрозу для окружающей среды, а максимизация плотности упаковки определяет суммарные операционные расходы и общую устойчивость логистической цепочки производителя [1, 3]. Таким образом, разработка новых и модернизация существующих алгоритмов оптимизации, способных находить эффективные решения в условиях высокой неопределенности и множественных ограничений, является критически важным фактором повышения эффективности работы терминалов, складов и логистических центров. А задача создания новых моделей, методов и алгоритмов оптимизации таких информационных процессов является актуальной научной задачей и составляет научно-техническую проблему, имеющую важное народно-хозяйственное значение.

Классификация типов задач упаковки. Целью решения задачи трехмерной упаковки является построение эффективного варианта размещения объектов в трехмерном пространстве с одновременной максимизацией плотности их размещения [4]. Точное решение данной задачи невозможно получить за полиномиальное время, поскольку данная задача относится к классу NP-полных и NP-трудных задач. Поэтому целью, в данном случае, является получение «квазиоптимального» решения за приемлемое время. Как показывает практика, наиболее эффективными подходами к решению задачи трехмерной упаковки являются метаэвристические алгоритмы, в том числе генетические, биоинспирированные, роевые методы [5–8]. Получение качественного решения в этом случае, как правило, достигается путем итеративного улучшения решений, в том числе в условиях неопределенности и в нечетких условиях.

Задачи трехмерной упаковки подразделяются на два класса: задачи оффлайн- и онлайн-упаковки. Задачи оффлайн-упаковки работают с предоставленным полным набором необходимых исходных данных о грузах, в них возможно выполнить моделирование последовательности упаковки грузов или сортировки грузов. В задачах онлайн-упаковки требуется незамедлительная погрузка поступающих поодиночке грузов в контейнеры в режиме реального времени, без полного предварительного знания о них, что повышает сложность задачи.

Задача упаковки также может быть классифицирована по количеству измерений: одно-, двух- и трехмерная упаковка. При одномерной упаковке решается задача распределения участков на ленте заданной длины, ширина ленты считается несущественной

или фиксированной. Примерами задачи одномерной упаковки являются: задача о рюкзаке, задача разрезания стальных прутьев определенной длины, задача распределения времени между процессами в операционных системах.

При двумерной упаковке решается задача распределения двумерных объектов (прямоугольной или произвольной формы) в двухмерных контейнерах. Типичным примером задачи двумерной упаковки является задача раскроя листов стекла или металла с целью минимизации отходов.

Самая сложная задача – трехмерная упаковка – подразумевает распределение объемных объектов в трехмерном пространстве. Как правило, при трехмерной упаковке рассматриваются объекты в форме параллелепипеда (которые зачастую моделируют упакованные товары), но возможны также варианты задачи трехмерной упаковки для объектов сложной формы. Модели задач трехмерной упаковки применяются, как правило, при погрузке в контейнеры с учетом практических особенностей различных предметных областей, но также возможно применение данных моделей в других прикладных областях, например, при 3D-печати методом запекания порошка. В более абстрактном понимании, разновидностью задачи трехмерной упаковки является задача построения расписаний, в которой одной из трех координат пространства упаковки является время.

Среди различных типов задач трехмерной упаковки в контейнеры можно выделить следующие четыре класса:

- ♦ упаковка в полосу. При этом используется один «контейнер», два измерения которого фиксированы, а третье является переменным. Целью является минимизация длины контейнера по переменному измерению;
- ♦ упаковка в контейнеры – выполняется поиск наименьшего числа контейнеров, одинаковой размерности и стоимости использования, в которое можно распределить все имеющиеся объекты;
- ♦ многоконтейнерная упаковка. Используется множество контейнеров разных размеров и стоимости. Цель – минимизация суммарной стоимости контейнеров, которые были использованы при решении задачи упаковки;
- ♦ задача о рюкзаке подразумевает распределение нескольких типов упаковываемых объектов (все объекты одного типа имеют одинаковый размер) в один контейнер фиксированного объема. Цель – максимизация плотности упаковки объектов.

Постановка задачи. Задача трехмерной упаковки может быть сформулирована следующим образом. Необходимо разместить в трехмерном пространстве ограниченного размера максимальное число элементов различной формы и объема. Исходными данными являются размер контейнера и размеры располагаемых элементов. Каждый i -й элемент описывается тремя параметрами: ширина w_i , длина l_i и высота h_i , $i = 1, 2, \dots, n$. Дана область контейнера, имеющая форму параллелепипеда с шириной W , длиной L и высотой H . Расположение каждого i -го элемента в контейнере описывается координатами двух его углов: $\langle x_1, y_1, z_1 \rangle$, наиболее близкого к началу координат, и $\langle x_2, y_2, z_2 \rangle$, наиболее удаленного от начала координат. Решение задачи представляется в следующей форме:

$$S = \{s_i = (\langle x_{1i}, y_{1i}, z_{1i} \rangle, \langle x_{2i}, y_{2i}, z_{2i} \rangle) | i = 1, 2, \dots, n\}.$$

Тогда целью является минимизация объема пустот в области упаковки и, следовательно, целевую функцию можно записать в следующем виде:

$$\text{ЦФ} = \frac{\sum_{i=1}^N v_i}{V} \rightarrow 1,$$

где v_i – объем i -го элемента, V – объем ограничивающего параллелепипеда, требующегося для упаковки всех элементов (называемый областью упаковки).

При размещении объектов учитываются следующие базовые ограничения:

1. Ни один объект не должен выходить за границы контейнера.
2. Суммарный объем объектов не должен превышать объем области упаковки.
3. Отсутствуют пересечения объектов друг с другом.
4. Объекты располагаются параллельно граням области упаковки.
5. Все точки основания всех объектов лежат либо на дне области упаковки, либо на верхних гранях других объектов.

В зависимости от особенностей конкретной предметной области может потребоваться введение дополнительных ограничений в используемую модель задачи упаковки. Рассмотрим некоторые из них:

- ◆ ограничение возможной ориентации объекта (например, необходимость перевезти товар вертикально);
- ◆ ограничение веса, который можно располагать над объектом;
- ◆ ограничения на расположение объекта при транспортировке (например, необходимость располагать крупногабаритные объекты на полу контейнера или фиксировать их положение ниже определенной высоты; установка определенных типов объектов рядом с дверцей контейнера);
- ◆ ограничения на устойчивость груза (необходимость обеспечить устойчивость груза при перевозке и при проведении погрузки или разгрузки контейнера);
- ◆ разные типы группировки объектов, в том числе:
- ◆ по получателю и по типу товара;
- ◆ по пункту назначения (когда контейнер предназначен для отправки грузов в несколько пунктов назначения, нужно располагать грузы, предназначенные для одного пункта, в тесном порядке друг с другом);
- ◆ разделение различных типов объектов в контейнере (например, необходимость отделить пищевые товары от остальных категорий товаров);
- ◆ равномерное распределение веса внутри контейнера (с помощью вычисления центра масс);
- ◆ ограничение по суммарному весу контейнера (при тяжелом грузе может оказаться более значимым ограничением, чем объем контейнера).

При этом важно отметить, что дополнительные ограничения не обязательно применяются одновременно: отдельные из них могут оказаться противоречивыми по отношению друг к другу, или их введение может быть необходимо из-за требований перевозчика (отправителя) в отдельно взятом случае.

Кроме того, выбор формы грузов сам по себе также можно рассматривать как отдельное ограничение. Очевидно, что подавляющее большинство задач и применяемых подходов рассматривает объекты в форме параллелепипеда, что обусловлено сложностью подбора математических моделей, позволяющих получить адекватное описание объектов сложной формы. Тем не менее в научной литературе описаны задачи, в которых используются объекты произвольной формы, например в форме цилиндра или шара. Так, в работе [9] предложена пространственная модель, основанная на частицах, для объектов в виде параллелепипеда и фигур сложной формы. В основе предлагаемой модели – описание объектов сложной формы в виде трехмерного облака точек с использованием переходов от разреженных и неравномерных облаков точек к выпуклым оболочкам с помощью метода грануляции. В работе [10] описана модель задачи упаковки крепежных изделий трех форм (шестигранных гаек, болтов с шестигранной головкой и цилиндров) в контейнеры переменной размерности. Предлагается иерархический подход, согласно которому исходная задача решается на трех уровнях:

- ◆ уровень 1 – упаковка крепежных изделий одинаковой формы в коробки малой размерности (задача трехмерной упаковки грузов сложной формы);
- ◆ уровень 2 – упаковка картонных коробок в ящики (задача упаковки в контейнеры);
- ◆ уровень 3 – упаковка ящиков в один контейнер (задача о рюкзаке).

Цель – максимизация плотности заполнения контейнера. При этом используются следующие ограничения:

- ◆ ограничение ориентации объектов;
- ◆ все объекты должны помещаться внутри контейнера;
- ◆ отсутствие пересечений (наложения) объектов друг с другом;
- ◆ возможность расположения объектов только параллельно граням контейнера;
- ◆ выполнение условия, что нижняя поверхность объектов, расположенных на втором и последующих уровнях относительно дна контейнера должна иметь в качестве опоры объекты предыдущего уровня.

Также возможно ограничение по соотношению площади соприкосновения объектов, расположенных на различных уровнях [10].

Обзор и анализ подходов к решению задачи трехмерной упаковки. Существует несколько различных классов алгоритмов, используемых для решения задачи трехмерной упаковки.

Точные алгоритмы. Точные алгоритмы выполняют поиск в области допустимых решений и гарантированно находят оптимальное решение. Применение точных алгоритмов ограничено задачами небольшой размерности, при небольшом количестве грузов и заданных ограничений, а также при наличии высокой однородности размещаемых грузов. В настоящее время точные алгоритмы практически не применяются для решения NP-сложных задач, но использование их в комбинации с эвристическими методами может повысить их скорость и качество получаемых решений [11, 12].

Традиционно, для решения задачи трехмерной упаковки, используются методы смешанно-целочисленного линейного программирования (MILP) [13].

Так, исследование [14] продемонстрировало, что возможно учитывать практические ограничения (например, ограничение по весу и ограничение динамической устойчивости) и применять данные методы для решения задач с размерностью до 100 объектов.

В работе [12] была представлена модель смешанно-целочисленного линейного программирования для решения задачи размещения в одном контейнере. Предложенная модель учитывает множество практических ограничений, в том числе, например, соблюдение приоритетов погрузки и вычисление центра масс. Данная модель позволяет получать оптимальные решения для задач размерностью до 16 грузов, что объясняется высокой вычислительной сложностью. В исследовании [11] был предложен точный алгоритм решения задачи о рюкзаке, основанный на методах линейного программирования. Данный алгоритм позволяет учитывать до 12 практических ограничений и успешно применяется для задач с размерностью до 110 элементов.

Эвристические алгоритмы. Высокая вычислительная сложность реальных практических задач трехмерной упаковки приводит к тому, что основным методом их решения являются эвристические алгоритмы. Основной идеей подавляющего большинства эвристических методов является стремление сократить или упростить пространство поиска решений, за счет оценки потенциально оптимальных и отсекация неоптимальных решений. Таким образом, одна большая исходная задача разбивается в процессе решения на множество подзадач, а получаемые результаты затем обобщаются с учетом специфических характеристик задачи. Общий подход, применяемый при решении с использованием эвристических методов, основан на конструктивном подходе, который состоит из трех этапов:

1. Генерация шаблона погрузки – определение структуры наборов грузов и допустимых объемов, в которые они могут быть размещены.

Различные шаблоны погрузки позволяют по-разному использовать объем упаковки (например, отдавая предпочтение определенной ориентации упаковки перед другими, ограничивая размещение определенными областями, обеспечивая наличие определенных точек соприкосновения между упакованными товарами). Общая идея заключается в создании из множества грузов неких укрупнённых структур (шаблонов, облегчающих упаковку) и последующей оптимизации размещения этих укрупнённых структур. Для генерации шаблона погрузки необходимо разработать стратегию размещения. Возможны следующие основные типы стратегии размещения:

♦ **построение слоев.** Построение слоев – это шаблон, при котором грузы размещаются горизонтально или вертикально (построение стен) слой за слоем. При такой схеме загрузки элементы сначала размещаются на дне контейнера для создания первого слоя. Последующие слои накладываются поверх предыдущего слоя или рядом с ним до тех пор, пока в контейнер больше не останется места. Существует множество различных стратегий построения слоев. Например, слои могут создаваться только с использованием однородных элементов [15]. В других случаях допускается включение в один слой неоднородных грузов разных типов. Или в процессе погрузки вначале создается несколько слоев однородных объектов, которые затем объединяются между собой по вертикали, в

результате чего получается общий гетерогенный слой [16]. Другая возможная стратегия состоит в создании нескольких слоев элементов одинаковой высоты и дальнейшем распределении их таким образом, чтобы высота слоя была близка к геометрическому центру контейнера [17]. В качестве критериев создания гетерогенных слоев в работе [18] было предложено использовать предельную величину нагрузки и вес грузов. В некоторых вариантах задачи трехмерной упаковки рассматриваются многокамерные контейнеры. В этом случае алгоритм построения слоев учитывает геометрические размеры всех камер. При этом, в случае, когда комбинируются изделия разных типов, плотность укладки слоев обычно считается показателем качества слоя [19].

♦ *построение блоков*. Стратегия построения блоков подразумевает формирование из грузов блоков в форме параллелепипедов и их последующую загрузку в контейнер. Блоки могут быть как однородными (состоящими из одинаковых грузов), так и неоднородными. Пример построения однородных блоков показан в работе [20], где для уменьшения размеров блоков и оптимизации их размещения в контейнере использован вероятностный метод. В работах [21, 22] приведен пример построения неоднородных блоков, при котором погрузка блоков в контейнер производится с помощью алгоритма поиска по дереву решений. В работе [23] используется похожая стратегия, согласно которой сначала формируются неоднородные блоки, которые затем располагаются с помощью алгоритма имитации отжига.

♦ *штабелирование*. В основе стратегии штабелирования лежит следующая идея: грузы сначала располагаются вертикальными стопками, а затем эти стопки размещаются на полу контейнера, что сводит рассматриваемую задачу к задаче двухмерной упаковки, которая имеет меньшую вычислительную сложность [24].

♦ *построение стены*. Стратегия построения стены описана в [25, 26], где используется модель классической задачи о рюкзаке и необходимо сформировать полосы коробок, а затем сгруппировать их в вертикальные слои.

2. Формирование последовательности упаковки – определение порядка расположения грузов, удовлетворяющих заданному набору ограничений.

3. Оценка решения и выполнение итераций – оценка качества получаемого решения и повторение алгоритма до тех пор, пока не будет достигнут заданный набор значений параметров качества.

Этапы формирования последовательности упаковки и оценки решения оказывают большое влияние на эффективность алгоритмов упаковки. При этом возможны два подхода к формированию последовательности упаковки: *статическая* и *динамическая* последовательность. Статическая последовательность задается в алгоритме до начала выполнения алгоритма и не изменяется в процессе упаковки [26]. Динамическая последовательность может быть скорректирована в процессе выполнения алгоритма в зависимости от результатов оценки значений параметров задачи [27].

Методы машинного обучения. Применение эвристических и гибридных методов позволяет улучшить качество получаемых результатов, а также исследовать большую часть пространства решений. Однако это достигается за счет увеличения вычислительной сложности алгоритмов. Потенциальной альтернативой эвристическим методам в настоящее время являются подходы, основанные на машинном обучении, которые в отдельных случаях позволяют получать лучшие результаты. Общая идея применения методов машинного обучения для решения задачи трехмерной упаковки состоит в использовании специальных классификаторов на основе известных примеров успешной предварительной упаковки контейнеров. Недостатком подходов, основанных на машинном обучении, является то, что они не могут гарантировать выполнение всех заданных ограничений для всех решений. Преимущество подходов на основе машинного обучения состоит в том, что решение может быть получено гораздо быстрее, чем это возможно при использовании точных или эвристических методов.

Пример подхода к решению задачи упаковки с использованием машинного обучения показан в работе [28], где для решения задачи трехмерного размещения грузов с минимизацией суммарного объема используется глубокое обучение с подкреплением в со-

четании с эвристическим методом. Метод глубокого обучения используется для определения и оптимизации порядка расположения грузов, а с помощью эвристического алгоритма определяется наилучшее положение и ориентация для каждого элемента. В работе [29] приведен модифицированный алгоритм из [28], улучшенный за счет использования многозадачного обучения (обучение с подкреплением и обучение с учителем) для одновременного определения порядка и ориентации объектов. В работе [30] предложена другая модель оптимизации, основанная на обучении с подкреплением, которая позволяет определить оптимальную последовательность выбора и ориентацию блоков. Данный подход предполагает наличие набора предварительно зафиксированных элементов и направлен на формирование последовательности упаковки. В исследовании [31] для оценки выбранной стратегии размещения использована глубокая нейронная сеть, а для улучшения этой стратегии используется алгоритм поиска по методу Монте-Карло.

Эволюционные и биоинспирированные алгоритмы решения задачи трехмерной упаковки. Согласно принятой в настоящее время точке зрения биоинспирированными (или вдохновленными природой) алгоритмами называется большая группа метаэвристических алгоритмов оптимизации, имитирующих процессы, протекающие как в живой, так и в неживой природе. Рассмотрим примеры биоинспирированных алгоритмов, используемые для решения задачи трехмерной упаковки.

Муравьиный алгоритм. Алгоритм муравьиной колонии основан на имитации поведения муравьиной колонии в процессе поиска пищи и взаимодействия между отдельными группами муравьев. В его основе лежит способность муравьев находить кратчайший путь и адаптироваться к изменяющимся внешним условиям. Муравьиная колония при этом рассматривается как многоагентная система, в которой особи (агенты) взаимодействуют между собой с помощью как прямого (химический и визуальный контакт), так и непрямого (стигмергия) обмена информацией. При движении каждый агент помечает свой путь феромоном (в алгоритме это некое числовое значение), что отражает приоритет выбора пути, который в дальнейшем используется другими агентами. После каждой итерации лучшие пути получают большее количество (значение) феромонов, а некоторая часть феромонов испаряется, что позволяет избежать преждевременной сходимости. К числу достоинств алгоритма относят устойчивость к попаданию в локальный оптимум, универсальность, масштабируемость и возможность организации параллельных вычислений. Недостатками являются необходимость прохождения большого количества итераций для сходимости и высокая зависимость от заданных настроек параметров алгоритма.

Метод роя частиц. В пространстве поиска решений случайным образом располагаются частицы, каждая из которых обладает координатой, скоростью и памятью. На каждой итерации вычисляется значение целевой функции для каждой частицы, а также определяется глобально лучшее решение, вдобавок, каждая частица запоминает лучшее из найденных ею решений. После этого производится обновление координат и скоростей частиц. На основе информации о текущем глобально лучшем решении производится корректировка скоростей и направлений движения частиц. Процесс поиска является итерационным и продолжается до достижения заданного максимального числа итераций или получения необходимой точности решения. Достоинствами метода роя частиц являются простота реализации, возможность организации параллельных вычислений, эффективность для решения задач большой размерности. Однако, данный алгоритм также требует тонкой настройки входных параметров, кроме того, его невозможно применять для решения задач с дискретной целевой функцией.

В работе [32] представлен подход к решению многокритериальной задачи двумерной упаковки с использованием нескольких ограничений. Для её решения предложен многокритериальный алгоритм роя частиц, который использует оптимальность по Парето для создания популяции решений. В процессе эволюции для увеличения разнообразия получаемых решений используется оператор мутации, что повышает гибкость алгоритма. Важным моментом является то, что в данном алгоритме направление движения частиц определяется на основе либо оценки пригодности каждого решения, либо глобальной оценки пригодности, но не с обеими оценками пригодности сразу.

Алгоритм искусственной пчелиной колонии. Алгоритм построен на имитации поведения пчел при поиске нектара и обмене информацией об источниках пищи с другими пчелами в улье. Источник пищи представляется в виде координаты в пространстве поиска и представляет собой вариант решения, а количество нектара интерпретируется как целевая функция задачи оптимизации. Для успешного применения данного алгоритма необходимо иметь возможность вычисления топологического расстояния между точками в пространстве решений. В начале работы алгоритма создается популяция пчёл, при этом пчёлы разделяются на два вида: пчелы-разведчики и пчелы-рабочие. Пчёлы-разведчики вылетают из улья в случайном направлении и производят поиск нектара на больших областях. Вернувшись в улей, они сообщают информацию о количестве обнаруженного нектара и его расположении. Затем производится оценка полученной информации на основе расчета функции пригодности и выделяются элитные участки. Затем в найденные элитные участки вылетают пчёлы-рабочие и добывают там нектар, а также исследуют их окрестности, а пчёлы-разведчики занимаются поиском новых участков. Если качество источника пищи не улучшается в течение заданного количества итераций, то текущий участок исключается из рассмотрения, а занятая работой на нем пчела-рабочий становится разведчиком. Таким образом, за счет комбинирования работы пчёл-разведчиков и пчел-рабочих, происходит итеративное улучшение текущей выборки решений. Количество пчёл разных видов, число и размеры участков, а также максимальное количество итераций являются настраиваемыми параметрами алгоритма. К преимуществам алгоритма можно отнести возможность быстрого обнаружения областей с высоким значением целевой функции, эффективность применения для решения многомерных задач и задач с непрерывной целевой функцией, а также возможность организации параллельных вычислений. В качестве недостатков алгоритма можно выделить высокую чувствительность к выбору значений управляющих параметров и отсутствие памяти, т.е. невозможность использования опыта предыдущих итераций для повышения скорости сходимости алгоритма.

Так, например, в работе [33] был предложен эвристический популяционный гибридный алгоритм, основанный на комбинации алгоритма искусственной пчелиной колонии и генетического алгоритма. В работе [34] описан алгоритм искусственной пчелиной колонии, включающий две эвристики загрузки для объединенной задачи трехмерной упаковки контейнеров и оптимизации маршрутов с учетом ограничений на грузоподъемность транспортных средств. При этом задача трехмерной упаковки решается с помощью применения различных эвристик, а для оптимизации маршрутов использовался алгоритм искусственной пчелиной колонии. Варианты решения в алгоритме искусственной пчелиной колонии формировались с помощью новой методики вещественного кодирования.

Светлячковый алгоритм. Метаэвристический алгоритм, вдохновленный движением роя светлячков к свету. Отдельный светлячок при этом интерпретируется как возможное решение. Яркость светлячка соответствует качеству решения, которое в случае с задачей упаковки определяется количеством использованных контейнеров или суммарным объемом. Алгоритм моделирует механизм движения менее ярких светлячков в направлении более ярких; этот процесс определяется функцией привлекательности, которая затухает по мере увеличения расстояния в пространстве решений. Такое поведение облегчает поиск глобального оптимума, позволяя оценивать различные последовательности упаковки и ориентацию предметов. Для задачи трехмерной упаковки решение обычно кодируется в виде последовательности перестановок грузов. Оператор движения адаптирован к этому комбинаторному процессу за счет использования таких операций, как сложные перестановки или последовательности мутации. Также используется детерминированная эвристика размещения (алгоритм максимального пространства или гильотинной резки) для преобразования последовательности перестановок в допустимый вариант упаковки и оценки качества полученного решения. Стохастическая и коллективная природа этого подхода повышает его способность избегать локальных экстремумов, что позволяет использовать его при решении задач трехмерной упаковки в ограниченном многомерном пространстве поиска (*highly constrained multi-modal landscape*), где необходимо учи-

тывать ограничения по весу, ориентации и стабильности. При этом можно отметить наличие серьезных трудностей при реализации из-за необходимости дополнительных вычислений при попарном сравнении светлячков, что приводит к квадратичной сложности алгоритма. Данная проблема может быть частично купирована за счет использования гибридизации светлячкового алгоритма с алгоритмами локального поиска (например, табу-поиск). К преимуществам алгоритма можно отнести высокую производительность при решении задач высокой размерности, возможность организации параллельных вычислений и гибкую структуру, которая позволяет легко комбинировать данный алгоритм с другими эвристическими методами. Однако, его эффективность в значительной степени зависит от качества предварительной настройки таких параметров, как коэффициент поглощения света и коэффициент рандомизации.

В работе [35] предложен комплексный подход к решению задачи трехмерной упаковки разнородных объектов в однотипные контейнеры. При упаковке рассматриваются объекты разных геометрических форм: куб, прямоугольная призма, цилиндр и сфера. Алгоритм позволяет учитывать различные практические ограничения, в том числе, ограничение стабильности, ограничения ориентации и несущей способности (load bearing) грузов. Моделирование показало, что проверка выполнения ограничений после размещения каждого объекта в контейнер дает лучшие результаты, чем проверка, выполняемая в конце процесса упаковки всего контейнера, хотя в этом случае увеличивается вычислительная сложность алгоритма. Также предусмотрена процедура предварительной сортировки и категоризации размещаемых объектов по их весу и размеру, что также позволяет улучшить качество получаемых результатов. После окончания работы светлячкового алгоритма, используется эвристика Best Fit для дозаполнения оставшегося свободного места в контейнере. Согласно приведенным данным, алгоритм позволяет в среднем на 22,84% улучшить использование объема контейнеров по сравнению с эвристикой Best Fit и генетическим алгоритмом. Тем не менее, при работе с фиксированным набором объектов для упаковки, иногда в контейнере остается незаполненное пространство (для заполнения которого затем опять-таки использовалась эвристика Best Fit), или, наоборот, некоторые объекты оставались неупакованными после окончания работы алгоритма (что приводит к удалению лучших решений). Кроме того, из-за ограничения стабильности, большинство объектов сферической формы исключалось из итогового варианта упаковки.

В работе [36] также рассмотрена задача трехмерной упаковки объектов сложной формы (в том числе, цилиндров и объектов сложной конфигурации). При этом размещаемые объекты представляются путем аппроксимации их формы с помощью вокселей. Для экономии времени для поиска подходящей позиции, была использована двухэтапная стратегия размещения. Сначала объект помещается в стартовую позицию, и затем выполняется попытка сдвига объекта во всех направлениях для того, чтобы он занял как можно более компактное положение. Таким образом решается задача минимизации имеющихся пустот. Гибридный светлячковый алгоритм используется для определения последовательности упаковки и ориентации объектов. Для инициализации популяции и настройки параметров алгоритма используется хаотический поиск, что позволяет увеличить разнообразие решений, а алгоритм имитации отжига применяется для выхода из локальных оптимумов. Полученные результаты показали, что предложенный гибридный светлячковый алгоритм выполняет упаковку объектов на 14-15% плотнее (по занимаемой высоте) по сравнению с упаковкой в случайном порядке. А по сравнению с классическими генетическим алгоритмом, алгоритмом роя частиц и светлячковым алгоритмом данный модифицированный алгоритм показал лучшие или сопоставимые по качеству результаты в 7 из 9 тестовых примеров. В случае разрешения на вращение объектов на 90 градусов полученные улучшения еще на 4%, но при этом увеличивается время вычислений.

Алгоритм стрекозы. Алгоритм стрекозы представляет собой метаэвристику, которая моделирует статическое и динамическое поведение роя стрекоз. Статический рой (небольшие группы, исследующие разные области) обеспечивает глобальное исследова-

ние пространства поиска, в то время как динамический рой (крупная группа, движущаяся согласованно) выполняет локальный поиск на перспективных участках. Механизм движения стрекоз похож на принцип, реализованный в алгоритме роя частиц, но дополнен сложными социальными правилами. В основе алгоритма лежат следующие пять управляющих социальных принципов:

- ◆ разделение – необходимость избегать столкновений с ближайшими соседями;
- ◆ согласование – поддержание одинаковой скорости с ближайшими особями;
- ◆ сплочённость – движение по направлению центра масс локальной группы;
- ◆ влечение к источнику пищи – движение в сторону источника пищи (лучшего найденного решения);
- ◆ стремление избегать врага – движение в противоположном направлении от наихудшего решения.

Перемещение каждой стрекозы в пространстве поиска определяется вектором, который определяется как взвешенная сумма этих пяти поведенческих факторов. В алгоритме происходит обновление текущей позиции каждого агента на основе использования вычисленного значения вектора шага. Для поддержания баланса между глобальным и локальным поиском, выполняется адаптивная настройка ключевых параметров, таких как инерционный вес и радиусы окрестности для определения соседей. В случае, если у стрекозы нет соседей в пределах её радиуса восприятия, то её позиция обновляется на основе принципа полёта Леви – особого типа случайного блуждания, усиливающего глобальный поиск. К основным преимуществам алгоритма стрекозы относятся его концептуальная простота, возможность поддержания баланса между глобальным и локальным поиском, а также эффективность при решении различных типов задач непрерывной и дискретной оптимизации. Главные недостатки – это высокая чувствительность к настройке параметров, а также вычислительная сложность особенно при решении задач большой размерности, что связано с необходимостью определения соседей для каждого агента на каждой итерации.

Алгоритм африканских буйволов. Алгоритм африканских буйволов – это метаэвристический алгоритм, имитирующий модель поведения, социальной организации и стратегию выживания стада африканских буйволов в их естественной среде обитания. При этом моделируется наблюдаемое у них сложное поведение, в том числе их коллективный образ жизни в больших группах, сезонные миграции, высоко скоординированные коллективные стратегии поиска источников пищи и воды, а также эффективного избегания хищников. В алгоритме реализован процесс обмена информацией между особями в естественной среде посредством звуковых сигналов. Под обучением, в данном случае, понимается процесс адаптации буйволов, основанный на их личном опыте и полученных коллективных знаниях. Каждый буйвол в популяции представляет собой потенциальное решение задачи оптимизации. Работа алгоритма начинается с инициализации стада (множество начальных решений), каждая особь которого занимает случайно выбранную позицию в пространстве поиска. На каждой итерации алгоритма каждый буйвол обновляет свое местоположение на основе сведений из следующих двух источников информации: своего личного лучшего опыта и глобально лучшего местоположения, обнаруженного всем стадом. При этом используются две стратегии поиска – совершенствование текущего лучшего решения ("оставаться и использовать") и исследование нового пространства решений ("двигаться и исследовать"), которые эффективно уравниваются за счет использования стохастических компонентов и специальных коэффициентов.

Отличительной особенностью алгоритма африканских буйволов является использование имитации звуков для управления стадом. Алгоритм итеративно совершенствует решения до тех пор, пока не будет выполнено заданное условие остановки. К преимуществам алгоритма можно отнести концептуальную простоту, небольшое количество настраиваемых параметров и эффективность в обеспечении баланса между исследованием и эксплуатацией (глобальным и локальным поиском), что позволяет применять его для решения различных дискретных и непрерывных задач. Однако алгоритм чувствителен к настройкам параметров, также существует проблема преждевременной сходимости в сложных условиях поиска с множественными экстремумами.

Известны примеры эффективного применения алгоритма африканских буйволов для NP-трудных задач, в частности, одномерных задач упаковки и раскроя. Так, в работе [37] алгоритм африканских буйволов использовался совместно с методом Ranked Order Value для конвертации решений из непрерывного пространства поиска в дискретные решения, необходимые для задачи упаковки. Для повышения эффективности алгоритма и учёта специфики решаемой задачи также используются различные эвристики упаковки (First Fit, Best Fit, First Fit Decreasing, Best Fit Decreasing). По сравнению с другими известными эвристическими алгоритмами (алгоритм муравьиной колонии, светлячковый алгоритм, квантовый и адаптивный алгоритмы кукушки) алгоритм африканских буйволов даёт лучшие результаты в 915 из 1210 тестовых задач одномерной упаковки [38].

В работе [39] был предложен эффективный подход к решению одномерной задачи раскроя на основе алгоритма африканских буйволов. Чтобы избежать преждевременной сходимости из-за частых перезапусков популяции, предложена модификация алгоритма с использованием процедуры «скрещивание и сохранение». В соответствии с данной процедурой, при перезапуске поиска выполняется скрещивание лучших из найденных на данный момент решений с целью создать потенциально более качественное решение. Это позволяет сохранять полученный в процессе поиска прогресс и целенаправленно улучшать решение перед тем, как алгоритм вернется к решению задачи исследования пространства. Предложенный подход позволяет получить лучшие по качеству решения по сравнению с некоторыми точными, эвристическими и метаэвристическими алгоритмами по критерию минимизации отходов в 60% тестовых задач, а по показателю использования материала в 50% тестовых задач.

Алгоритм серых волков. Алгоритм серых волков – это роевой метаэвристический алгоритм, основанный на модели социальной иерархии и коллективного охотничьего поведения стаи серых волков. Он моделирует иерархическую структуру стаи, разделяя популяцию существующих решений на четыре типа:

- ◆ Альфа (α) – наилучшее решение.
- ◆ Бета (β) – второе по качеству решение.
- ◆ Дельта (δ) – третье решение.
- ◆ Омега (ω) – все остальные решения, которые должны следовать за лидерами.

Основная процедура алгоритма воспроизводит коллективный процесс охоты, включающий поиск, окружение и атаку добычи. Управление процессом поиска и оптимизации выполняется на основе учета позиций, занимаемых тремя ведущими волками (α , β и δ). Остальные волки (Омега) на каждой итерации корректируют свои позиции на основе расчета местоположения лидеров, таким образом, вся популяция движется в направлении к наиболее перспективным областям пространства поиска. Ключевым элементом алгоритма серых волков является его адаптивный механизм поддержания баланса между глобальным (оператор «поиска») и локальным поиском (оператор «атаки»), на основе параметра, значение которого уменьшается с течением итераций. Такая стратегия способствует широкому исследованию пространства поиска на начальном этапе и смещению фокуса на локальное улучшение решений на поздних стадиях. Алгоритм характеризуется концептуальной простотой и относительно малым количеством настраиваемых параметров.

К достоинствам алгоритма серых волков относятся высокая скорость сходимости и эффективность в нахождении удовлетворительных решений для разнообразных задач оптимизации. Однако, при этом высока вероятность преждевременной сходимости к локальным оптимумам и риск снижения разнообразия популяции. Это может приводить к стагнации поиска и получению квазиоптимальных решений, особенно при решении сложных, многокритериальных задач. Для устранения выявленных недостатков существует большое количество модификаций алгоритма серых волков, в том числе, на основе гибридизации с другими алгоритмами, использования хаотических отображений и полета Леви для повышения эффективности поиска и устойчивости.

Так, в работе [40] представлен дискретный алгоритм серых волков для решения двухмерной задачи упаковки (упаковка в полосу). В решаемой задаче не учитывается гильотинное ограничение и невозможен поворот объектов. Поскольку стандартный алго-

ритм серых волков работает в непрерывном пространстве, для решения задачи упаковки, были модифицированы операторы поиска и атаки. Также, вместо евклидова расстояния, была использована новая метрика для расчета расстояния между двумя решениями (последовательностями упаковки). Была предложена улучшенная стратегия размещения на основе эвристики Best Fit. При этом после оценки решений возможны пять вариантов размещения прямоугольника в доступном пространстве, что позволяет более плотно заполнять область упаковки. Также для того, чтобы избежать преждевременного перехода к локальному поиску и застревания в локальных оптимумах был изменен критерий остановки фазы поиска: в модифицированной версии алгоритма поиск ведется всеми волками, что повышает разнообразие решений. Проведено тестирование алгоритма на 737 бенчмарках с классическими метаэвристиками и современными алгоритмами, специализирующимися на задаче упаковки. Представленный алгоритм превзошел результаты узкоспециализированных алгоритмов на наборах данных CX, 2SP и ZDF, а также показал конкурентоспособные результаты на остальных бенчмарках, но проиграл конкурентам при решении некоторых задач средней размерности. Несмотря на модификации, в алгоритме сохранено основное правило – следование за лидерами (альфа, бета, дельта), что может приводить к преждевременной сходимости. Кроме того, алгоритм характеризуется высокой вычислительной сложностью.

Генетический алгоритм. Генетический алгоритм является метаэвристическим методом поиска, вдохновлённым теорией эволюции и генетикой. Вначале формируется множество решений задачи (популяция), при этом каждое решение специальным образом кодируется в «хромосому» – последовательность генов (разрядов), каждый из которых является числом. Распространённым подходом, применяемым для моделирования решения задачи трехмерной упаковки, является кодирование в гене информации о порядковом номере груза, последовательности погрузки и возможности вращения груза, а обратное декодирование хромосомы в решение производится посредством применения детерминированного эвристического алгоритма в качестве стратегии размещения каждого объекта в пространстве контейнера [41, 42]. Генетический алгоритм состоит из трех операторов: селекция, кроссовер и мутация. Алгоритм выполняется итерационно и повторяется до тех пор, пока не будет выполнен критерий остановки, либо не пройдет заданное число итераций. В процессе эволюции моделируется естественный отбор, как инструмент устранения из популяции самых неудачных по качеству решений. Качество решений оценивается с помощью функции приспособленности. Операция кроссовера позволяет создавать новые решения и позволяет распространять признаки лучших решений среди популяции. Оператор мутации производит случайную замену разрядов решений. Благодаря этим трем операторам на каждом шаге алгоритма производится постепенное улучшение качества как отдельных решений, так и популяции всей в целом [43].

К отличительным чертам генетических алгоритмов можно отнести возможность выхода из локального экстремума за счет применения многонаправленного поиска, повышение эффективности генетических операторов за счет минимизации вычислительной сложности одной итерации алгоритма, использование только одной целевой функции (без необходимости вычисления вспомогательных функций) для оценки качества найденных решений [44]. Благодаря этому генетические алгоритмы эффективно применяются для решения различных видов задач, включая дискретные, непрерывные и недифференцируемые [45]. К недостаткам генетических алгоритмов можно отнести низкую эффективность при наличии вычислительно сложных функций приспособленности, что связано с многократным вычислением их значений для каждой особи на каждой итерации алгоритма. Также результаты алгоритма в значительной степени зависят от качества настройки управляющих параметров алгоритма.

В литературе имеется множество примеров успешного применения генетических алгоритмов для решения различных типов задач, в том числе задачи трехмерной упаковки. Так, в работе [42] предложен гибридный генетический алгоритм, позволяющий учитывать такие факторы, как стабильность и ориентация грузов, при решении оффлайн-задачи трехмерной упаковки. Для компенсации существующих недостатков в виде недостаточных возможностей локального поиска и небольшой скорости сходимости в роли

оператора мутации был использован алгоритм табу-поиска. В работе [46] использована «жадная» эвристика, согласно которой грузы сперва складываются в стопки, а затем их потенциальное расположение оценивается с использованием функции взвешенной оценки. Генетический алгоритм используется для автоматической настройки функции взвешенной оценки и улучшения качества упаковки. В работе [47] предложен адаптивный генетический алгоритм на основе эвристики построения стены, в котором вероятности выполнения операторов кроссовера и мутации адаптивно изменяются в зависимости от значения индивидуальной приспособленности.

В работе [48] генетический алгоритм был применен для минимизации количества палет, максимизации плотности упаковки грузов и минимизации гетерогенности грузов в каждой палете. Авторами предложена модель, позволяющая учитывать 8 различных ограничений, в том числе ограничения ориентации, устойчивости грузов, расположения грузов с учетом заказов и взаимного скрепления грузов.

В работе [49] был предложен подход, использующий изменение размера контейнера с целью повышения вместимости и возможности упаковки всех подготовленных грузов. При этом рассматривалась задача трехмерной упаковки на открытом пространстве, которая подразумевает отсутствие ограничений на размер разнородных контейнеров. Целью представленного алгоритма является минимизация суммарной стоимости за счет использования сконфигурированных контейнеров свободной формы и шаблонов погрузки. В данном случае была решена комбинированная задача трехмерной упаковки на открытом пространстве и многоконтейнерной упаковки на основе использования генетического алгоритма и внутренней детерминированной эвристики.

Обобщенная схема метаэвристического алгоритма для решения задачи трехмерной упаковки. Анализ особенностей выполнения и принципов работы различных метаэвристических и эволюционных алгоритмов позволил сформулировать обобщенный принцип работы метаэвристического алгоритма, отражающий последовательность действий при решении задачи трехмерной упаковки (рис. 1):

1. Инициализация алгоритма, ввод исходных данных. В состав исходных данных для решения задачи трехмерной упаковки должны включаться сведения об объектах, требующих упаковки: спецификация объектов, их габаритные размеры, важные особенности и возможные ограничения на размещение; сведения о имеющихся в наличии контейнерах: количество контейнеров, их габаритные размеры, особенности погрузки и возможные ограничения.

2. Создание и инициализация начального множества решений (популяции). В случае, если для решения задачи упаковки используется генетический алгоритм, то необходимо разработать методику кодирования / декодирования решений, составляющих начальную популяцию, т.е. каждая хромосома должна отображать определенную последовательность упаковки рассматриваемых объектов. В случае же, когда упаковка выполняется с использованием роевых методов необходимо определить (вычислить) начальные координаты всех членов коллектива (популяции решений).

3. Вычисление значений функции пригодности и оценка качества решений, составляющих текущее множество решений.

4. Выбор из текущей популяции решений заданного числа особей (хромосом). Выбор может производиться как случайным образом, так и на основе различных эвристик, например, на основе рулетки, по результатам «турнира», на основе ранжирования, «элитная» и т.д.

5. Выполнение предусмотренных алгоритмом операций и получение новых решений. В случае с генетическим алгоритмом – это могут быть генетические операторы: кроссинговер, мутация, инверсия и др. В биоинспирированных методах тип операторов для генерации новых решений зависит от используемого в данном решении метода. Также важным моментом, независимо от применяемого метода является необходимость поддержания баланса между поиском новых решений и исследованием существующих областей пространства поиска, между глобальным (операторы исследования) и локальным (операторы использования/эксплуатации) поиском.

6. Проверка выполнения заданного критерия останова. В случае невыполнения имеющегося критерия – возврат к п. 3 данного алгоритма для проведения оценки новых решений.

7. В случае, если заданный критерий выполнен – завершение работы алгоритма и сохранение наилучшего среди полученных вариантов решений.

8. Декодирование полученного решения к виду однозначного описания последовательности упаковки объектов в контейнер с помощью детерминированной эвристики.

Сравнительный анализ эффективности применения метаэвристических алгоритмов. Можно отметить проведенные исследования сравнительной эффективности различных метаэвристических алгоритмов при решении задачи о рюкзаке [50]. В качестве объекта исследования рассматривались генетический алгоритм, табу-поиск и модифицированный алгоритм искусственной пчелиной колонии на основе принципа Парето, а также три эвристических алгоритма. Исследование эффективности работы рассматриваемых алгоритмов проводилось на 15 типовых задачах, предложенных Loh и Nee в работе [51]. Средние арифметические показатели процента плотности размещения грузов в пространстве упаковки представлены на рис. 2.

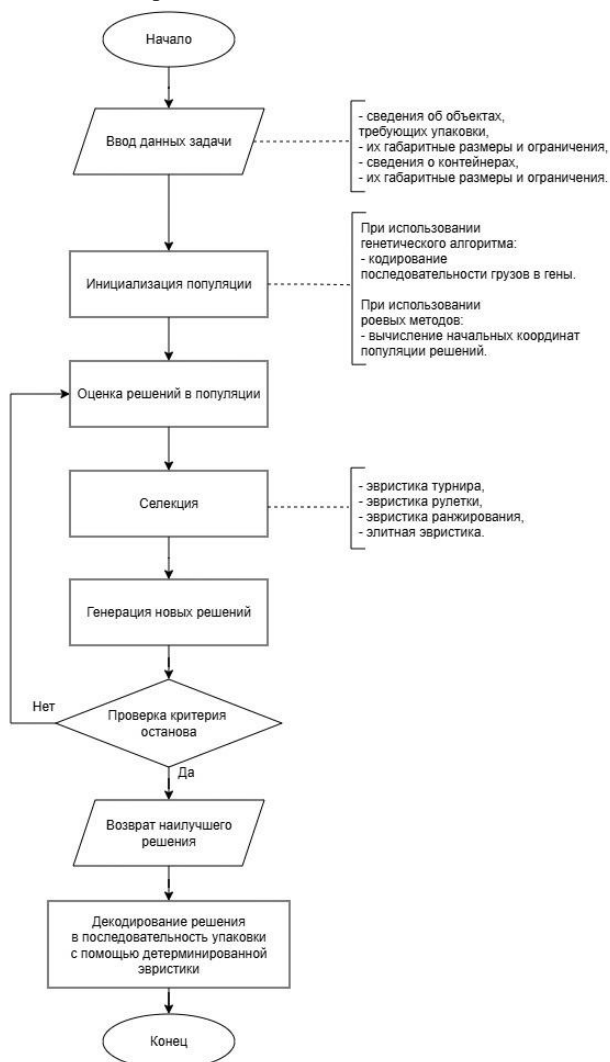


Рис. 1. Обобщенная схема метаэвристического алгоритма для решения задачи трехмерной упаковки

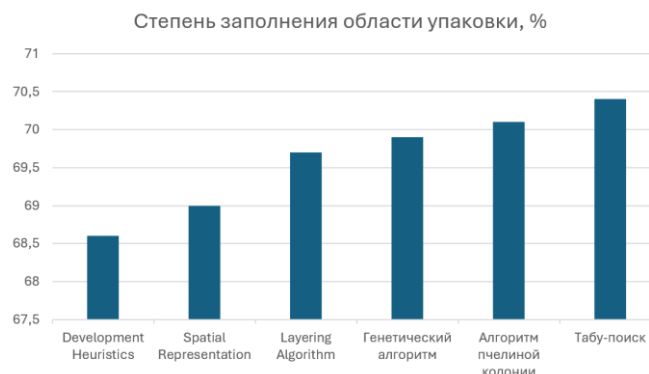


Рис. 2. Средние показатели результатов сравнительного анализа метаэвристических алгоритмов по результатам 15 тестов Loh и Nee

Полученные результаты тестов показывают, что метаэвристические алгоритмы в целом позволяют получить результаты, более близкие к оптимальным по количеству размещенных грузов и проценту заполнения области упаковки, чем эвристические методы. Это подтверждает, что для решения NP-трудной задачи трехмерной упаковки целесообразно разрабатывать методы, основанные на метаэвристических подходах.

Заключение. Таким образом, в ходе проведенного сравнительного анализа были рассмотрены основные типы алгоритмов решения задачи трехмерной упаковки, выделены их особенности, преимущества и недостатки. Проведенный анализ показал, что биоинспирированные методы, как правило, могут эффективно применяться для решения NP-сложных и NP-трудных задач, в частности задачи трехмерной упаковки, поскольку позволяют избегать ловушек локального экстремума и получать наборы квазиоптимальных решений за полиномиальное время. Кроме того, было показано, что методы машинного обучения также могут применяться для решения современных задач упаковки, поскольку позволяют получить допустимое решение за приемлемое время. В результате сравнительного анализа применения различных типов метаэвристических алгоритмов была разработана обобщенная схема применения данной категории методов для эффективного решения трехмерной упаковки, что позволит в дальнейшем определить наиболее перспективное направление решения поставленной задачи.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Liang K., Shan M., Liu H., et al. Intelligent optimization of e-commerce order packing using deep reinforcement learning with heuristic strategies // Applied Soft Computing. – 2025. – Issue 179.
2. Xiong H., Ding K., Ding W., et al. Towards reliable robot packing system based on deep reinforcement learning // Advanced Engineering Informatics. – 2023. – Issue 57.
3. Yang J., Liang K., Liu H., et al. Optimizing e-commerce warehousing through open dimension management in a three-dimensional bin packing system // PeerJ Computer Science. – 2023. – Issue 9. – URL: <https://peerj.com/articles/cs-1613/>.
4. Alpert Charles J., Mehta Dinesh P., Sapatnekar Sachin S. Handbook of algorithms for physical design automation. – CRC Press, New York, USA, 2009.
5. Гладков Л.А., Кравченко Ю.А., Курейчик В.В., Родзин С.И. Интеллектуальные системы: модели и методы метаэвристической оптимизации: монография. – Чебоксары: Среда, 2024.
6. Гладков Л.А., Курейчик В.В., Курейчик В.М. Генетические алгоритмы. – М.: Физматлит, 2025.
7. Карпенко А.П. Современные алгоритмы поисковой оптимизации. Алгоритмы, вдохновленные природой. – М.: МГТУ им. Баумана, 2014.
8. Гладков Л.А., Курейчик В.В., Курейчик В.М., Сороколетов П.В. Биоинспирированные методы в оптимизации. – М.: Физматлит, 2009.
9. Liu H., Zhou L., Yang J., et al. The 3D bin packing problem for multiple boxes and irregular items based on deep Q-network // Applied Intelligence. – 2023. – Issue 53. – P. 23398-23425.
10. Wu H., Leung S.C.H., Si Y. Three-stage heuristic algorithm for three-dimensional irregular packing problem // Applied Mathematical Modelling. – 2017. – Issue 41. – P. 431-444.

11. Nascimento O.X.D., de Queiroz T.A., Junqueira L., et al. Practical constraints in the container loading problem: Comprehensive formulations and exact algorithm // *Computers & Operations Research*. – 2021. – Issue 128.
12. Ocloo V.E. A new mathematical model for a 3D container packing problem. – Brandenburg: Brandenburg University of Technology, 2020.
13. Chen C.S., Lee S.M., Shen Q.S. An analytical model for the container loading problem // *European Journal of Operational Research*. – 1995. – Issue 80, No. 1. – P. 68-76.
14. Junqueira J., Morabito R., Yamashita D.S. Three-dimensional container loading models with cargo stability and load bearing constraints // *Computers & Operations Research*. – Issue 39, No. 1. – P. 74-85.
15. Ramos A.G., Silva E., Oliveira J.F. A new load balance methodology for container loading problem in road transportation // *European Journal of Operational Research*. – 2018. – Issue 266, No. 3. – P. 1140-1152.
16. Hasan J., Kaabi J., Harrath Y. Multi-objective 3D bin-packing problem // *Proceedings of the 8th International Conference on Modeling Simulation and Applied Optimization (ICMSAO)*. – 2019. – P. 1-5.
17. Costa M.D.G., Captivo M.E. Weight distribution in container loading: A case study // *International Transactions in Operational Research*. – 2016. – Issue 23. – P. 239-263.
18. Gzara F., Elhedhli S., Yildiz B.C. The pallet loading problem: Three-dimensional bin packing with practical constraints // *European Journal of Operational Research*. – 2020. – Issue 287, No. 3. – P. 1062-1074.
19. Júnior R.R., Yanasse H.H., Morabito R., et al. A hybrid approach for a multi-compartment container loading problem // *Expert Systems with Applications*. – 2019. – Issue 137. – P. 471-492.
20. Araújo O.C.B.D., and V.A. Armentano. A multi-start random constructive heuristic for the container loading problem // *Pesquisa Operacional*. – 2007. – Issue 27, No. 2. – P. 311-331.
21. Zhu W., Lim A. A new iterative-doubling Greedy-Lookahead algorithm for the single container loading problem // *European Journal of Operational Research*. – 2012. – Issue 222, No. 3. – P. 408-417.
22. Fanslau T., Bortfeldt A. A tree search algorithm for solving the container loading problem // *INFORMS Journal on Computing*. – 2009. – Issue 22, No. 2. – P. 222-235.
23. Da Silva E.F., Leao A.A.S., Toledo F.M.B., et al. A matheuristic framework for the three-dimensional single large object placement problem with practical constraints // *Computers & Operations Research*. – 2020. – Issue 124.
24. Toffolo T.A.M., Esprit E., Wauters T. A two-dimensional heuristic decomposition approach to a three-dimensional multiple container loading problem // *European Journal of Operational Research*. – 2017. – Issue 257, No. 2. – P. 526-538.
25. Liu S., Tan W., Xu Z., et al. A tree search algorithm for the container loading problem // *Computers & Industrial Engineering*. – 2014. – Issue 75. – P. 20-30.
26. Chien C.-F., Deng J.-F. A container packing support system for determining and visualizing container packing patterns // *Decision Support Systems*. – 2004. – Issue 37, No. 1. – P. 23-24.
27. Zhao H., She Q., Zhu C. Online 3D bin packing with constrained deep reinforcement learning // *Proceedings of AAAI Technical Track on Application Domains*. – Issue 35, No. 1. – P. 741-749.
28. Hu H., Zhang X., Yan X. et al. Solving a new 3D bin packing problem with deep reinforcement learning method // *arXiv*. – 2017. – arXiv:1708.05930
29. Duan L., Hu H., Qian Y. et al. A multi-task selected learning approach for solving 3D flexible bin packing problem // *arXiv*. – 2018. – arXiv:1804.06896.
30. Hu R., Xu J., Chen B. TAP-Net: Transport-and-pack using reinforcement learning // *ACM Transactions on Graphics (TOG)*. – 2020. – Issue 39, No. 6. – P. 1-15.
31. Laterre A., Fu Y., Jabri M.K., et al. Ranked reward: Enabling self-play reinforcement learning for combinatorial optimization // *arXiv*. – 2018. – arXiv:1807.01672.
32. Liu D. On solving multiobjective bin packing problems using evolutionary particle swarm optimization // *European Journal of Operational Research*. – 2008. – Issue 190, No. 2. – P. 357-382.
33. Олейник А.С., Полупанова Е.Е. Решение задачи трёхмерной упаковки алгоритмом последовательной гибридизации // *Прикладная математика: современные проблемы математики, информатики и моделирования: Матер. III всероссийской научно-практической конференции молодых ученых, Краснодар, 21–24 апреля 2021 года. Т. I*. – 2021. – С. 183-188.
34. Wu B., Lin Jg., Dong M. Artificial Bee Colony Algorithm for Three-Dimensional Loading Capacitated Vehicle Routing Problem // *Proceedings of 20th International Conference on Industrial Engineering and Engineering Management*. – 2013. – P. 815-825.
35. Rajesh Kanna S.K., Udaiyakumar K.C. A Complete Framework For Multi-Constrained 3 D Bin Packing Optimization Using Firefly Algorithm // *International Journal of Pure and Applied Mathematics*. – 2017. – Issue 114. – P. 267-282. – URL: <https://www.semanticscholar.org/paper/A-Complete-Framework-For-Multi-Constrained-3-D-Bin-Kanna-Udaiyakumar/badaf2002a67581a1574a2e82e55deb65a611a6f>.

36. Zhao Ch., Jliang L., Teo K. A hybrid chaos firefly algorithm for three-dimensional irregular packing problem // *Journal of Industrial and Management Optimization*. – 2020. – Issue 13. – P. 1-21. – URL: https://www.researchgate.net/publication/328761857_A_hybrid_chaos_firefly_algorithm_for_three-dimensional_irregular_packing_problem.
37. Gherboudj A. African Buffalo Optimization for One Dimensional Bin Packing Problem // *International Journal of Swarm Intelligence Research*. – 2019. – Issue 10, No. 4. – P. 38-52. – URL: <https://www.igi-global.com/gateway/article/234698>.
38. Layeb A., Benayad Z. A novel firefly algorithm based ant colony optimization for solving combinatorial optimization problems // *International Journal of Computer Science and Applications*. – 2014. – Issue 11. – P. 19-37. – URL: https://www.researchgate.net/publication/268215316_A_novel_firefly_algorithm_based_ant_colony_optimization_for_solving_combinatorial_optimization_problems.
39. Montiel-Arrieta L.J., Barragan-Vite I., Seck-Tuoh-Mora J.C., et al. Minimizing the total waste in the one-dimensional cutting stock problem with the African buffalo optimization algorithm // *PeerJ Computer Science*. – 2023. – Issue 9. – DOI: 10.7717/peerj-cs.1728. – URL: <https://pubmed.ncbi.nlm.nih.gov/38192486/>.
40. Wang P., Rao Y., Luo Q. An Effective Discrete Grey Wolf Optimization Algorithm for Solving the Packing Problem // *IEEE Access*. – 2020. – DOI:10.1109/ACCESS.2020.3004380. – URL: https://www.researchgate.net/publication/342402646_An_Effective_Discrete_Grey_Wolf_Optimization_Algorithm_for_Solving_the_Packing_Problem.
41. Kang K., Moon I., Wang H. A hybrid genetic algorithm with a new packing strategy for the three-dimensional bin packing problem // *Applied Mathematics and Computation*. – 2012. – No. 219. – P. 1287-1299.
42. Wang X., Gong L., Zhao H., Li B., Tian M. A 3D Offline Packing Algorithm considering Cargo Orientation and Stability // *International Journal of Distributed Sensor Networks*. – 2023.
43. Arora J.S. Introduction to Optimum Design. – 3 ed. – Iowa City, 2012. – URL: [https://mu.menofia.edu.eg/PrtlFiles/Staff/1045/Portal/Files/Introduction%20to%20Optimum%20Design,%203rd%20Ed\(3\).pdf](https://mu.menofia.edu.eg/PrtlFiles/Staff/1045/Portal/Files/Introduction%20to%20Optimum%20Design,%203rd%20Ed(3).pdf).
44. Zhang D.-F., Chen Sh.-D., Liu Y.-J. An improved heuristic recursive strategy based on genetic algorithm for the strip rectangular packing problem // *Acta Automatica Sinica*. – 2007. – No. 33.
45. Goncalves J.F., Resende M.G.C. A biased random key genetic algorithm for 2D and 3D bin packing problems // *International Journal of Production Economics*. – 2013. – Issue 145, No. 2. – P. 500-510.
46. Olsson J., Larsson T., Quttineh N.-H. Automating the planning of container loading for atlas copco: Coping with real-life stacking and stability constraints // *European Journal of Operational Research*. – Issue 280, No. 3. – P. 1018-1034.
47. Xiang X. Yu C., Xu H., et al. Optimization of heterogeneous container loading problem with adaptive genetic algorithm // *Complexity*. – 2018. – Issue 6. – P. 1-12.
48. Ananno A.A., Ribeiro L. A Multi-Heuristic Algorithm for Multi-Container 3-D Bin Packing Problem Optimization Using Real World Constraints // *IEEE Access*. – 2024. – Issue 12. – P. 42105-42130.
49. Yang Y., Wu Z., Hao X., et al. Two-layer heuristic for the three-dimensional bin design and packing problem // *Engineering Optimization*. – 2023. – DOI: 10.1080/0305215X.2023.2269868.
50. Phongmoo S., Leksakul K., Charoenchai N., et al. Artificial Bee Colony Algorithm with Pareto-Based Approach for Multi-Objective Three-Dimensional Single Container Loading Problems / S. Phongmoo, // *Applied Sciences*. – 2023. – Issue 13, No. 11. – DOI:10.3390/app13116601. – URL: <https://www.mdpi.com/2076-3417/13/11/6601>.
51. Loh T., Nee A. A Packing Algorithm for Hexahedral Boxes // *Proceedings of the Conference of Industrial Automation*. – 1992. – P. 115-126.

REFERENCES

1. Liang K., Shan M., Liu H., et al. Intelligent optimization of e-commerce order packing using deep reinforcement learning with heuristic strategies, *Applied Soft Computing*, 2025, Issue 179.
2. Xiong H., Ding K., Ding W., et al. Towards reliable robot packing system based on deep reinforcement learning, *Advanced Engineering Informatics*, 2023, Issue 57.
3. Yang J., Liang K., Liu H., et al. Optimizing e-commerce warehousing through open dimension management in a three-dimensional bin packing system, *PeerJ Computer Science*, 2023, Issue 9. Available at: <https://peerj.com/articles/cs-1613/>.
4. Alpert Charles J., Mehta Dinesh P., Sapatnekar Sachin S. Handbook of algorithms for physical design automation. CRC Press, New York, USA, 2009.
5. Gladkov L.A., Kravchenko Yu.A., Kureychik V.V., Rodzin S.I. Интеллектуальные системы: модели и методы метаэвристической оптимизации: монография [Intelligent systems: models and methods of metaheuristic optimization: monograph]. Cheboksary: Sreda, 2024.

6. Gladkov L.A., Kureychik V.V., Kureychik V.M. Geneticheskie algoritmy [Genetic algorithms]. Moscow: Fizmatlit, 2025.
7. Karpenko A.P. Sovremennye algoritmy poiskovoy optimizatsii. Algoritmy, vdokhnovlennyye prirodoy [Modern search engine optimization algorithms. algorithms inspired by nature]. Moscow: MGTU im. Bauman, 2014.
8. Gladkov L.A., Kureychik V.V., Kureychik V.M., Sorokoletov P.V. Bioinspirirovannyye metody v optimizatsii [Bioinspired methods in optimization]. Moscow: Fizmatlit, 2009.
9. Liu H., Zhou L., Yang J., et al. The 3D bin packing problem for multiple boxes and irregular items based on deep Q-network, *Applied Intelligence*, 2023, Issue 53, pp. 23398-23425.
10. Wu H., Leung S.C.H., Si Y. Three-stage heuristic algorithm for three-dimensional irregular packing problem, *Applied Mathematical Modelling*, 2017, Issue 41, pp. 431-444.
11. Nascimento O.X.D., de Queiroz T.A., Junqueira L., et al. Practical constraints in the container loading problem: Comprehensive formulations and exact algorithm, *Computers & Operations Research*, 2021, Issue 128.
12. Ocloo V.E. A new mathematical model for a 3D container packing problem. Brandenburg: Brandenburg University of Technology, 2020.
13. Chen C.S., Lee S.M., Shen Q.S. An analytical model for the container loading problem, *European Journal of Operational Research*, 1995, Issue 80, No. 1, pp. 68-76.
14. Junqueira J., Morabito R., Yamashita D.S. Three-dimensional container loading models with cargo stability and load bearing constraints, *Computers & Operations Research*, Issue 39, No. 1, pp. 74-85.
15. Ramos A.G., Silva E., Oliveira J.F. A new load balance methodology for container loading problem in road transportation, *European Journal of Operational Research*, 2018, Issue 266, No. 3, pp. 1140-1152.
16. Hasan J., Kaabi J., Harrath Y. Multi-objective 3D bin-packing problem, *Proceedings of the 8th International Conference on Modeling Simulation and Applied Optimization (ICMSAO)*, 2019, pp. 1-5.
17. Costa M.D.G., Captivo M.E. Weight distribution in container loading: A case study, *International Transactions in Operational Research*, 2016, Issue 23, pp. 239-263.
18. Gzara F., Elhedhli S., Yildiz B.C. The pallet loading problem: Three-dimensional bin packing with practical constraints, *European Journal of Operational Research*, 2020, Issue 287, No. 3, pp. 1062-1074.
19. Júnior R.R., Yanasse H.H., Morabito R., et al. A hybrid approach for a multi-compartment container loading problem, *Expert Systems with Applications*, 2019, Issue 137, pp. 471-492.
20. Araújo O.C.B.D., and V.A. Armentano. A multi-start random constructive heuristic for the container loading problem, *Pesquisa Operacional*, 2007, Issue 27, No. 2, pp. 311-331.
21. Zhu W., Lim A. A new iterative-doubling Greedy-Lookahead algorithm for the single container loading problem, *European Journal of Operational Research*, 2012, Issue 222, No. 3, pp. 408-417.
22. Fanslau T., Bortfeldt A. A tree search algorithm for solving the container loading problem, *INFORMS Journal on Computing*, 2009, Issue 22, No. 2, pp. 222-235.
23. Da Silva E.F., Leao A.A.S., Toledo F.M.B., et al. A matheuristic framework for the three-dimensional single large object placement problem with practical constraints, *Computers & Operations Research*, 2020, Issue 124.
24. Toffolo T.A.M., Esprit E., Wauters T. A two-dimensional heuristic decomposition approach to a three-dimensional multiple container loading problem, *European Journal of Operational Research*, 2017, Issue 257, No. 2, pp. 526-538.
25. Liu S., Tan W., Xu Z., et al. A tree search algorithm for the container loading problem, *Computers & Industrial Engineering*, 2014, Issue 75, pp. 20-30.
26. Chien C.-F., Deng J.-F. A container packing support system for determining and visualizing container packing patterns, *Decision Support Systems*, 2004, Issue 37, No. 1, pp. 23-24.
27. Zhao H., She Q., Zhu C. Online 3D bin packing with constrained deep reinforcement learning, *Proceedings of AAAI Technical Track on Application Domains*, Issue 35, No. 1, pp. 741-749.
28. Hu H., Zhang X., Yan X. et al. Solving a new 3D bin packing problem with deep reinforcement learning method, *arXiv*, 2017. arXiv:1708.05930.
29. Duan L., Hu H., Qian Y. et al. A multi-task selected learning approach for solving 3D flexible bin packing problem, *arXiv*, 2018. arXiv:1804.06896.
30. Hu R., Xu J., Chen B. TAP-Net: Transport-and-pack using reinforcement learning, *ACM Transactions on Graphics (TOG)*, 2020, Issue 39, No. 6, pp. 1-15.
31. Laterre A., Fu Y., Jabri M.K., et al. Ranked reward: Enabling self-play reinforcement learning for combinatorial optimization, *arXiv*, 2018. arXiv:1807.01672.
32. Liu D. On solving multiobjective bin packing problems using evolutionary particle swarm optimization, *European Journal of Operational Research*, 2008, Issue 190, No. 2, pp. 357-382.

33. Oleynik A.S., Polupanova E.E. Reshenie zadachi trekhmernoy upakovki algoritmom posledovatel'noy gibridizatsii [Solution of the three-dimensional packing problem using the sequential hybridization algorithm], *Prikladnaya matematika: sovremennye problemy matematiki, informatiki i modelirovaniya: Mater. III vserossiyskoy nauchno-prakticheskoy konferentsii molodykh uchenykh, Krasnodar, 21–24 aprelya 2021 goda* [Applied Mathematics: Modern Problems of Mathematics, Computer Science and Modeling: Proceedings of the III All-Russian Scientific and Practical Conference of Young Scientists, Krasnodar, April 21–24, 2021]. Vol. I, 2021, pp. 183-188.
34. Wu B., Lin Jg., Dong M. Artificial Bee Colony Algorithm for Three-Dimensional Loading Capacitated Vehicle Routing Problem, *Proceedings of 20th International Conference on Industrial Engineering and Engineering Management*, 2013, pp. 815-825.
35. Rajesh Kanna S.K., Udaiyakumar K.C. A Complete Framework For Multi-Constrained 3 D Bin Packing Optimization Using Firefly Algorithm, *International Journal of Pure and Applied Mathematics*, 2017, Issue 114, pp. 267-282. Available at: <https://www.semanticscholar.org/paper/A-Complete-Framework-For-Multi-Constrained-3-D-Bin-Kanna-Udaiyakumar/badaf2002a67581a1574a2e82e55deb65a611a6f>.
36. Zhao Ch., Jjiang L., Teo K. A hybrid chaos firefly algorithm for three-dimensional irregular packing problem, *Journal of Industrial and Management Optimization*, 2020, Issue 13, pp. 1-21. Available at: https://www.researchgate.net/publication/328761857_A_hybrid_chaos_firefly_algorithm_for_three-dimensional_irregular_packing_problem.
37. Gherboudj A. African Buffalo Optimization for One Dimensional Bin Packing Problem, *International Journal of Swarm Intelligence Research*, 2019, Issue 10, No. 4, pp. 38-52. Available at: <https://www.igi-global.com/gateway/article/234698>.
38. Layeb A., Benayad Z. A novel firefly algorithm based ant colony optimization for solving combinatorial optimization problems, *International Journal of Computer Science and Applications*, 2014, Issue 11, pp. 19-37. Available at: https://www.researchgate.net/publication/268215316_A_novel_firefly_algorithm_based_ant_colony_optimization_for_solving_combinatorial_optimization_problems.
39. Montiel-Arrieta L.J., Barragan-Vite I., Seck-Tuoh-Mora J.C., et al. Minimizing the total waste in the one-dimensional cutting stock problem with the African buffalo optimization algorithm // *PeerJ Computer Science*. – 2023. – Вып. 9. – DOI:10.7717/peerj-cs.1728. Available at: <https://pubmed.ncbi.nlm.nih.gov/38192486/>.
40. Wang P., Rao Y., Luo Q. An Effective Discrete Grey Wolf Optimization Algorithm for Solving the Packing Problem, *IEEE Access*, 2020. DOI:10.1109/ACCESS.2020.3004380. Available at: https://www.researchgate.net/publication/342402646_An_Effective_Discrete_Grey_Wolf_Optimization_Algorithm_for_Solving_the_Packing_Problem.
41. Kang K., Moon I., Wang H. A hybrid genetic algorithm with a new packing strategy for the three-dimensional bin packing problem, *Applied Mathematics and Computation*, 2012, No. 219, pp. 1287-1299.
42. Wang X., Gong L., Zhao H., Li B., Tian M. A 3D Offline Packing Algorithm considering Cargo Orientation and Stability, *International Journal of Distributed Sensor Networks*, 2023.
43. Arora J.S. Introduction to Optimum Design. 3 ed. Iowa City, 2012. Available at: [https://mu.menofia.edu.eg/PrtlFiles/Staff/1045/Portal/Files/Introduction%20to%20Optimum%20Design,%203rd%20Ed\(3\).pdf](https://mu.menofia.edu.eg/PrtlFiles/Staff/1045/Portal/Files/Introduction%20to%20Optimum%20Design,%203rd%20Ed(3).pdf).
44. Zhang D.-F., Chen Sh.-D., Liu Y.-J. An improved heuristic recursive strategy based on genetic algorithm for the strip rectangular packing problem, *Acta Automatica Sinica*, 2007, No. 33.
45. Goncalves J.F., Resende M.G.C. A biased random key genetic algorithm for 2D and 3D bin packing problems, *International Journal of Production Economics*, 2013, Issue 145, No. 2, pp. 500-510.
46. Olsson J., Larsson T., Quttineh N.-H. Automating the planning of container loading for atlas copco: Coping with real-life stacking and stability constraints, *European Journal of Operational Research*, Issue 280, No. 3, pp. 1018-1034.
47. Xiang X. Yu C., Xu H., et al. Optimization of heterogeneous container loading problem with adaptive genetic algorithm, *Complexity*, 2018, Issue 6, pp. 1-12.
48. Ananno A.A., Ribeiro L. A Multi-Heuristic Algorithm for Multi-Container 3-D Bin Packing Problem Optimization Using Real World Constraints, *IEEE Access*, 2024, Issue 12, pp. 42105-42130.
49. Yang Y., Wu Z., Hao X., et al. Two-layer heuristic for the three-dimensional bin design and packing problem, *Engineering Optimization*, 2023. DOI: 10.1080/0305215X.2023.2269868.
50. Phongmoo S., Leksakul K., Charoenchai N., et al. Artificial Bee Colony Algorithm with Pareto-Based Approach for Multi-Objective Three-Dimensional Single Container Loading Problems, *Applied Sciences*, 2023, Issue 13, No. 11. DOI: 10.3390/app13116601. Available at: <https://www.mdpi.com/2076-3417/13/11/6601>.
51. Loh T., Nee A. A Packing Algorithm for Hexahedral Boxes, *Proceedings of the Conference of Industrial Automation*, 1992, pp. 115-126.

Сорокин Михаил Максимович – Южный федеральный университет; e-mail: sorokin@sfedu.ru; г. Таганрог, Россия; кафедра систем автоматизированного проектирования им. Виктора Михайловича Курейчика; аспирант.

Гладков Леонид Анатольевич – Южный федеральный университет; e-mail: lagladkov@sfedu.ru; г. Таганрог, Россия; кафедра систем автоматизированного проектирования им. Виктора Михайловича Курейчика; профессор.

Гладкова Надежда Викторовна – Южный федеральный университет; e-mail: nvgladkova@sfedu.ru; г. Таганрог, Россия; кафедра систем автоматизированного проектирования им. Виктора Михайловича Курейчика; старший преподаватель.

Sorokin Mikhail Maximovich – Southern Federal University; e-mail: sorokin@sfedu.ru; Taganrog, Russia; the Department of Computer Aided Design named after V.M. Kureichik; postgraduate student.

Gladkov Leonid Anatol'evich – Southern Federal University; e-mail: lagladkov@sfedu.ru; Taganrog, Russia; the Department of Computer Aided Design named after V.M. Kureichik; professor.

Gladkova Nadezhda Viktorovna – Southern Federal University; e-mail: nvgladkova@sfedu.ru; Taganrog, Russia; the Department of Computer Aided Design named after V.M. Kureichik; senior teacher.

УДК 004.056.55

DOI 10.18522/2311-3103-2026-4-150-159

В.С. Стародубцев, Л.К. Бабенко**АЛГОРИТМЫ СНИЖЕНИЯ ВРЕМЕННЫХ ЗАТРАТ НА ВЫПОЛНЕНИЕ ОПЕРАЦИЙ КРИПТОСИСТЕМЫ ДОМИНГО-ФЕРРЕРА**

Проведён анализ литературы по тематике полностью гомоморфного шифрования. Представлено краткое описание полностью гомоморфной криптографической системы Доминго-Феррера, основанной на задаче факторизации чисел, приведены временные характеристики этапов атаки с известным открытым текстом на данную криптосистему. Проанализированы временные характеристики операций криптосистемы, описаны методы и средства их практической реализации. Предложены новые алгоритмы реализации операций криптосистемы Доминго-Феррера, позволяющие снизить время их выполнения. Обоснование оценок временных затрат операций криптосистемы формируется на основе теоретических расчётов, а также результатов экспериментальных исследований. Целью исследования является снижение времени выполнения операций криптосистемы Доминго-Феррера посредством разработки алгоритмов их модификации, учитывающих особенности практической реализации. Основным результатом настоящей работы является снижение времени выполнения следующих операций криптосистемы Доминго-Феррера: шифрования – на 10-15%, расшифрования – в 2 раза, гомоморфного умножения – в 64 раза для цепочки из 200 умножений при использовании степени полиномов представления шифротекста $d=100$ и незначительном увеличении временных затрат на генерацию ключа. Проведённое исследование представляет собой значимый вклад в развитие полностью гомоморфной криптосистемы Доминго-Феррера, основанной на задаче факторизации целых чисел. Настоящая работа обладает практической значимостью, поскольку позволяет значительно повысить производительность гомоморфных вычислений данной криптосистемы. Полученные результаты могут стать основой для разработки эффективных (с точки зрения вычислительных затрат и обеспечения уровня безопасности) программно-аппаратных комплексов облачных вычислений, использующих полностью гомоморфную криптосистему Доминго-Феррера для обеспечения конфиденциальности обрабатываемой информации.

Информационная безопасность; гомоморфное шифрование; гомоморфная схема шифрования; полностью гомоморфное шифрование; криптосистема Доминго-Феррера; криптоанализ.

V.S. Starodubcev, L.K. Babenko**ALGORITHMS FOR REDUCING THE TIME REQUIRED TO PERFORM OPERATIONS OF THE DOMINGO-FERRER CRYPTOSYSTEM**

An analysis of the literature on the topic of fully homomorphic encryption is carried out. A brief description of the completely homomorphic Domingo-Ferrer cryptographic system based on the number factorization problem is presented, and the time characteristics of the stages of an attack with a known plaintext on this cryptosystem are given. The time characteristics of cryptosystem operations are analyzed,

methods and means of their practical implementation are described. New algorithms for implementing the operations of the Domingo-Ferrer cryptosystem are proposed to reduce their execution time. The justification of estimates of the time costs of cryptosystem operations is formed on the basis of theoretical calculations, as well as the results of experimental studies. The aim of the study is to reduce the execution time of the Domingo-Ferrer cryptosystem by developing algorithms for their modification, taking into account the specifics of practical implementation. The main result of this work is a reduction in the execution time of the following operations of the Domingo-Ferrer cryptosystem: encryption by 10-15%, decryption by 2 times, homomorphic multiplication by 64 times for a chain of 200 multiplications using the degree of polynomials of the ciphertext representation $d=100$ and a slight increase in the time spent on key generation. The conducted research represents a significant contribution to the development of a fully homomorphic Domingo-Ferrer cryptosystem based on the integer factorization problem. This work has practical significance because it significantly improves the performance of homomorphic calculations of this cryptosystem. The results obtained can become the basis for the development of efficient (in terms of required computing costs and the level of security provided) cloud computing software and hardware systems using a fully homomorphic Domingo-Ferrer cryptosystem to ensure the confidentiality of processed information.

Information security; homomorphic encryption; homomorphic encryption scheme; fully homomorphic encryption; Domingo-Ferrer cryptosystem; cryptanalysis.

Введение. В настоящее время гомоморфное шифрование является одним из наиболее перспективных направлений благодаря своей возможности обеспечивать конфиденциальность при обработке информации с помощью облачных вычислений [1, 2]. Постоянный рост объемов обрабатываемой информации вынуждает компании постоянно наращивать свои вычислительные мощности. Однако в большинстве случаев развертывание собственной инфраструктуры связано со значительными денежными затратами, поэтому наиболее часто компании не приобретают, а вместо этого арендуют необходимое оборудование у провайдеров облачных вычислений. Это позволяет значительно сократить как денежные, так и временные затраты на развертывание необходимого окружения, поскольку технология облачных вычислений обеспечивает высокую гибкость и масштабируемость используемых решений [3].

Однако, несмотря на все вышеобозначенные преимущества, область применения облачных вычислений существенно ограничена. Подобная технология обладает существенным недостатком – в большинстве случаев облачная инфраструктура представляет собой среду, в которой не гарантируется конфиденциальность обрабатываемых сведений [4]. Следовательно, компании, обрабатывающие информацию ограниченного доступа, вынуждены отказываться от использования облачных вычислений. Одним из решений данной проблемы является использование гомоморфных операций, которые обрабатывают информацию в зашифрованном виде.

Впервые идея выполнения гомоморфных вычислений была предложена Ривестом и др. в работе [5], когда выяснилось, что числа, зашифрованные с помощью асимметричного шифра RSA [6], возможно перемножать в зашифрованном виде. Результат вычислений над шифртекстами был эквивалентен результату аналогичных операций, выполненных над открытыми текстами. Это термин авторы обозначили «Privacy Homomorphism», а их работа легла в основу концепции гомоморфного шифрования. Вскоре стало ясно, что для формирования полноценной системы гомоморфных вычислений требуется разработать полностью гомоморфную криптосистему, поддерживающую выполнение и операций сложения, и умножения. Впоследствии различными авторами было предложено множество схем, однако большая часть из них оказалась либо небезопасной, либо, подобно RSA, лишь частично гомоморфной [4] (поддерживала выполнение только одной гомоморфной операции – сложения или умножения).

Значительный прорыв в области гомоморфного шифрования совершил Крейг Джендри, когда в 2009 году в своей работе [7] предложил новую криптографическую систему. Его криптосистема оказалась первой схемой полностью гомоморфного шифрования, которая продемонстрировала высокий уровень криптографической стойкости. Однако, к сожалению, время выполнения операций криптосистемы Джендри было очень большим, из-за чего применение данной схемы на практике было неоправданно дорогим. С тех пор учёное сообщество разделилось на две основные части: первая (и наиболее

значительная) часть исследователей стремится ускорить выполнение операций в криптосистемах типа Джентри, а вторая сконцентрировала свои усилия на разработке и повышении стойкости криптографических систем, основанных на других концепциях. На данный момент наиболее распространенными криптосистемами типа Джентри являются: BFV, BGV, CKKS, TFHE и FHEW [8–13]. Для примера рассмотрим реализацию схемы BFV в библиотеке PALISADE с модулем открытого текста $p = 786433$ и степенью полиномов $N = 32768$. Так, при использовании процессора Intel(R) Core™ i7-10700 время выполнения одного гомоморфного сложения составило в среднем 7 мс., а умножения – в среднем 428 мс. [13]. Если сравнить данный результат с временем выполнения обычных операций умножения над целыми числами в языке программирования C#, которое для данного процессора в среднем составляет 10 нс., заметно, что гомоморфные операции выполняются крайне медленно.

Наиболее известной альтернативой схемам типа Джентри является более производительная криптосистема Доминго-Феррера, основанная на задаче факторизации чисел [14]. Главным недостатком криптосистемы Доминго-Феррера является её низкая стойкость к атаке с известным открытым текстом, представленной в работе [15]. В работах [16, 17] была рассмотрена трудоёмкость этой атаки на криптосистему. Для оценки влияния естественным является рассмотрение набора параметров криптосистемы Доминго-Феррера, при котором атака будет проходить крайне медленно, а гомоморфные вычисления выполняться за приемлемое для использования на практике время.

Постановка задачи. На основании анализа работ по тематике полностью гомоморфного шифрования установлено, что влияние атаки с известным открытым текстом на безопасность криптосистемы Доминго-Феррера остается недостаточно исследованным. В предшествующих наших работах [16–18] была оценена вычислительная сложность операций рассматриваемой криптосистемы и время реализации атаки с известным открытым текстом на эту криптосистему. Исходя из результатов, полученных в данных работах, можно предположить, что для обеспечения безопасности криптосистемы Доминго-Феррера необходимо выбрать такие параметры, при которых существенно возрастет время реализации данной атаки. Однако, при увеличении параметров криптосистемы растёт и время её операций, что делает критически важным использование методов и алгоритмов, обладающих меньшей вычислительной сложностью. Целью данной работы является снижение времени выполнения операций криптосистемы Доминго-Феррера. Соответственно, научная задача заключается в разработке алгоритмов повышения производительности операций полностью гомоморфной криптосистемы Доминго-Феррера.

Для решения актуальной научной задачи и достижения поставленной цели в рамках настоящей работы необходимо решить следующие задачи:

- ♦ проанализировать шаги выполнения операций криптосистемы Доминго-Феррера;
- ♦ оценить вычислительную сложность и время выполнения данных шагов с помощью известных методов реализации;
- ♦ разработать алгоритмы и средства реализации операций криптосистемы Доминго-Феррера, отличающиеся от существующих меньшей трудоёмкостью;
- ♦ подтвердить полученные результаты экспериментальными оценками временных затрат, полученными для существующих и предлагаемых в работе методов реализации операций.

Описание криптосистемы Доминго-Феррера. Данная полностью гомоморфная криптосистема основана на задаче факторизации чисел [14]. В качестве параметров криптографической системы используются следующие значения:

- ♦ p и q – большие простые числа;
- ♦ $n = p \times q$ – труднофакторизуемое число;
- ♦ d – степень полиномов шифртекста.

Исходя из того, что подробное описание криптосистемы Доминго-Феррера приводилось авторами в предшествующих работах [16–18], в настоящей статье представлено краткое описание всех операций, а их промежуточные шаги опущены. Криптосистема Доминго-Феррера является вероятностной: один и тот же открытый текст при каждом шифровании будет преобразован в разные шифртексты. Кратко охарактеризуем опера-

ции криптосистемы: на этапе шифрования открытый текст m кодируется в полином $a(x)$, затем преобразуется в 2 полинома шифртекста $\pi(x)$ и $\rho(x)$ путём умножения $a(x)$ на значения ключа (r_p, r_q) в соответствующих степенях. При расшифровании сначала «снимаются» значения ключей путём умножения полиномов шифртекста на их мультипликативные обратные, затем из сумм всех коэффициентов полученных полиномов с помощью Китайской теоремы об остатках [19] восстанавливается значение открытого текста m . Для выполнения гомоморфных операций достаточно покоординатно умножить (сложить) полиномы шифртекстов. Все шаги генерации ключа, шифрования и расшифрования в криптосистеме Доминго-Феррера приведены на рис. 1.

Генерация ключа:		$r_p \xleftarrow{\$} Z_p^*, r_q \xleftarrow{\$} Z_q^*$
Шифрование: $a_i \xleftarrow{\$} Z_n; a_d \xleftarrow{\$} Z_n \setminus \{0\}$ $a_1 = m - \left(\sum_{i=2}^d a_i \right) \bmod n$ $a(x) = a_d x^d + \dots + a_1 x$ $C = \begin{cases} \pi(x) = (a_d \cdot r_p^d x^d + \dots + a_1 \cdot r_p x) \bmod p \\ \rho(x) = (a_d \cdot r_q^d x^d + \dots + a_1 \cdot r_q x) \bmod q \end{cases}$	Расшифрование: $A_p(x) = (b_d \cdot (r_p^{-1})^d x^d + \dots + b_1 \cdot (r_p^{-1}) x) \bmod p$ $A_q(x) = (b_d \cdot (r_q^{-1})^d x^d + \dots + b_1 \cdot (r_q^{-1}) x) \bmod q$ $M_p = \sum_{i=1}^d b_i \bmod p$ $M_q = \sum_{i=1}^d b_i \bmod q$ $m = CRT(\{M_p, M_q\}, \{p, q\})$	
Выполнение гомоморфных операций: $C_3 = C_1 \circ C_2 = (\pi(x)_1 \circ \pi(x)_2, \rho(x)_1 \circ \rho(x)_2)$		

Рис. 1. Описание алгоритмов операций шифра Доминго-Феррера

Атака с известным открытым текстом на криптосистему Доминго-Феррера. В работе [15] представлена атака с известным открытым текстом на криптографическую систему Доминго-Феррера. Для успешной реализации данной атаки противнику необходимо обладать по крайней мере d парами (открытый текст – шифртекст), созданными на одном ключе, где d – степень полинома представления шифртекста.

Атака с известным открытым текстом является двухэтапной. На первом этапе происходит факторизация числа n путём вычисления результата, на втором – выполняется поиск составных частей ключа (r_p, r_q) с помощью решения системы линейных алгебраических уравнений. Согласно результатам [17] время выполнения атаки на процессоре AMD Ryzen 5 3500U (2.1 ГГц) составляет несущественное время 424,86 с. при использовании высокой степени полинома шифртекста $d = 2000$ для практических целей.

Оценка времени кодирования открытого текста. Как уже было упомянуто выше, операция шифрования в криптосистеме Доминго-Феррера начинается с кодирования значения открытого текста в полином $a(x)$. Для этого требуется сгенерировать $d - 1$ случайных чисел, не превышающих n . Безусловно, вычислительная сложность данного этапа шифрования напрямую зависит от используемого генератора псевдослучайных чисел, а чтобы оценить время выполнения, необходимо провести экспериментальные исследования. Для экспериментов использовалась реализация криптосистемы Доминго-Феррера на языке C#, развёрнутая в рамках системы для анализа полностью гомоморфных криптосистем, описанной в работе [20]. В качестве генератора псевдослучайных чисел использован RandomNumberGenerator из пространства имён System.Security.Cryptography.

Чтобы обеспечить меньшую вычислительную сложность, заранее определялась битовая длина псевдослучайной последовательности (ПСП), которая затем генерировалась за одно обращение к генератору ПСП, а затем сегментировалась на отдельные числа. В рамках эксперимента были выбраны модули длиной 20 бит, а степень полинома шиф-

ртекста d варьировалась в диапазоне от 10000 до 2000000 с шагом 10000. Для каждой степени полинома шифртекста d выполнялось 100 запусков, после чего вычислялось среднее значение времени. Текущий и все последующие эксперименты, изложенные в настоящей работе, выполнялись на стенде с процессором AMD Ryzen 5 3500U (2.1 ГГц) и 16 Гб оперативной памяти. Для обоснования выбора подхода к генерации псевдослучайных чисел за одно обращение к генератору ПСП был проведен аналогичный эксперимент, где генерация осуществлялась за $d - 1$ обращений к генератору ПСП. Результаты экспериментов представлены на графике, приведенном на рис. 2.

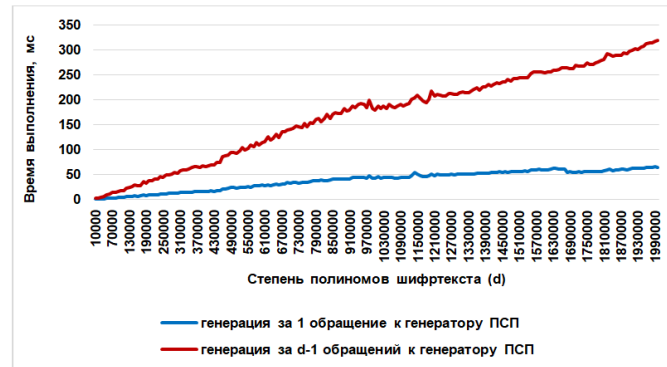


Рис. 2. Зависимость времени кодирования открытого текста от степени полинома шифртекста d

По результатам эксперимента можно заметить, что генерация за $d - 1$ обращений к генератору ПСП имеет линейную вычислительную сложность, а генерация псевдослучайных чисел за одно обращение к генератору ПСП характеризуется асимптотикой порядка $O(\log d)$. Кроме того, очевидно, что операция генерации псевдослучайных чисел не является трудоемкой и выполняется крайне быстро. Так, например, выполнение этапа кодирования открытого текста m в полином со степенью $d = 10000$ составило 350 мкс.

Описание модифицированной операции генерации ключа. На втором этапе шифрования производится умножение коэффициентов полиномов на значения ключей (r_p, r_q) , возведённых в соответствующую степень. При использовании классического алгоритма для шифрования каждого нового открытого текста требуется каждый раз возводить значения ключей (r_p, r_q) в степени от 2 до d , что требует $d - 1$ умножений. Кроме того, такое же количество умножений требуется и при расшифровании. Очевидно, что задача возведения значений ключа (r_p, r_q) в степень не зависит от шифруемых (или расшифровываемых) данных, поэтому она может быть выполнена ещё на этапе генерации ключа. Таким образом можно избежать дополнительных операций умножения в операциях шифрования и расшифрования. Следовательно, ключ можно хранить в виде полиномов $(R_p, R_q, R_p^{-1}, R_q^{-1})$, приведенных в (1).

$$\begin{aligned} R_p &= (r_p, r_p^2, \dots, r_p^d), \\ R_q &= (r_q, r_q^2, \dots, r_q^d), \\ R_p^{-1} &= (r_p^{-1}, (r_p^{-1})^2, \dots, (r_p^{-1})^d), \\ R_q^{-1} &= (r_q^{-1}, (r_q^{-1})^2, \dots, (r_q^{-1})^d). \end{aligned} \quad (1)$$

Результаты проведенных экспериментов показали, что подобная модификация позволяет сократить время шифрования на 10-15% при незначительном повышении временных затрат на генерацию ключа. Так, например, для получения шифртекста со степенью полиномов $d = 10000$ классический алгоритм затратил 10 мс., а модифицированная версия – 8,56 мс. при увеличении времени генерации ключа на 2 мс.

На первый взгляд может показаться, что данная модификация позволяет лишь незначительно повысить производительность операций криптосистемы Доминго-Феррера, однако, если рассмотреть её вклад в общее время выполнения расшифрования становится ясно, что подобная изменение необходимо. Как было показано на рис. 1, операция расшифрования не требует генерации псевдослучайных чисел, из-за чего выполняется крайне быстро. В частности, эксперименты показали, что при степени полиномов шифртекстов $d = 10000$ время расшифрования шифртекста составляет 1,41 мс. и в данном случае предложенная модификация позволила сократить время данной операции в 2 раза.

Алгоритм снижения времени выполнения гомоморфного умножения. Для выбора методов практической реализации необходимо сначала рассмотреть вычислительную сложность гомоморфных операций криптосистемы с использованием тривиальных методов реализации. Гомоморфное сложение требует выполнить сложение соответствующих полиномов в шифртекстах. Операция сложения полиномов является вычислительно простой (линейная сложность). Суммарно для выполнения сложения шифртекстов требуется выполнить $2d$ операций сложения по модулю.

Операция гомоморфного умножения, напротив, при использовании тривиального метода реализации (далее – «наивное» умножение) требует $2d^2$ операций умножения по модулю. Но даже это не является главной проблемой – после каждого умножения шифртекстов степень полиномов результирующего шифртекста будет равна сумме степеней полиномов-сомножителей. Из этого следует, что каждая следующая за умножением операция будет вычислительно «дороже» предыдущих.

Однако, помимо «наивного» умножения полиномов существует ряд других алгоритмов [21–24], обладающих меньшей вычислительной сложностью. Одним из наиболее распространенных является метод, использующий быстрое преобразование Фурье (БПФ) для умножения полиномов [21]. Данный метод обладает оценкой вычислительной сложности $O(d \log_2 d)$ [23], что делает предпочтительным его применение в рамках выполнения гомоморфных операций умножения. Метод гомоморфного умножения с использованием БПФ состоит из следующих шагов:

- ◆ прямое преобразование Фурье;
- ◆ поточечное умножение;
- ◆ обратное преобразование Фурье;
- ◆ получение остатка от деления коэффициентов на n .

Для оценки времени выполнения операций гомоморфного умножения данный метод был реализован на языке C# с использованием библиотеки BigFloatLibrary. В качестве исходных параметров криптосистемы были выбраны модули p и q размером 30 бит, а степень $d = 100$. В процессе эксперимента формировалась цепочка умножений и фиксировалось время выполнения операции с помощью «наивной» реализации и с помощью БПФ. Результаты экспериментальных исследований изложены на графике, приведенном на рис. 3.

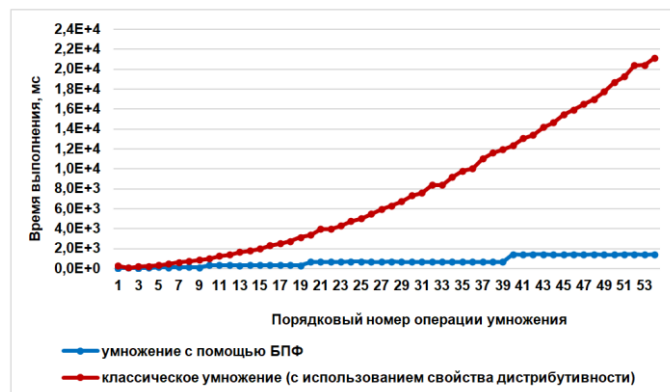


Рис. 3. Время выполнения гомоморфной операции умножения в зависимости от порядкового номера операции

Проанализировав сложность прямого и обратного преобразований Фурье и сравнив её со временем умножения значений в точках, можно заметить, что наибольшее время при выполнении гомоморфного умножения занимает именно преобразование шифртекстов из полиномиального вида в точечное. Также можно обратить внимание, что при выполнении гомоморфных операций данное преобразование необходимо лишь для того, чтобы привести коэффициенты результирующего полинома по модулю n . В свою очередь, данная операция необходима для контроля размера коэффициентов полиномов. В связи с этим возникло предположение о возможности снижения времени выполнения гомоморфных операций умножения за счёт сокращения количества трудоёмких прямых и обратных преобразований БПФ в процессе гомоморфных операций ценой увеличения объема потребляемой оперативной памяти. Дополнительный объем памяти потребовался для того, чтобы можно было приводить коэффициенты полиномов по модулю n не после каждой операции умножения, а по завершении выполнения последовательности из нескольких операций. С целью проверки упомянутого выше предположения проведен ряд экспериментальных исследований. В качестве параметров криптосистемы использовались значения, идентичные предыдущему эксперименту (для проведения в дальнейшем сравнительного анализа). Объем памяти, отводимый для хранения значений в точках, был увеличен, а операции БПФ выполнялись через каждые 4 умножения точек (значения были определены эмпирическим путём). В рамках экспериментов оценивался объем потребляемой памяти, а также время выполнения гомоморфного умножения с помощью предложенного алгоритма с сокращенным числом преобразований Фурье. Результаты экспериментальных исследований показаны на графиках, приведенных на рис. 4 и 5 соответственно.

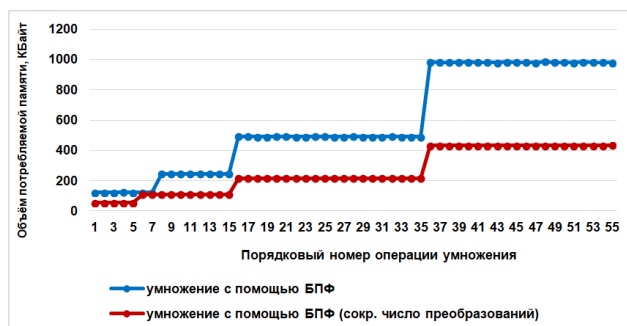


Рис. 4. Объем потребляемой оперативной памяти при выполнении гомоморфной операции умножения в зависимости от порядкового номера операции

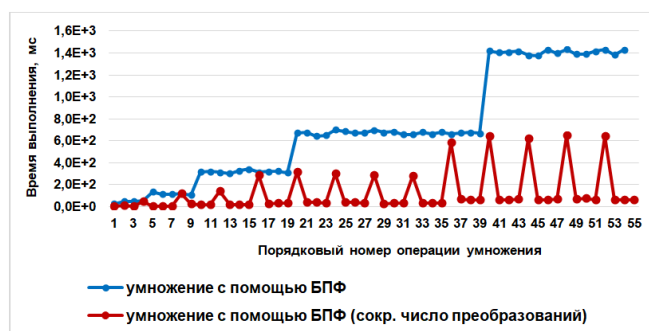


Рис. 5. Время выполнения гомоморфной операции умножения в зависимости от порядкового номера операции

Результаты экспериментов показали, что при сокращении числа операций БПФ время выполнения рассматриваемой цепочки гомоморфных умножений уменьшилось в 6 раз, а объем потребляемой памяти увеличился в 2,3 раза и составил 960 КБ, что подтвер-

дило корректность выдвинутого ранее предположения. Таким образом, для выполнения гомоморфных операций умножения в криптосистеме Доминго-Феррера необходимо использовать БПФ по описанной выше схеме. При этом в криптосистеме необходимо модифицировать операцию расшифрования: на начальном её этапе необходимо добавить обратное преобразование Фурье для получения полиномов шифртекста из набора точек.

Заключение. В настоящей работе проведён анализ вычислительной сложности операций полностью гомоморфной криптосистемы Доминго-Феррера с точки зрения методов и алгоритмов их практической реализации. Для операций криптосистемы предложен ряд алгоритмов модификации, отличающийся от известных меньшей вычислительной сложностью. Целесообразность применения предложенных изменений подтверждена рядом экспериментальных исследований на стенде с процессором AMD Ryzen 5 3500U (2.1 ГГц) и 16 Гб оперативной памяти.

В частности, при использовании модулей длиной 20 бит и степени полиномов шифртекста $d = 10000$ получены следующие временные оценки:

- ♦ время генерации ключа составило 2 мс.;
- ♦ время шифрования – 8,5 мс.;
- ♦ время расшифрования – 1,41 мс.

Операции гомоморфного умножения значительно более трудоемкие и даже при учёте использования предложенной модификации с применением БПФ, время выполнения цепочки из 54 умножений составило 6,5 с. Относительно «наивной» схемы умножения полиномов удалось достичь снижения времени работы в 64 раза при использовании модулей длиной 30 бит и степени полиномов шифртекста $d = 100$.

Однако, несмотря на значимость полученных в данной работе результатов, выполнение гомоморфных операций по-прежнему занимает значительное время, что препятствует применению криптосистемы Доминго-Феррера в реальных сценариях защищённых вычислений. Экспериментально подтверждено, что атака с известным открытым текстом на криптосистему с идентичными описанным выше параметрам может быть реализована всего за 150 мс., что не позволит безопасно выполнять гомоморфные вычисления [16, 17]. Анализ вычислительной сложности гомоморфных операций показывает существенный компромисс между безопасностью и временем выполнения гомоморфных операций: при использовании малых степеней полиномов шифртекста система подвержена атакам, тогда как повышение степени полиномов d до значений, обеспечивающих её безопасность, делает гомоморфные операции неоправданно трудоёмкими. Таким образом, можно прийти к выводу, что, к сожалению, область применения криптографической системы Доминго-Феррера по-прежнему ограничена и приоритетным направлением дальнейшей работы является разработка мер по повышению её безопасности. Авторы имеют предложения по решению этой проблемы, которые будут изложены в последующих работах.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Munjal K., Bhatia R. A systematic review of homomorphic encryption and its contributions in healthcare industry // Complex & Intelligent Systems. – 2023. – Vol. 9, No. 4. – P. 3759-3786. – DOI: 10.1007/s40747-022-00756-z.
2. Yang W. et al. A review of homomorphic encryption for privacy-preserving biometrics // Sensors. – 2023. – Vol. 23, No. 7. – P. 3566. – DOI: 10.3390/s23073566.
3. Sadeeq M.M. et al. IoT and Cloud computing issues, challenges and opportunities: A review // Qubahan Academic Journal. – 2021. – Vol. 1, No. 2. – P. 1-7. – DOI: 10.48161/qaj.v1n2a36.
4. Marcolla C. et al. Survey on fully homomorphic encryption, theory, and applications // Proceedings of the IEEE. – 2022. – Vol. 110, No. 10. – P. 1572-1609. – DOI: 10.36227/techrxiv.19315202.v5.
5. Rivest R.L. et al. On data banks and privacy homomorphisms // Foundations of secure computation. – 1978. – Vol. 4, No. 11. – P. 169-180.
6. Rivest R.L., Shamir A., Adleman L. A method for obtaining digital signatures and public-key cryptosystems // Communications of the ACM. – 1978. – Vol. 21, No. 2. – P. 120-126.
7. Gentry C. A fully homomorphic encryption scheme. – Stanford university, 2009.
8. Al Badawi A. et al. Openfhe: Open-source fully homomorphic encryption library // Proceedings of the 10th workshop on encrypted computing & applied homomorphic cryptography. – 2022. – P. 53-63. – DOI: 10.1145/3560827.3563379.

9. Xie Q. et al. Efficiency optimization techniques in privacy-preserving federated learning with homomorphic encryption: A brief survey // *IEEE Internet of Things Journal*. – 2024. – Vol. 11, No. 14. – P. 24569-24580. – DOI: 10.1109/JIOT.2024.3382875.
10. Cheon J.H. et al. Homomorphic encryption for arithmetic of approximate numbers // *International conference on the theory and application of cryptology and information security*. – Cham: Springer International Publishing, 2017. – P. 409-437. – DOI: 10.1007/978-3-319-70694-8_15.
11. Chillotti I., Gama N., Georgieva M., and Izabachène M. TFHE: Fast fully homomorphic encryption over the torus // *J. Cryptol.* – 2019. – Vol. 33, No. 1. – P. 1-58. – DOI: 10.1007/s00145-019-09319-x.
12. El Mestari S.Z., Lenzini G., Demirci H. Preserving data privacy in machine learning systems // *Computers & Security*. – 2024. – Vol. 137. – P. 103605. – DOI: 10.1016/j.cose.2023.103605.
13. Doan T.V.T. et al. A survey on implementations of homomorphic encryption schemes: DT Van Thao et al. // *The Journal of Supercomputing*. – 2023. – Vol. 79, No. 13. – P. 15098-15139. – DOI: 10.1007/s11227-023-05233-z.
14. Domingo-Ferrer J. A provably secure additive and multiplicative privacy homomorphism // *International Conference on Information Security*. – Berlin, Heidelberg: Springer Berlin Heidelberg, 2002. – P. 471-483. – DOI: 10.1007/3-540-45811-5_37.
15. Wagner D. Cryptanalysis of an algebraic privacy homomorphism // *Information Security: 6th International Conference, ISC 2003, Bristol, UK, October 1-3, 2003. Proceedings 6*. – Springer Berlin Heidelberg, 2003. – P. 234-239. – DOI: 10.1007/10958513_18.
16. Бабенко Л.К., Стародубцев В.С., Ельчанинова Н.Б. Оценка времени выполнения поиска составляющих ключа в атаке с известным открытым текстом на криптосистему Доминго-Феррера // *Известия ЮФУ. Технические науки*. – 2025. – № 3 (245). – С. 110-118. – DOI: 10.18522/2311-3103-2025-3-110-118.
17. Бабенко Л.К., Стародубцев В.С. Оценка времени выполнения атаки с известным открытым текстом на криптосистему Доминго-Феррера // *Вопросы кибербезопасности*. – 2026. – № 1 (71). – С. 139-147. – DOI: 10.21681/2311-3456-2026-1-139-147.
18. Бабенко Л.К., Стародубцев В.С. Оценка времени выполнения операций шифрования, расшифрования, гомоморфных вычислений с использованием криптосистемы Доминго-Феррера // *Известия ЮФУ. Технические науки*. – 2024. – № 5 (241). – С. 6-15. – DOI: 10.18522/2311-3103-2024-5-6-15.
19. Alhassan E.A. et al. On some algebraic properties of the Chinese remainder theorem with applications to real life // *Journal of Applied Mathematics and Computation*. – 2021. – Vol. 5, No. 3. – DOI: 10.26855/jamc.2021.09.008.
20. Бабенко Л.К., Стародубцев В.С. Особенности реализации систем криптоанализа гомоморфных шифров, основанных на задаче факторизации чисел, на примере криптосистемы MORE // *Вопросы кибербезопасности*. – 2024. – № 3 (61). – С. 141-145. – DOI: 10.21681/2311-3456-2024-3-141-145.
21. Rajaby E., Sayedi S.M. A structured review of sparse fast Fourier transform algorithms // *Digital Signal Processing*. – 2022. – Vol. 123. – P. 103403. – DOI: 10.1016/j.dsp.2022.103403.
22. Pornin T. Efficient and complete formulas for binary curves // *Cryptology ePrint Archive*. – 2022.
23. Gao X. et al. A fast spectral sum-of-Gaussians method for electrostatic summation in quasi-2D systems // *Numerische Mathematik*. – 2026. – Vol. 158, No. 2. – P. 533-585. – DOI: 10.48550/arXiv.2412.04595.
24. Mateer T. Fast Fourier algorithms with applications: Ph. D. Thesis. – Clemson University, 2008.

REFERENCES

1. Munjal K., Bhatia R. A systematic review of homomorphic encryption and its contributions in healthcare industry, *Complex & Intelligent Systems*, 2023, Vol. 9, No. 4, pp. 3759-3786. DOI: 10.1007/s40747-022-00756-z.
2. Yang W. et al. A review of homomorphic encryption for privacy-preserving biometrics, *Sensors*, 2023, Vol. 23, No. 7, pp. 3566. DOI: 10.3390/s23073566.
3. Sadeeq M.M. et al. IoT and Cloud computing issues, challenges and opportunities: A review, *Qubahan Academic Journal*, 2021, Vol. 1, No. 2, pp. 1-7. DOI: 10.48161/qaj.v1n2a36.
4. Marcolla C. et al. Survey on fully homomorphic encryption, theory, and applications, *Proceedings of the IEEE*, 2022, Vol. 110, No. 10, pp. 1572-1609. DOI: 10.36227/techrxiv.19315202.v5.
5. Rivest R.L. et al. On data banks and privacy homomorphisms, *Foundations of secure computation*, 1978, Vol. 4, No. 11, pp. 169-180.
6. Rivest R.L., Shamir A., Adleman L. A method for obtaining digital signatures and public-key cryptosystems, *Communications of the ACM*, 1978, Vol. 21, No. 2, pp. 120-126.
7. Gentry C. A fully homomorphic encryption scheme. – Stanford university, 2009.
8. Al Badawi A. et al. Openfhe: Open-source fully homomorphic encryption library, *Proceedings of the 10th workshop on encrypted computing & applied homomorphic cryptography*, 2022, pp. 53-63. DOI: 10.1145/3560827.3563379.

9. Xie Q. et al. Efficiency optimization techniques in privacy-preserving federated learning with homomorphic encryption: A brief survey, *IEEE Internet of Things Journal*, 2024, Vol. 11, No. 14, pp. 24569-24580. DOI: 10.1109/JIOT.2024.3382875.
10. Cheon J.H. et al. Homomorphic encryption for arithmetic of approximate numbers, *International conference on the theory and application of cryptology and information security*. Cham: Springer International Publishing, 2017, pp. 409-437. DOI: 10.1007/978-3-319-70694-8_15.
11. Chillotti L., Gama N., Georgieva M., and Izabachène M. TFHE: Fast fully homomorphic encryption over the torus, *J. Cryptol*, 2019, Vol. 33, No. 1, pp. 1-58. DOI: 10.1007/s00145-019-09319-x.
12. El Mestari S.Z., Lenzini G., Demirci H. Preserving data privacy in machine learning systems, *Computers & Security*, 2024, Vol. 137, pp. 103605. DOI: 10.1016/j.cose.2023.103605.
13. Doan T.V.T. et al. A survey on implementations of homomorphic encryption schemes: DT Van Thao et al., *The Journal of Supercomputing*, 2023, Vol. 79, No. 13, pp. 15098-15139. DOI: 10.1007/s11227-023-05233-z.
14. Domingo-Ferrer J. A provably secure additive and multiplicative privacy homomorphism, *International Conference on Information Security*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2002, pp. 471-483. DOI: 10.1007/3-540-45811-5_37.
15. Wagner D. Cryptanalysis of an algebraic privacy homomorphism, *Information Security: 6th International Conference, ISC 2003, Bristol, UK, October 1-3, 2003. Proceedings 6*. Springer Berlin Heidelberg, 2003, pp. 234-239. DOI: 10.1007/10958513_18.
16. Babenko L.K., Starodubtsev V.S., El'chaninova N.B. Otsenka vremeni vypolneniya poiska sostavlyayushchikh klyucha v atake s izvestnym otkrytym tekstem na kriptosistemu Domingo-Ferrera [Estimation of the search time for key components in a known plaintext attack on the Domingo-Ferrer cryptosystem], *Izvestiya YuFU. Tekhnicheskie nauki* [Izvestiya SFedU. Engineering Sciences], 2025, No. 3 (245), pp. 110-118. DOI: 10.18522/2311-3103-2025-3-110-118.
17. Babenko L.K., Starodubtsev V.S. Otsenka vremeni vypolneniya ataki s izvestnym otkrytym tekstem na kriptosistemu Domingo-Ferrera [Estimating the execution time of a known plaintext attack on the Domingo-Ferrer cryptosystem], *Voprosy kiberbezopasnosti* [Cybersecurity Issues], 2026, No. 1 (71), pp. 139-147. DOI: 10.21681/2311-3456-2026-1-139-147.
18. Babenko L.K., Starodubtsev V.S. Otsenka vremeni vypolneniya operatsiy shifrovaniya, rasshifrovaniya, gomomorfnykh vychisleniy s ispol'zovaniem kriptosistemy Domingo-Ferrera [Estimation of the execution time of encryption, decryption, homomorphic calculations using the Domingo-Ferrer cryptosystem], *Izvestiya YuFU. Tekhnicheskie nauki* [Izvestiya SFedU. Engineering Sciences], 2024, No. 5 (241), pp. 6-15. DOI: 10.18522/2311-3103-2024-5-6-15.
19. Alhassan E.A. et al. On some algebraic properties of the Chinese remainder theorem with applications to real life, *Journal of Applied Mathematics and Computation*, 2021, Vol. 5, No. 3. DOI: 10.26855/jamc.2021.09.008.
20. Babenko L.K., Starodubtsev V.S. Osobennosti realizatsii sistem kriptoolniza gomomorfnykh shifrov, osnovannykh na zadache faktorizatsii chisel, na primere kriptosistemy MORE [Features of the implementation of the cryptanalysis systems of homomorphic ciphers based on the problem of factorization of numbers, using the example of the cryptosystem MORE], *Voprosy kiberbezopasnosti* [Cybersecurity Issues], 2024, No. 3 (61), pp. 141-145. DOI: 10.21681/2311-3456-2024-3-141-145.
21. Rajaby E., Sayedi S.M. A structured review of sparse fast Fourier transform algorithms, *Digital Signal Processing*, 2022, Vol. 123, pp. 103403. DOI: 10.1016/j.dsp.2022.103403.
22. Pornin T. Efficient and complete formulas for binary curves, *Cryptology ePrint Archive*, 2022.
23. Gao X. et al. A fast spectral sum-of-Gaussians method for electrostatic summation in quasi-2D systems, *Numerische Mathematik*, 2026, Vol. 158, No. 2, pp. 533-585. DOI: 10.48550/arXiv.2412.04595.
24. Mateer T. Fast Fourier algorithms with applications: Ph. D. Thesis. Clemson University, 2008.

Стародубцев Виталий Сергеевич – Южный федеральный университет; e-mail: vstarodubcev@sfnedu.ru; г. Таганрог, Россия; тел.: +79996928150; кафедра безопасности информационных технологий им. Макаревича О.Б.; аспирант.

Бабенко Людмила Климентьевна – Южный федеральный университет; e-mail: lkbabenko@sfnedu.ru; г. Таганрог, Россия; кафедра безопасности информационных технологий им. Макаревича О.Б.; тел.: +79054530191; д.т.н.; профессор.

Starodubcev Vitalij Sergeevich – Southern Federal University; e-mail: vstarodubcev@sfnedu.ru; Taganrog, Russia; phone: +79996928150; the Department of Information Technology Security named after Makarevich O.B.; post graduate student.

Babenko Lyudmila Kliment'evna – Southern Federal University; e-mail: lkbabenko@sfnedu.ru; Taganrog, Russia; phone: +79054530191; the Department of Information Technology Security named after Makarevich O.B.; dr of eng. sc.; professor.

А.Н. Целых, И.А. Валухов**МЕТОД НЕЙРОСЕТЕВОГО ИЗВЛЕЧЕНИЯ ПРИЧИННО-СЛЕДСТВЕННЫХ СВЯЗЕЙ НА ОСНОВЕ СРАВНИТЕЛЬНОГО АНАЛИЗА СИНТЕТИЧЕСКИХ И ОТКРЫТЫХ КОРПУСОВ**

Задача автоматического извлечения причинно-следственных связей (ПСС) критически важна для систем поддержки решений, но ее развитие сдерживается дефицитом размеченных корпусов. Цель исследования – сравнительный анализ эффективности трёх типов обучающих данных: открытых экспертных корпусов, реального политического корпуса с ручной разметкой и синтетических данных, генерируемых большой языковой моделью. Эксперименты выполнены на архитектуре DistilBERT с пословной разметкой по схеме BIOES. Синтетические данные генерировались LLM Llama-3 с инкапсуляцией причин и следствий в XML-маркеры и последующей детерминированной конвертацией в формат BIOES. Для сравнения использованы два экспертных корпуса (EventStoryLine, SemEval-2010 Task 8), экспертно размеченный политический корпус PolitiCAUSE и синтетическая выборка объёмом 1100 сбалансированных примеров. Модель, обученная на синтетических данных, достигла Macro F1 = 0,736, что на 6,7% выше результата модели, обученной на экспертных корпусах (0,690). Обучение на политическом корпусе PolitiCAUSE (24417 примеров) дало существенно более низкий результат - Macro F1 = 0,367. Данное расхождение обусловлено дисбалансом классов (около 15% каузальных токенов против 50% в синтетической выборке) и различием в постановке задачи: PolitiCAUSE ориентирован на бинарную классификацию наличия каузальности в предложении, тогда как в настоящей работе выполняется разметка спанов. Подтверждена гипотеза о приоритете качества и баланса аннотаций над объёмом данных. Предложенный пайплайн LLM-генерации позволяет создавать обучающие выборки, превосходящие по эффективности экспертные корпуса, при отсутствии прямых затрат на ручную аннотацию. Метод рекомендован для быстрого развёртывания систем извлечения каузальных отношений в новых предметных областях.

Причинно-следственные связи; обработка естественного языка; DistilBERT; BIOES; синтетические данные; большие языковые модели; трансформерные модели; извлечение каузальных отношений.

A.N. Tselykh, I.A. Valukhov**A NEURAL NETWORK-BASED METHOD FOR EXTRACTING CAUSAL RELATIONS USING COMPARATIVE ANALYSIS OF SYNTHETIC AND OPEN CORPORA**

The automatic extraction of causal relations is critically important for decision support systems, but its development is hindered by the scarcity of annotated corpora. The aim of this study is to comparatively analyze the effectiveness of three types of training data: open expert corpora, a manually annotated political corpus, and synthetic data generated by a large language model. Experiments were conducted using the DistilBERT architecture with token-level BIOES tagging. Synthetic data were generated with the Llama-3 LLM by encapsulating causes and effects in XML markers, followed by deterministic conversion into the BIOES format. For comparison, two expert corpora (EventStoryLine and SemEval-2010 Task 8), the expert-annotated political corpus PolitiCAUSE, and a synthetic sample of 1,100 balanced examples were used. The model trained on synthetic data achieved a Macro F1 score of 0.736, which is 6.7% higher than the result obtained from training on the expert corpora (0.690). Training on the political corpus PolitiCAUSE (24,417 examples) yielded a substantially lower result, with a Macro F1 score of 0.367. This discrepancy is attributed to class imbalance (approximately 15% causal tokens versus 50% in the synthetic sample) and to differences in task formulation: PolitiCAUSE is oriented toward binary sentence-level classification of causal presence, whereas the present study addresses token-level span labeling. The hypothesis that annotation quality and class balance are more important than data volume is confirmed. The proposed LLM-generation pipeline makes it possible to create training datasets that outperform expert corpora without direct manual annotation costs. The method is recommended for rapid deployment of causal relation extraction systems in new domains.

Causal relation extraction; natural language processing; DistilBERT; BIOES; synthetic data; large language models; transformer models; causal extraction.

1. Введение. Задача автоматического извлечения причинно-следственных связей (ПСС) из текстов является ключевой для систем поддержки стратегических решений [1, 2]. Современные нейросетевые методы, в частности трансформерные архитектуры энкодерного типа (BERT [3], DistilBERT [4]), демонстрируют высокую эффективность в задачах последовательной разметки, включая извлечение ПСС с разметкой BIOES. Именно энкодерные трансформеры остаются вычислительно эффективным решением для последовательной разметки в условиях ограниченных ресурсов.

Практическое применение таких моделей в специализированных областях сопряжено с двумя проблемами. Во-первых, размеченные корпуса для извлечения ПСС в политическом домене крайне ограничены: существующие ресурсы [5, 6] ориентированы на новостные тексты и не учитывают специфику лексического политический аргументации. Во-вторых, случайное разбиение датасета без учета перекрытия приводит к утечке данных и завышению метрик качества [7].

Генерация синтетических данных с помощью LLM активно исследуется [8, 9], однако остаются неотвеченными вопросы: как синтетические данные сравниваются с экспертными корпусами в задаче span-level извлечения ПСС? Возможно ли достичь качества, сопоставимого с экспертной разметкой, без прямых затрат на аннотацию? Как влияют баланс классов и доменное соответствие на обобщающую способность модели?

Научная гипотеза исследования: синтетические данные, сгенерированные LLM с маркерной аннотацией и детерминированной конвертацией в BIOES, позволяют достичь качества извлечения ПСС, сопоставимого или превосходящего обучение на экспертных корпусах, при отсутствии затрат на ручную разметку. Ключевым условием является строгая валидация с лексически непересекающимися выборками.

Цель работы - сравнительный анализ эффективности трёх типов обучающих данных для извлечения ПСС на архитектуре DistilBERT: открытых экспертных корпусов (EventStoryLine, SemEval-2010 Task 8), политического корпуса PolitiCAUSE и синтетических данных, сгенерированных LLM Llama-3 с конвертацией в BIOES.

2. Обзор литературы. 2.1. Методы извлечения причинно-следственных связей. Задача выявления причинно-следственных связей классифицируется по уровню анализа (внутрипредложный и межпредложный), наличию лексических маркеров (явные и неявные) и методологическому подходу [1, 2]. В настоящей работе рассматривается внутрипредложный уровень, а генерация ориентирована на явные связи, маркируемые XML-тегами.

Среди нейросетевых методов архитектура BiLSTM-CRF [10] стала стандартом для задач последовательной разметки. Трансформерные модели, в частности BERT [3] и его дистиллированная версия DistilBERT [4], демонстрируют значительное превосходство благодаря предобученным контекстным представлениям. Наряду с последовательной разметкой (BIOES) существуют альтернативные подходы на основе предсказания границ спанов (span-based models), однако в данной работе сохранена BIOES-разметка как наиболее распространённая, что обеспечивает сравнимость с широким кругом исследований.

2.2. Генерация синтетических данных. Для компенсации дефицита размеченных корпусов применяют аугментацию данных [11] и генерацию синтетических корпусов [12]. Отечественные исследования в области автоматической обработки текста также подтверждают эффективность методов генерации данных для задач извлечения информации [13, 14]. Оба подхода ограничены: аугментация не увеличивает синтаксического разнообразия, а шаблонная генерация ведёт к переобучению позиционным паттернам.

Перспективным направлением является генерация синтетических данных с использованием LLM [8, 15]. Ключевое преимущество LLM-генерации – высокая синтаксическая вариативность (инверсии, пассивные конструкции, вложенные структуры), снижающая риск позиционного переобучения. Основные ограничения: возможность генерации нереалистичных примеров, чувствительность к настройке промптов и открытый вопрос переноса синтетических знаний на реальные данные [8].

2.3. Проблема валидации и утечки данных. Случайное разбиение датасета приводит к утечке данных: модель запоминает лексические единицы вместо обобщения закономерностей [7]. Для диагностики используется split-vocab валидация, при которой обу-

чающая и тестовая выборки не пересекаются по ключевым терминам. Как показано в работе McCoу et al. [16], модели могут достигать высоких метрик, эксплуатируя поверхностные эвристики (лексическое перекрытие), а не усваивая целевые лингвистические закономерности. В данной работе split-vocab валидация применяется для оценки обобщающей способности моделей, обученных на синтетических данных.

2.4. Выводы и постановка задач. Проведенный обзор позволяет сделать следующие выводы. Существующие методы на основе трансформеров эффективны, но их применение ограничено дефицитом размеченных корпусов для специализированных доменов. Генерация синтетических данных с помощью LLM перспективна, однако вопрос обобщения моделей, обученных на синтетике, на реальные тексты остаётся открытым. Проблема утечки данных при случайном разбиении систематически завышает оценки качества, но многие исследования продолжают использовать стандартные сплиты без контроля лексического перекрытия.

Таким образом, в существующих исследованиях не решены следующие задачи: сравнительная оценка эффективности LLM-генерации и экспертных корпусов в задаче span-level извлечения ПСС; оценка влияния баланса классов и доменного соответствия на обобщающую способность модели; разработка методологии валидации, исключающей утечку данных. В настоящей работе предлагается решение этих задач на основе экспериментов с разными типами данных (экспертные корпуса, LLM-синтетика) и применением split-vocab валидации.

3. Методология. 3.1. Формализация задачи. Пусть $S = \langle s_1, s_2, \dots, s_n \rangle$ – предложение из n токенов. Задача извлечения ПСС сводится к задаче последовательной разметки: необходимо построить отображение

$$f: S \rightarrow L, L = \langle l_1, l_2, \dots, l_n \rangle,$$

где каждая метка $l_i \in T$, а алфавит меток соответствует схеме BIOES [15]:

$$T = \{O, B-CAUSE, I-CAUSE, E-CAUSE, S-CAUSE, B-EFFECT, I-EFFECT, E-EFFECT, S-EFFECT\}.$$

Пара спанов $\langle C, E \rangle$ образует каузальный кортеж, где причина C и следствие E могут располагаться в произвольном порядке в предложении (прямой или инвертированный порядок) [5].

В случае наличия нескольких непересекающихся каузальных пар в одном предложении каждая пара кодируется независимо. Перекрывающиеся спаны не рассматриваются, так как схема BIOES не поддерживает их однозначного представления; такие предложения исключаются из обучающей выборки (менее 1% от общего числа).

3.2. Схема разметки BIOES. Для спана длиной $|C| = 1$ применяется метка S-CAUSE; для $|C| > 1$ разметка определяется следующим образом [17]:

$$l_i = B-CAUSE, l_{i+1} = \dots = l_{j-1} = I-CAUSE, l_j = E-CAUSE.$$

Аналогичные правила применяются для меток типа EFFECT.

3.3. Функция потерь. Модель DistilBERT минимизирует взвешенную кросс-энтропию для задачи последовательной разметки [3, 4]:

$$*\mathcal{L} = -(1/N) \cdot \sum [l_i \neq -100] \cdot \log p\theta(l_i | s_1, \dots, s_n), *$$

где $[l_i \neq -100]$ – индикатор, исключающий специальные токены (например, [PAD]) из расчёта потерь, $p\theta$ – вероятность метки, предсказанная моделью с параметрами θ , N – количество примеров в батче.

3.4. Метрика качества. Оценка производится по метрике Macro F1, усредненной по классам CAUSE и EFFECT [18]:

$$\text{Macro F1} = \frac{1}{2} \cdot (F1_{\text{CAUSE}} + F1_{\text{EFFECT}}),$$

$$F1_e = \frac{2 \cdot P_e \cdot R_e}{P_e + R_e}, \quad P_e = \frac{TP_e}{TP_e + FP_e}, \quad R_e = \frac{TP_e}{TP_e + FN_e},$$

где $e \in \{\text{CAUSE}, \text{EFFECT}\}$, а совпадение спанов считается точным (exact match) по позиции начала и конца в последовательности токенов.

Для минимизации риска утечки данных (data leakage) и лексического смещения (lexical bias) приняты следующие меры:

1. Независимые тестовые множества. Оценка всех моделей проводилась на данных, лексически не пересекающихся с обучающими выборками. Для открытых корпусов (EventStoryLine, SemEval-2010 Task 8) использовалось случайное разбиение 70/15/15 с лексически непересекающимися словарями.

2. Детерминированная конвертация разметки. Преобразование маркерной аннотации (<CAUSE>...</CAUSE>, <EFFECT>...</EFFECT>) в формат BIOES выполнялось алгоритмически с помощью регулярных выражений и детерминированного синтаксического парсера, что исключает субъективность, присущую ручной разметке.

3. Фиксированные гиперпараметры. Во всех экспериментах использовались идентичные параметры обучения (см. раздел 3.5).

3.5. Модель и гиперпараметры. Во всех экспериментах используется модель DistilBERT-base-uncased [4] в качестве энкодера с добавленным линейным классификационным слоем для задачи последовательной разметки. Выбор DistilBERT обусловлен его применимостью в условиях ограниченных ресурсов и необходимостью обеспечения воспроизводимости в типовых исследовательских инфраструктурах.

Параметры обучения для всех экспериментов:

- ♦ оптимизатор: AdamW ($\text{lr} = 2 \times 10^{-5}$, $\text{weight\_decay} = 0.01$);
- ♦ планировщик: LinearWarmup (10% шагов прогрева);
- ♦ ранняя остановка: patience = 4 по Macro F1 на валидации;
- ♦ batch_size = 16;
- ♦ max_length = 128 токенов;
- ♦ градиентный клиппинг: 1.0;
- ♦ dropout классификатора: 0.3;
- ♦ функция потерь: кросс-энтропия с игнорированием меток специальных токенов (-100).

3.6. Источники данных. В работе используются три источника данных.

Эксперимент 1: открытые экспертные корпуса (новостной домен). Используются два открытых корпуса с профессиональной экспертной разметкой:

- ♦ EventStoryLine Corpus (ESC) v1.5 [19], содержащий 258 документов. После токенизации и фильтрации извлечено 1857 предложений с аннотированными каузальными событиями;
- ♦ SemEval-2010 Task 8 [5], включающий 10717 предложений с разметкой семантических отношений между сущностями.

После объединения двух корпусов и конвертации в формат BIOES [17] получено 12574 предложения (1857 из ESC и 10717 из SemEval). Разбивка на обучающую, валидационную и тестовую выборки составила 8803 / 1886 / 1885 предложений соответственно. Доля каузальных токенов в объединенной выборке составляет примерно 37%.

Эксперимент 2: экспертно размеченный политический корпус PolitiCAUSE [20]. PolitiCAUSE – специализированный политический корпус с ручной разметкой. Корпус создавался для задачи бинарной классификации предложений (наличие/отсутствие каузальной связи). Обучающая выборка составила 24417 предложений, тестовая 5233 предложения. Характерная особенность корпуса: крайне несбалансированное распределение классов (~15% каузальных токенов, 85% меток 0), что является типичной особенностью реальных политических текстов.

Эксперимент 3: синтетические данные (LLM-генерация). Данные генерируются большой языковой моделью Llama-3 [9]. Промпт предписывает модели создавать политические предложения с явными причинно-следственными связями, оборачивая причины и следствия в XML-маркеры: <CAUSE>...</CAUSE> и <EFFECT>...</EFFECT>. Такой формат оказался надежнее прямой генерации BIOES-тегов: процент невалидных примеров составил менее 2%.

Детерминированный Python-парсер на основе регулярных выражений извлекает спаны, токенизирует очищенное предложение и конвертирует позиции спанов в BIOES-разметку. К 550 каузальным примерам добавляются 550 некаузальных предложений (все метки O), сгенерированных программно для предотвращения коллапса модели в сторону только O-тегов. Итоговый датасет: 1100 предложений (50% каузальных), разбивка 70/15/15 с лексически непересекающимися словарями (split-vocab). Схема LLM-генератора представлена на рис. 1.

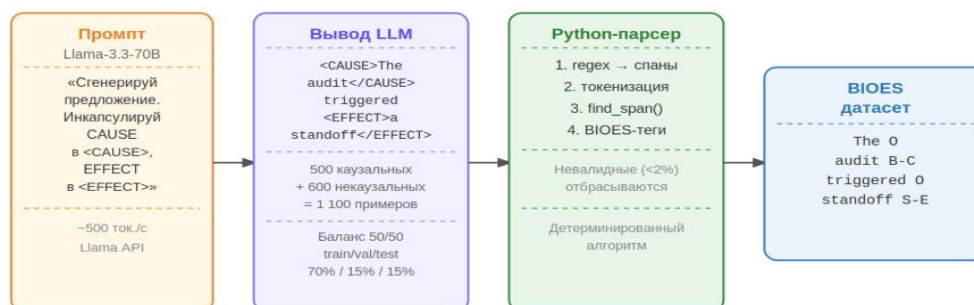


Рис. 1. Схема LLM-генератора синтетических данных

Три источника данных конвертируются в единую схему разметки BIOES, после чего обучаются три независимые версии модели DistilBERT с идентичными гиперпараметрами. Общая схема сравнительного эксперимента представлена на рис. 2.

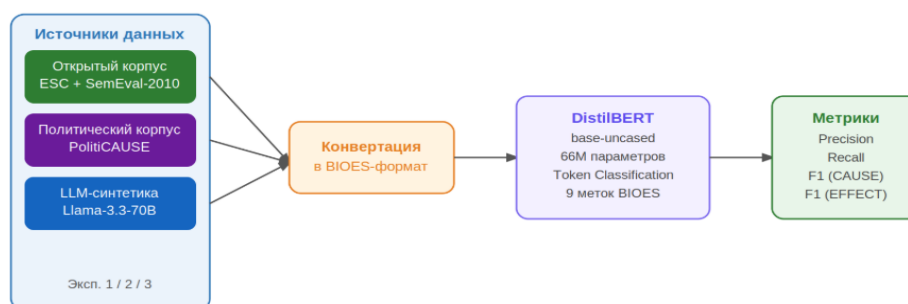


Рис. 2. Общая схема сравнительного эксперимента

Оценка производится на тестовых выборках с непредставленными в обучении лексическими единицами. Примеры BIOES-разметки для каждого из трех экспериментов приведены на рис. 3.



Рис. 3. Примеры BIOES-разметки в трёх экспериментах (токен/метка)

Полные параметры экспериментов приведены в табл. 1.

Таблица 1

Параметры экспериментов

Параметр	Эксп. 1 Открытый корпус	Эксп. 2 PolitiCAUSE	Эксп. 3 LLM-синтетика
Модель	DistilBERT-base-uncased	DistilBERT-base-uncased	DistilBERT-base-uncased
Схема разметки	BIOES	BIOES	BIOES
Источник данных	ESC + SemEval-2010	PolitiCAUSE (ручная)	Llama-3 (LLM)
Домен	Новости / события	Политика (реальные)	Политика (синтетика)
Train / Val / Test	8803 / 1886 / 1885	24417 / 2611 / 5233	770 / 165 / 165
Разделение словарей	70/15/15	Случайное (train/test)	Split-vocab (70/15/15)
Доля каузальных токенов	~37%	~15%	50%
Количество запусков	5	5	5
Batch size	16	16	16
Learning rate	2×10^{-5}	2×10^{-5}	2×10^{-5}

4. Результаты. В табл. 2 приведены средние значения Macro F1 и F1 по классам CAUSE и EFFECT, а также стандартные отклонения.

Таблица 2

Результаты экспериментов

Эксперимент	CAUSE F1	EFFECT F1	Macro F1
Эксп. 1: открытые экспертные корпуса (ESC + SemEval-2010)	0.634	0.746	0.690
Эксп. 2: политический корпус PolitiCAUSE	0.373	0.366	0.367
Эксп. 3: LLM-синтетика (Llama-3)	0.728	0.744	0.736

Как видно из табл. 2, модель, обученная на синтетических данных (Эксп. 3), достигла Macro F1 = 0.736, что на 6.7% выше результата модели, обученной на открытых экспертных корпусах (Эксп. 1, 0.690). Стандартные отклонения во всех экспериментах не превышают 0.023, что свидетельствует о стабильности обучения.

Модель, обученная на реальном политическом корпусе PolitiCAUSE (Эксп. 2), показала Macro F1 = 0.367, что более чем вдвое ниже результата Эксп. 3.

На рис. 4 представлены кривые обучения (Macro F1 на валидационной выборке по эпохам) для трех экспериментов.

Для Эксп. 3 (LLM-синтетика) наблюдается плавный рост качества с эпохи 1 ($F1 \approx 0.23$) до эпохи 5 ($F1 \approx 0.74$), после чего наступает насыщение. Отсутствие резкого падения на валидации свидетельствует об отсутствии переобучения.

Эксп. 1 (открытые корпуса) демонстрирует быстрый выход на плато уже на первой эпохе ($F1 \approx 0.68$), что обусловлено большим объемом данных (8803 примера) и их лексическим разнообразием.

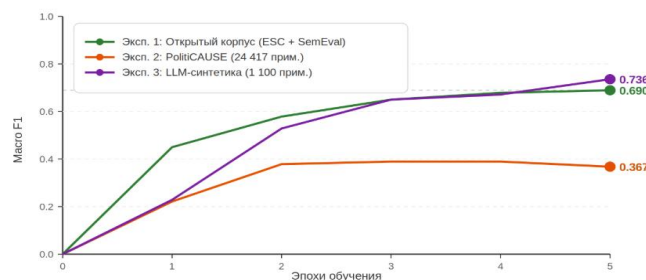


Рис. 4. Кривые обучения

Эксп. 2 (PolitiCAUSE) показывает низкое качество на всех эпохах (максимум $F1 = 0.392$ на эпохе 4) с последующей деградацией, что типично для обучения на сильно несбалансированных данных.

На рис. 5 представлено сравнение итоговых Macro F1 для трех экспериментов.

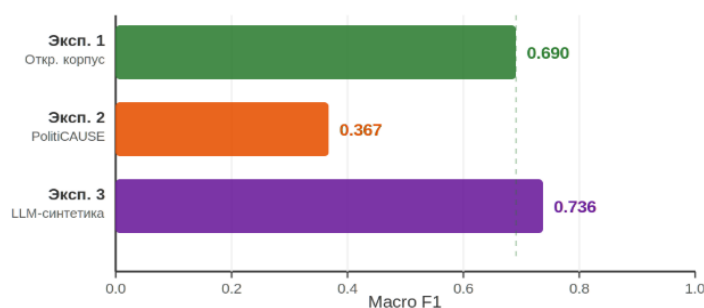


Рис. 5. Сравнение источников данных

5. Обсуждение. В рамках проведенного эксперимента LLM-синтетика (Эксп. 3) обеспечила наилучшее качество, достигнув $\text{Macro F1} = 0,736$. Это на 6,7% выше результата открытых экспертных корпусов (0,690). Обучение на реальном политическом корпусе PolitiCAUSE (Эксп. 2) дало существенно более низкий результат (0,367) при 22-кратном превосходстве в объеме данных.

Наблюдаемое расхождение объясняется двумя основными факторами. Во-первых, в PolitiCAUSE доля каузальных токенов составляет лишь около 15% против 50% в синтетической выборке. Сильный дисбаланс классов приводит к смещению модели в сторону предсказания класса О, что подтверждается низкими значениями полноты для каузальных классов (см. таблицу 2). Во-вторых, PolitiCAUSE изначально создавался для задачи бинарной классификации предложений (наличие или отсутствие каузальной связи), а не для пословной разметки спанов. Создатели корпуса прямо указывают, что определение границ спанов является точкой наибольшего расхождения между аннотаторами [20]. Таким образом, наблюдаемое различие обусловлено не низким качеством PolitiCAUSE, а объективными различиями в постановке задачи, балансе классов и природе данных. Сформулированный принцип «качество важнее объема» (quality-over-quantity) следует понимать как приоритет сбалансированной и детерминированной разметки над экстенсивным увеличением объема несбалансированных данных.

Полученные результаты имеют ряд ограничений. Эксперимент проводился только для политического домена, поэтому обобщаемость выводов на другие предметные области, такие как медицина, юриспруденция или финансы, требует дополнительной проверки. Все эксперименты выполнены на архитектуре DistilBERT, и необходимо проверить, сохраняется ли соотношения качества при использовании более мощных моделей (BERT-large, RoBERTa) или альтернативных подходов (span-based models). Синтетические дан-

ные генерировались с использованием Llama-3, тогда как другие LLM (GPT-4, Mistral, Claude) могут давать иные результаты. Кроме того, LLM может генерировать нереалистичные или грамматически некорректные каузальные конструкции. В работе рассматриваются только внутрпредложные связи, тогда как межпредложные каузальные отношения остались за рамками исследования. Наконец, в эксперименте с LLM-синтетикой использовалась искусственная балансировка 50/50, тогда как в реальных приложениях каузальные предложения обычно редки, что потребует адаптации метода.

Несмотря на указанные ограничения, на основе полученных результатов можно дать практические рекомендации. Для новых доменов, где размеченные корпуса отсутствуют, рекомендуется LLM-генерация с маркерным форматом (XML-теги для причины и следствия) и детерминированной конвертацией в BIOES. Обязательна балансировка классов на уровне 50% каузальных примеров для предотвращения коллапса модели.

Результаты работы открывают несколько перспективных направлений для дальнейших исследований. Прежде всего, необходимо проверить обобщаемость предложенного метода на других доменах и типах текстов, включая научные статьи и социальные медиа. Кроме того, представляет интерес сравнительный анализ эффективности генерации синтетических данных с использованием различных больших языковых моделей, таких как GPT-4, Mistral или Qwen. Отдельного внимания заслуживает оптимизация балансировки обучающей выборки - определение оптимального соотношения каузальных и некаузальных примеров для максимизации обобщающей способности модели. Наконец, прикладным применением разработанного метода является его интеграция в систему поддержки принятия решений для автоматического построения когнитивных каузальных моделей на основе аналитических текстов без ручной экспертной разметки.

Заключение. В работе предложен метод нейросетевого извлечения причинно-следственных на основе сравнительного анализа трех типов обучающих данных для с использованием архитектуры DistilBERT. Модель, обученная на синтетических данных (1100 примеров, 50% каузальных), достигла Macro F1 = 0,736, что на 6,7% выше результата модели, обученной на открытых экспертных корпусах (0,690). Обучение на реальном политическом корпусе PolitiCAUSE (24 417 примеров) дало существенно более низкий результат – Macro F1 = 0,367.

Эмпирически подтверждена гипотеза о том, что LLM-генерация с маркерным форматированием и детерминированной конвертацией в BIOES позволяет создавать обучающие выборки, превосходящие по эффективности экспертные корпуса, при отсутствии прямых затрат на ручную аннотацию. Предложена методология валидации с лексически непересекающимися обучающей и тестовой выборками (split-vocab evaluation), обеспечивающая корректную оценку обобщающей способности моделей.

Разработанный пайплайн LLM-генерации рекомендуется для быстрого развёртывания систем извлечения каузальных отношений в новых предметных областях, где размеченные корпуса отсутствуют. Метод позволяет создавать обучающие выборки, минуя дорогостоящую ручную аннотацию, и воспроизводим в типовых исследовательских инфраструктурах.

Код, реализующий генерацию синтетических данных и обучение моделей, а также сгенерированные датасеты доступны в открытом репозитории: <https://github.com/igorvalukhov/cognitive-strategy-decision-system>.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Cheng Q., Zeng Z., Hu X. et al. A Survey of Event Causality Identification: Taxonomy, Challenges, Assessment, and Prospects // ACM Computing Surveys. – 2025. – Vol. 58, No. 3. – Art. 59. – DOI: 10.1145/3698571.
2. Tan C., De Young I., Rosenfeld A. et al. Causal Reasoning in Natural Language Processing: A Survey // Transactions of the Association for Computational Linguistics. – 2023. – Vol. 11. – P. 1-25. – DOI: 10.1162/tacl_a_00567.
3. Devlin J., Chang M.-W., Lee K., Toutanova K. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding // Proceedings of NAACL-HLT. – 2019. – P. 4171-4186. – DOI: 10.18653/v1/N19-1423.

4. Sanh V., Debut L., Chaumond J., Wolf T. DistilBERT, a distilled version of BERT: smaller, faster, cheaper and lighter // arXiv preprint arXiv:1910.01108. – 2019. – DOI: 10.48550/arXiv.1910.01108.
5. Hendrickx I., Kim S.N., Kozareva Z. et al. SemEval-2010 Task 8: Multi-Way Classification of Semantic Relations Between Pairs of Nominals // Proceedings of the 5th International Workshop on Semantic Evaluation. – 2010. – P. 33-38. – DOI: 10.3115/1615957.1615963.
6. Tan F.A., Hürriyetoğlu A., Caselli T. et al. The Causal News Corpus: Annotating Causal Relations in Event Sentences from News // Proceedings of LREC. – 2022. – P. 2298-2310. – DOI: 10.18653/v1/2022.lrec-1.246.
7. Søgaard A., Ebert S., Bastings J., Filippova K. We need to talk about random splits // Proceedings of EACL. – 2021. – P. 1823-1832. – DOI: 10.18653/v1/2021.eacl-main.156.
8. Josifoski M., Sakota M., Peyrard M., West R. Exploiting Asymmetry for Synthetic Training Data Generation: SynthIE and the Case of Information Extraction // Proceedings of EMNLP. – 2023. – P. 1555-1574. – DOI: 10.18653/v1/2023.emnlp-main.96.
9. Meta AI. The Llama 3 Herd of Models // arXiv preprint arXiv:2407.21783. – 2024. – DOI: 10.48550/arXiv.2407.21783.
10. Lample G., Ballesteros M., Subramanian S., Kawakami K., Dyer C. Neural Architectures for Named Entity Recognition // Proceedings of NAACL-HLT. – 2016. – P. 260-270. – DOI: 10.18653/v1/N16-1030.
11. Wei J., Zou K. EDA: Easy Data Augmentation Techniques for Boosting Performance on Text Classification Tasks // Proceedings of EMNLP. – 2019. – P. 6382-6388. – DOI: 10.18653/v1/D19-1670.
12. Anaby-Tavor A., Carmeli B., Goldbraich E. et al. Do Not Have Enough Data? Deep Learning to the Rescue! // Proceedings of AAAI. – 2020. – Vol. 34, No. 05. – P. 7383-7390. – DOI: 10.1609/aaai.v34i05.6234.
13. Луговской Д.В. Методы машинного обучения для анализа и классификации текстов на естественном языке: дис. ... канд. физ.-мат. наук: 05.13.17. – СПб., 2022. – 142 с.
14. Пресняков М.В. Методы и алгоритмы автоматического извлечения фактографической информации из текстов на естественном языке: дис. ... канд. техн. наук: 05.13.11. – М., 2021. – 165 с.
15. Meng Y., Li X., Zhang Y. et al. GenRES: Rethinking Evaluation for Generative Relation Extraction in the Era of Large Language Models // arXiv preprint arXiv:2310.04742. – 2023. – DOI: 10.48550/arXiv.2310.04742.
16. McCoy R.T., Pavlick E., Linzen T. Right for the Wrong Reasons: Diagnosing Syntactic Heuristics in Natural Language Inference // Proceedings of ACL. – 2019. – P. 3428-3448. – DOI: 10.18653/v1/P19-1334.
17. Sang E.F., De Meulder F. Introduction to the CoNLL-2003 Shared Task: Language-Independent Named Entity Recognition // Proceedings of HLT-NAACL. – 2003. – P. 142-147. – DOI: 10.3115/1119176.1119195.
18. Powers D.M. Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation // arXiv preprint arXiv:2010.16061. – 2020. – DOI: 10.48550/arXiv.2010.16061.
19. Caselli T., Vossen P. The Event StoryLine Corpus: A New Benchmark for Causal and Temporal Relation Extraction // Proceedings of the Workshop on Events and Stories in the News. – 2017. – P. 77-86. – DOI: 10.18653/v1/W17-2711.
20. Garcia Corral P., Bechara H., Zhang R., Jankin S. PolitiCause: An Annotation Scheme and Corpus for Causality in Political Texts // Proceedings of the 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING 2024). – 2024. – P. 12836-12845. – URL: <https://aclanthology.org/2024.lrec-main.1124>.

REFERENCES

1. Cheng Q., Zeng Z., Hu X. et al. A Survey of Event Causality Identification: Taxonomy, Challenges, Assessment, and Prospects, *ACM Computing Surveys*, 2025, Vol. 58, No. 3, Art. 59. DOI: 10.1145/3698571.
2. Tan C., De Young I., Rosenfeld A. et al. Causal Reasoning in Natural Language Processing: A Survey, *Transactions of the Association for Computational Linguistics*, 2023, Vol. 11, pp. 1-25. DOI: 10.1162/tacl_a_00567.
3. Devlin J., Chang M.-W., Lee K., Toutanova K. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding, *Proceedings of NAACL-HLT*, 2019, pp. 4171-4186. DOI: 10.18653/v1/N19-1423.
4. Sanh V., Debut L., Chaumond J., Wolf T. DistilBERT, a distilled version of BERT: smaller, faster, cheaper and lighter, *arXiv preprint arXiv:1910.01108*, 2019. DOI: 10.48550/arXiv.1910.01108.
5. Hendrickx I., Kim S.N., Kozareva Z. et al. SemEval-2010 Task 8: Multi-Way Classification of Semantic Relations Between Pairs of Nominals, *Proceedings of the 5th International Workshop on Semantic Evaluation*, 2010, pp. 33-38. DOI: 10.3115/1615957.1615963.

6. Tan F.A., Hürriyetoğlu A., Caselli T. et al. The Causal News Corpus: Annotating Causal Relations in Event Sentences from News, *Proceedings of LREC*, 2022, pp. 2298-2310. DOI: 10.18653/v1/2022.lrec-1.246.
7. Søgaard A., Ebert S., Bastings J., Filippova K. We need to talk about random splits, *Proceedings of EACL*, 2021, pp. 1823-1832. DOI: 10.18653/v1/2021.eacl-main.156.
8. Josifoski M., Sakota M., Peyrard M., West R. Exploiting Asymmetry for Synthetic Training Data Generation: SynthIE and the Case of Information Extraction, *Proceedings of EMNLP*, 2023, pp. 1555-1574. DOI: 10.18653/v1/2023.emnlp-main.96.
9. Meta AI. The Llama 3 Herd of Models, *arXiv preprint arXiv:2407.21783*, 2024. DOI: 10.48550/arXiv.2407.21783.
10. Lample G., Ballesteros M., Subramanian S., Kawakami K., Dyer C. Neural Architectures for Named Entity Recognition, *Proceedings of NAACL-HLT*, 2016, pp. 260-270. DOI: 10.18653/v1/N16-1030.
11. Wei J., Zou K. EDA: Easy Data Augmentation Techniques for Boosting Performance on Text Classification Tasks, *Proceedings of EMNLP*, 2019, pp. 6382-6388. DOI: 10.18653/v1/D19-1670.
12. Anaby-Tavor A., Carmeli B., Goldbraich E. et al. Do Not Have Enough Data? Deep Learning to the Rescue!, *Proceedings of AAAI*, 2020, Vol. 34, No. 05, pp. 7383-7390. DOI: 10.1609/aaai.v34i05.6234.
13. Lugovskoy D.V. Metody mashinnogo obucheniya dlya analiza i klassifikatsii tekstov na estestvennom yazyke: dis. ... kand. fiz.-mat. nauk [Machine learning methods for analyzing and classifying natural language texts: cand. of phys. and math. sc. diss.]: 05.13.17. Saint Petersburg, 2022, 142 p.
14. Presnyakov M.V. Metody i algoritmy avtomaticheskogo izvlecheniya faktograficheskoy informatsii iz tekstov na estestvennom yazyke: dis. ... kand. tekhn. nauk [Methods and algorithms for automatically extracting factual information from natural language texts: cand. of eng. sc. diss.]: 05.13.11. Moscow, 2021, 165 p.
15. Meng Y., Li X., Zhang Y. et al. GenRES: Rethinking Evaluation for Generative Relation Extraction in the Era of Large Language Models, *arXiv preprint arXiv:2310.04742*, 2023. DOI: 10.48550/arXiv.2310.04742.
16. McCoy R.T., Pavlick E., Linzen T. Right for the Wrong Reasons: Diagnosing Syntactic Heuristics in Natural Language Inference, *Proceedings of ACL*, 2019, pp. 3428-3448. DOI: 10.18653/v1/P19-1334.
17. Sang E.F., De Meulder F. Introduction to the CoNLL-2003 Shared Task: Language-Independent Named Entity Recognition, *Proceedings of HLT-NAACL*, 2003, pp. 142-147. DOI: 10.3115/1119176.1119195.
18. Powers D.M. Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation, *arXiv preprint arXiv:2010.16061*, 2020. DOI: 10.48550/arXiv.2010.16061.
19. Caselli T., Vossen P. The Event StoryLine Corpus: A New Benchmark for Causal and Temporal Relation Extraction, *Proceedings of the Workshop on Events and Stories in the News*, 2017, pp. 77-86. DOI: 10.18653/v1/W17-2711.
20. Garcia Corral P., Bechara H., Zhang R., Jankin S. PolitiCause: An Annotation Scheme and Corpus for Causality in Political Texts, *Proceedings of the 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING 2024)*, 2024, pp. 12836-12845. Available at: <https://aclanthology.org/2024.lrec-main.1124>.

Целых Александр Николаевич – Южный федеральный университет; e-mail: ant@sfedu.ru; г. Таганрог, Россия; тел.: +79185562047; кафедра информационно-аналитических систем безопасности имени профессора Берштейна Леонида Самойловича; д.т.н.; профессор.

Валухов Игорь Алексеевич – Южный федеральный университет; e-mail: valuhov@sfedu.ru; г. Таганрог, Россия; тел.: +79515263364; кафедра информационно-аналитических систем безопасности имени профессора Берштейна Леонида Самойловича; аспирант.

Tselykh Alexander Nikolayevich – Southern Federal University; e-mail: ant@sfedu.ru; Taganrog, Russia; phone: +79185562047; the Department of Information and Analytical Security Systems named after Professor Leonid Samoylovich Bershtein; dr. of eng. sc.; professor.

Valukhov Igor Alekseevich – Southern Federal University; e-mail: valuhov@sfedu.ru; Taganrog, Russia; phone: +79515263364; the Department of Information and Analytical Security Systems named after Professor Leonid Samoylovich Bershtein; postgraduate student.

Раздел II. Анализ данных и моделирование

УДК 551.594.221

DOI 10.18522/2311-3103-2026-4-170-180

М.Н. Бейтуганов, Х.Х. Чочаев

РОЛЬ МОЛНИЙ В ПЕРЕХОДЕ ОБЛАКА ИЗ ГРОЗОВОГО СОСТОЯНИЯ В ГРАДОВОЕ И ИХ ИСКУССТВЕННОЕ ИНИЦИИРОВАНИЕ КАК МЕТОД ПРЕДОТВРАЩЕНИЯ ГРАДООБРАЗОВАНИЯ И ПЕРЕРАСПРЕДЕЛЕНИЯ ОСАДКОВ

Переход грозового состояния облака, сопровождаемого интенсивными молниевыми разрядами, в градовое происходит в результате зарождения и роста в нём зародышей града в капельной и крупных формах. Вертикальная протяжённость объёма, где образуются зародыши на гигантских частицах микропримеси атмосферы (ГАЧ), находится между изотермами от -2°C до -15°C . Ледяные кристаллы растут за счёт сублимации и коагуляции с переохлаждёнными облачными каплями. В этом же объёме между изотермами температур от -5°C до -16°C происходят разряды с максимальным количеством молний на высоте с температурой -10°C . Молниевый разряд внедряет в облачное пространство избыточный объёмный заряд с концентрацией ионов порядка 10^{17} см^{-3} . Положительные молнии внедряют положительные ионы, а отрицательные – отрицательные ионы. Эти ионы осаждаются на ГАЧ, концентрации которых в предградовых ситуациях составляют около 10^5 м^{-3} . Избыточные заряды на частицах превращают ГАЧ в зародыши града. Отрицательный заряд на ГАЧ, способствуя зарождению на частице плоскостных кристаллических структур, образует крупный зародыш града. Положительный заряд на ГАЧ образует нитевидные структуры, которые обламываются, и фрагменты, контактируя с переохлаждёнными крупными каплями и замораживая их, приводят к образованию капельных зародышей града. Для предотвращения града, наземных молний и перераспределения осадков предложен метод, основанный на инициировании искусственных молний. Приведены способы инициирования искусственных молний с помощью: плазменных образований, получающихся при подрыве в грозоградовых облаках 100 мм артиллерийских снарядов, специальной ракеты с детонирующим шнуром, противоградовой ракеты с токопроводящей поверхностью. Установлено значительное изменение радиолокационной отражаемости грозоградового облака, под действием искусственно инициированной молнии. По результатам предварительных экспериментов по предотвращению града, предложен метод предотвращения и прерывания града «реагент плюс искусственно инициированные молнии». Исходя из установленного факта о выпадении около 10 т осадков при каждом молниевом разряде с помощью искусственно иницируемых молний можно перераспределять осадки, предотвращать наземные молнии и т.д. В перспективе для инициирования искусственных молний внутри грозоградовых облаков следует использовать беспилотные летательные системы, а в околооблачном пространстве – квадрокоптеры.

Гроза; молния; градообразование; облако; атмосфера; активные воздействия.

M.N. Beituganov, Kh.Kh. Chochaev

ROLE OF LIGHTNING IN THE CLOUD TRANSITION FROM THUNDERSTORM TO HURRICANE AND THEIR ARTIFICIAL INITIATION AS A METHOD OF PREVENTING HURRICANE AND PRECIPITATIONS REDISTRIBUTION

The transition of the cloud, thunderstorm state, accompanied by intensive lightning discharges, to hail occurs as a result of the nucleation and growth of hail nuclei in the drop and grain forms. The vertical extent of the volume where nuclei where are formed on giant particles of micro-impurities in the atmosphere (GPA) is between the isotherms from -2°C to -15°C . Ice crystals grow due to sublimation and coagulation with supercooled cloud droplets. In the same volume, between the isotherms of temperatures

from -5°C to -16°C , discharges occur with the maximum number of lightning strikes at an altitude of -10°C . Lightning strikes introduce an excess volume charge with an ion concentration of about 10^{17} cm^{-3} into the cloud space. Positive lightning strikes introduce positive ions, while negative lightning strikes introduce negative ions. These ions are deposited on the GPA, which have concentrations of about 10^5 m^{-3} in pre-hail situations. The negative charge on the GPA, contributing to the nucleation of planar crystal structures on the particle, forms a granular hailstone. The positive charge on the GPA forms filamentous structures that break off, and the fragments, when they come into contact with supercooled large droplets and freeze them, lead to the formation of droplet hailstones. To prevent hail, ground lightning, and the redistribution of precipitation, a method based on the initiation of artificial lightning has been proposed. Methods of initiating artificial lightning are given using: plasma formations resulting from the detonation of 100 mm artillery shells in thunderstorm-hail clouds, a special rocket with a detonating cord, and an anti-hail rocket with a conductive surface. A significant change in the radar reflectivity of a thunderstorm-hail cloud has been observed under the influence of an artificially initiated lightning strike. Based on the results of preliminary experiments on hail prevention, the method of preventing and interrupting hail using "reagent plus artificially initiated lightning" has been proposed. Given the established fact that approximately 10 tons of precipitation are released with each lightning strike, it is possible to redistribute precipitation, prevent ground-based lightning, and more using artificially initiated lightning. In the future, unmanned aerial systems should be used to initiate artificial lightning within thunderstorm clouds, while quadcopters should be used in the vicinity of clouds.

Thunderstorm; lightning; urban development; cloud; atmosphere; active influences.

Введение. Известно, что градовое состояние облака наступает как стадия развития грозового облака. Практически град выпадает из зрелых грозовых облаков, в которых происходят электрические разряды. Эти разряды могут быть видимыми или слышимыми внутриоблачными или наземными молниями. Кроме них, в облаках происходят так называемые инципиентные молнии, т.е. электрические разряды длиной меньше, чем длина ступени лидера молнии. Стримерные зоны лидеров этих типов разрядов внедряют в пространства, которые они пронизывают по пути своего распространения, избыточный объёмный заряд (ИОЗ). Заряд ионов в ИОЗ такой же что и заряд головки распространяющегося лидера: отрицательный лидер внедрит отрицательный ИОЗ, положительный – положительный ИОЗ. Концентрация ионов в ИОЗ достигает 10^{17} м^{-3} . Объём пространства, занимаемый стримерной зоной лидера, в частности, молнии определяется протяжённостью молнии, достигающей 100–300 м [1]. Поперечное сечение этого пространства можно оценить исходя из величины угла вершины, конусообразной стримерной зоны лабораторной искры порядка 15° [2]. Тогда соответствующие этим стримерам радиусы будут равны 26 и 78 м. Таким образом, объём, занимаемый ИОЗ будет представлять собой тело вращения с длиной, равной длине извилистой молнии и диаметрами от 26 до 78 м.

Цель работы – разработка методов предотвращения и прерывания града: «реагент плюс искусственно инициированные молнии», а также перераспределения осадков.

Градообразование в грозовом облаке. Процесс перехода грозового состояния облака в градовое происходит в результате зарождения и превращения в град капельных и крупяных зародышей града. Зародыши града образуются в интервале температур от -2°C до -15°C на крупных частицах микропримеси, которые условно разделены на гигантские (ГАЧ) с размерами от 1 до 30 мкм и сверхгигантские (СГАЧ) с размерами от 30 до 60 мкм и больше [3–6].

Вероятность выпадения града возрастает с увеличением ударов молний до 70 в минуту, но при большей частоте ударов града не наблюдалось [7]. В работе [8] установлено, что вероятность наступления суровых погодных условий и, следовательно, образования града очень быстро возрастала с увеличением количества положительных наземных молний. Возрастание числа положительных наземных молний сопровождалось выпадением града. По мнению авторов работы [9] в случае высокой интенсивности положительных молниевых разрядов на землю они происходят вокруг ядра с высокой отражательной способностью или во всей области с высокой отражаемостью. Общая грозовая активность имеет тенденцию приводить к неблагоприятным погодным условиям на земле или к, так называемой, суровой погоде [10], когда шторм переходит в последнюю стадию при зарождении фидерных облаков вокруг материнского облака.

О масштабах облачного пространства, занимаемого одним молниевым разрядом, можно судить по установленным в работах [11, 12] вертикальных и горизонтальных проекциях молний, которые составляют около 13-14 км.

Внутриоблачные молнии развиваются не симметрично, что объясняется различием скорости прорастания отрицательного и положительного ветвей [13]. Это приводит к большей длине отрицательной ветви (корня), чем к положительной ветви кроны. Тогда в стримерные зоны лидеров этих ветвей будут внедрять соответствующие ИОЗ с одноименными знаками. Как результат, в корне будут зарождаться капельные зародыши града, а в кроне – крупные.

Основным процессом, приводящим к зарождению зародышей града в грозовом облаке и превращающим грозовое облако в градовое мы считаем процесс зарядки ГАЧ и СГАЧ в ИОЗ, во внедряемые в облачную среду внутриоблачными и наземными молниями [11]. В работе [14] установлено, что молниевые разряды в грозовом облаке происходят между изотермами -5°C и -16°C , а большинство из них – вблизи изотермы -10°C . Концентрации ГАЧ и СГАЧ в предградовые дни могут достигать значения 10^4 – 10^5 м^{-3} [3]. Так как электрическое поле грозового облака восстанавливается за несколько секунд после молниевых разрядов, то следует учесть два механизма зарядки частиц – диффузионный и зарядка в электростатическом поле, которые даются соответственно формулами (1). В результате на ГАЧ и СГАЧ начинают оседать избыточные заряды [12]:

$$N_1 = \frac{dkT}{2e^2 \ln(1 + \pi dcNe^2t / 2kT)}, \quad (1)$$

$$N_2 = \frac{2\varepsilon_0 d^2 E}{4e(\varepsilon_0 + 2)}$$

где d – диаметр частицы, k – постоянная Больцмана, t – время зарядки, N_1 и N_2 – средние концентрации положительных и отрицательных ионов, c – среднеквадратическая скорость иона, T – температура в градусах Кельвина, ε_0 – диэлектрическая проницаемость частицы, E – напряжённость электростатического поля e – заряд электрона.

По оценочным расчётам по первой формуле из (1) на частицу с $d = 60$ мкм из-за диффузии за $t=1$ с осадут около $4,8 \cdot 10^4$ ионов, а на частицу с $d = 30$ мкм – около $2 \cdot 10^3$ ионов. Такие заряды на частицах получены при $T = 273^{\circ}\text{K}$, концентрацией $N = 10^{17} \text{ м}^{-3}$, $c = 4,5 \cdot 10^2 \text{ м/с}$. $\varepsilon_0 = 4$.

При зарядке в электростатическом поле на частицу с $d = 60$ мкм, $E = 100 \text{ кВ/м}$, на частицу осадут $1,25 \cdot 10^5$ ионов, а на частицу с $d = 30$ мкм – $3 \cdot 10^4$ ионов. При оценке роли заряженных гигантских частиц в образовании зародышей града следует учесть следующее.

Исследованиями института кристаллографии установлено, что зарождение кристаллических структур на поверхности подложек зависит от знака заряда на ней [13]. Так, только отрицательный заряд на поверхности способствует зарождению на ней плоскостных структур, которые впоследствии растут блочно, а при положительных зарядах образуются нитевидные структуры, не способные к блочному росту. С учётом этих исследований можно заключить, что на отрицательно заряженных ГАЧ и СГАЧ ледяные кристаллы, вырастая блочно, превращаются напрямую в крупные зародыши. На положительно заряженном ГАЧ и СГАЧ образуются нитевидные кристаллики, которые обламываются и их фрагменты, сталкиваясь с переохлажденными каплями, замораживают их, образуя капельные зародыши града. Часть крупных зародышей, контактируя с переохлажденными каплями, могут также образовать капельные зародыши града. Ясно, что только крупные частицы, на которых осадут большее количество ионов, превратятся в зародыши града.

Таким образом, переход грозовой стадии развития облака в градовую обусловлен формированием зародышей града на заряжающихся в ИОЗ гигантских частиц микропримеси атмосферы. Совпадением интервалов температур, где происходят молниевые разряды – между изотермами -5°C и -16°C и образование капельных и крупных зародышей града – между изотермами температур от -2°C до -15°C , можно объяснить главную роль молниевых разрядов в переходе облака из грозового состояния в градовое.

Интенсивные молниевые разряды, внедряющие в облачную среду множество ИОЗ и создающие большее количество зародышей града, способствуют перераспределению водности между ними и, тем самым, уменьшают размеры градин. В этом случае зрелое грозовое облако не переходит в градовое облако. Если же число молниевых разрядов недостаточно для образования необходимого количества зародышей града, то зародыши вырастая, превращаются в градины и наступает градовая стадия развития. В качестве предварительного условия образования града сотрудниками Высокогорного геофизического института указывается интенсивность молниевых разрядов 5 разрядов в минуту [14]. Если интенсивность меньше 5, то активное воздействие на градовое облако можно прервать. Указывая на взаимную обусловленность молниевых разрядов и града, в работе [15, 16] приводятся данные о том, что грозовое облако становится не градоопасным, если число внутриоблачных молний превышает 71 разрядов в минуту. На отсутствие града в грозо градовых облаках при интенсивностях до 10 разрядов в минуту и выше 70 разрядов в минуту указано также в этой работе. На основании двух последних работ можно предположить, что увеличением интенсивности молниевых разрядов выше 70 разрядов в минуту можно прервать град из грозоградовых облаков.

Комбинированный метод предотвращения града. Сравнение вышеуказанных величин интенсивностей представляет особый интерес в связи с предлагаемым нами комбинированным методом предотвращения града – «реагент плюс искусственно инициируемые молнии» [17]. Основанием предложенного метода являются экспериментально доказанные возможности инициирования в грозоградовых облаках искусственных молний (ИМ) с помощью: взрывающихся в них артиллерийских снарядов [18] и специальных ракет с детонирующим шнуром [19], противоградовых ракет с токопроводящей поверхностью (ТПП) [18].

В работе [19] ИМ инициировались плазменным образованием, получающимся при подрыве в грозоградовом облаке 100мм артиллерийских снарядов. Было установлено, что после инициированной молнии происходит колоссальное изменение радиолокационной отражаемости (РЛО) облака, выражающееся в снижении высоты с максимальной РЛО: высота отражаемости 35 dBZ снизилась примерно на 1км, а 45 dBZ на 2км (рис. 1). Подрыв снаряда производился на высоте около 7 км между верхним + Р зарядом и нижним – N зарядом. Возникновение ИМ идентифицировалось четырьмя приборами: микрофоном, фиксировавшим запуск снаряда и его подрыв в облаке по звуку; радиоприёмником, регистрировавшим атмосферик от ИМ; флюксометром, регистрировавшим изменения напряжённости поля у поверхности земли; измерителем тока короны с заземлённого острия. По совпадению изменений во времени измеряемых параметров, идентифицировался факт инициирования ИМ. В результате в этих экспериментах 11 снарядами были инициированы 9 ИМ.

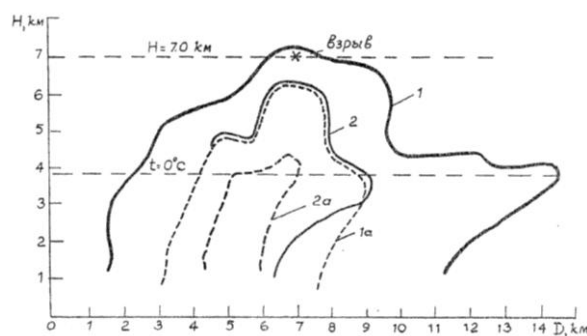


Рис. 1. Картина изменения РЛО под действием инициированной ИМ: кривые 1 и 2 соответствуют отражаемостям 35 dBZ и 45 dBZ, 1а и 2а через 2÷3 мин после инициирования ИМ

Как видно, кроме снижения высот РЛО, площади обоих контуров РЛО уменьшаются значительно.

Применимость ракет с ТПП для инициирования ИМ, а также их влияние на РЛО грозового облака экспериментально исследованы в работе [18, 20]. Выбор объектов воздействия проводился по высоте отражаемости 45 dBZ. Изменения РЛО в этих экспериментах показаны на рис. 2 и 3.

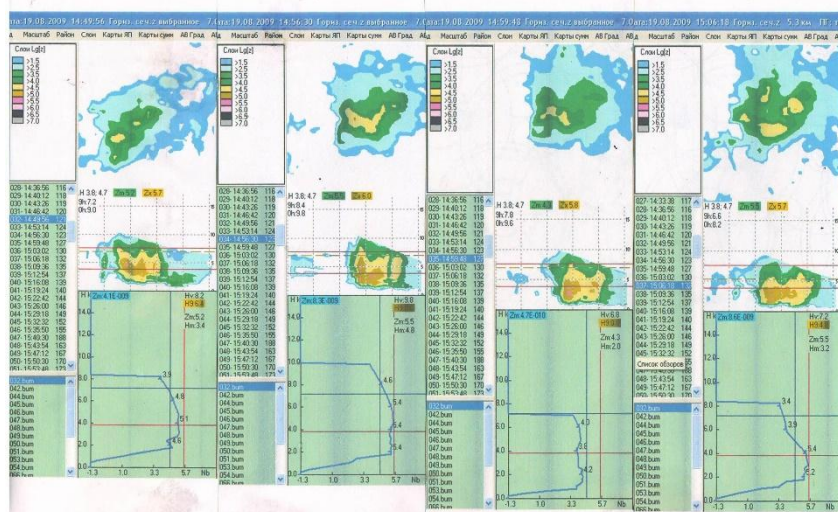


Рис. 2. Изменения РЛО через 3-4 мин после инициирования ИМ 3 противоградовыми ракетами с ТПП

При угле запуска около 60 градусов ракета входит в грозовое облако и находится там около 30–35 с. Тогда по траектории полёта ракеты и при её самоликвидации, она может инициировать ИМ много раз. Трудность их идентификации связана с неизвестностью траекторий иницируемых молний. Отмечаемые грозопеленгатором LG 8000 последовательные по времени молнии могут быть в указанном промежутке времени нахождения ракеты в облаке связаны именно с иницированием ИМ последней. На начальном участке полёта ракеты до 2,2 км, т.е. когда она не «вошла» в зону, где напряжённости поля от грозового облака велики, не должны появляться ИМ. Запуски противоградовых ракет с ТПП производились с ПВ 11 вблизи г. Баксан. Были запущены 3 ракеты с 15:55:00 до 15:57:53. Картины РЛО на рисунке по времени запаздывают на 1 час. В таблице цифрами в скобках отмечены количество точек, соединением которых получены траектории, иницированных ИМ.

Таблица

Пространственно-временные характеристики молний

Время, мин, с	Длина, км	Время, мин, с	Длина, км	Время, мин, с	Длина, км
55:00		56:00(2)	2,9	57:00	
55:15(2)	1,41	56:06(2)	1,78	57:02(7)	7,94
55:21(3)	3,96	56:18		57:06	
55:26		56:23		57:09	
55:28(7)	8,5	56:30(4)	4,05	57:27	
55:32		56:32		57:34	
55:35(5)	10,2	56:37(5)	9,5	57:37(2)	3,3
55:51		56:38(3)	7,8	57:44	
		56:44		57:48	
				57:53(2)	2,23

Первая ракета инициирует молнии по всей траектории полёта от объёмных зарядов, существующих по траектории полёта. Вторая ракета, двигаясь в среде с пониженной напряжённостью поля, начинает инициировать молнии на более дальних от ПВ расстояниях, так как электростатическое поле облака не успевает восстановиться ко времени появления ракеты в этой области. Третья ракета также инициирует ИМ на более дальних расстояниях. Последняя молния, скорее всего, инициирована ракетой при её самоликвидации, т.е. от плазменного образования – продуктов детонации. Таким образом, в объёме грозоградового облака с горизонтальной проекцией около 18 км^2 тремя ракетами инициированы 12 внутриоблачных (ВО) ИМ, с суммарной длиной около 57 км. Первая ракета за 51 с инициировала 4 ИМ длиной 21,4 км, вторая ракета за 44 с – 5 ИМ длиной 26 км и третья за 53 с – 3 ИМ длиной 8,2 км. За эти же 3 мин зарегистрированы 15 молний облака-земля (О-З) на этой же площади. Траектории ИМ восстанавливались соединением точек, где одновременно по секундам проявлялись составляющие молний в различных точках на горизонтальной проекции объёма облака. Такой подход к определению траекторий ИМ, по-видимому, правомочен, так как в основе разработки грозопеленгатора лежат исследования конца семидесятых годов, когда разделение молний О-З и ВО молний проводилось сравнением мощностей электромагнитного излучения молний-атмосфериков, получаемых горизонтальным и вертикальным антеннами. В частности, мы полагаем, что проявляемые одновременно в различных точках элементы извилистой молнии являются вертикальными участками траектории одной и той же молнии.

Некоторым подтверждением такого подхода является время восстановления электростатического поля грозоградового облака для последующей молнии, достигающей несколько секунд. Изменение картины РЛЮ со временем под действием искусственно инициированных 12 внутриоблачных молний и 15 наземных молний приведены на рис. 2.

Объём облачного пространства, в который происходит внедрение ИОЗ, состоит из объёма, охватываемого стримерными зонами инициированных ИМ и, объёма, занимаемого кроной наземных молний (О-З). Принято считать, что $2/3$ длины молнии О-З находится внутри грозового облака. Тогда для средней молнии, длиной около 9 км получим, что крона «проникает» в облако на 6 км. Объём облака, разряжающаяся одной наземной молнией, может достигать кубические километры. Считая, что за 3 мин общая длина проникающих в облако частей молний составляет 90 км, получим длину всех молний около 150 км. Для оценки объёма облака, в котором происходит внедрение ИОЗ стримерной зоной лидера молнии (СЗЛМ) можно выбрать средний радиус стримерной зоны, равный около 45 м. Тогда получим приблизительный объём облачного пространства, в котором внедряется ИОЗ, равный $6,76 \cdot 10^8 \text{ м}^3$. Выше отмечено, что в штормовые дни концентрации ГАЧ и СГАЧ в грозовых облаках могут достигать 10^4 м^{-3} и более. Допуская, что в процессе осаднения на эти частицы ионов из ИОЗ они становятся источниками капельных и крупяных зародышей града, получим около $7 \cdot 10^{12}$ шт. и, которые конкурируя с естественными зародышами града за облачную водность, предотвратят град. Более мелкие частицы ЧМА также осаждают на себя ионы из ИОЗ. В зависимости от знака заряда и окружающей температуры они становятся облачными ядрами конденсации или сублимации. Свободные отрицательные ионы становятся ядрами конденсации, притягивая дипольные молекулы водяного пара, что приведёт к «высушиванию» объёма с ИОЗ и образованию облачных капель. Второй опыт по воздействию на градоопасное облако инициированием ИМ приведён на рис. 3 [18].

Ниже приведен фрагмент регистрации ИМ и молний О-З грозопеленгатором (рис. 4). Масштаб фрагмента 1 см: 1 км. В этом опыте двумя ракетами за 1 мин 43 с инициированы 4 ИМ, протяжённостью 9,8 км. За это же время зафиксированы 11 молний О-З суммарной протяжённостью около 65 км.

Объём пространства, в которое внедряется ИОЗ, так же, как и выше, можно оценить исходя из суммарной протяжённости обеих молний и радиуса СЗЛМ, равной около 45 м, т.е. около $1,2 \cdot 10^8 \text{ м}^3$. Зарядка в этом объёме частиц ГАЧ и СГАЧ с концентрацией 10^4 м^{-3} , приводя к образованию капельных и крупяных зародышей града, способствует освоению капельной водности, и тем самым, сокращает время жизни облака. Как видно, из измене-

ний картин РЛО облака под действием наземных и ИМ, грозоградное облако осаждается приблизительно за 20 мин, не переходя в градовое облако. По сравнению с используемым в противоградовой защите методом получения конкурирующих зародышей града внесением частиц кристаллизующего реагента (ЧКР) с размерами до 1 мкм, зарядившиеся в ИОЗ ГАЧ и СГАЧ с размерами от 30 мкм до 60 мкм и более, становятся источниками двух типов зародышей града, минуя время роста кристалла до 40-50 мкм [20]. Время образования зародышей града на ГАЧ и СГАЧ около 3-4 мин, поэтому реакция облака на процессы осадкообразования может колебаться между 4 и 10 мин. Опыт по применению предложенного метода предотвращения града «реагент плюс искусственно инициированная молния» провели 1 сентября 2021г. [17].

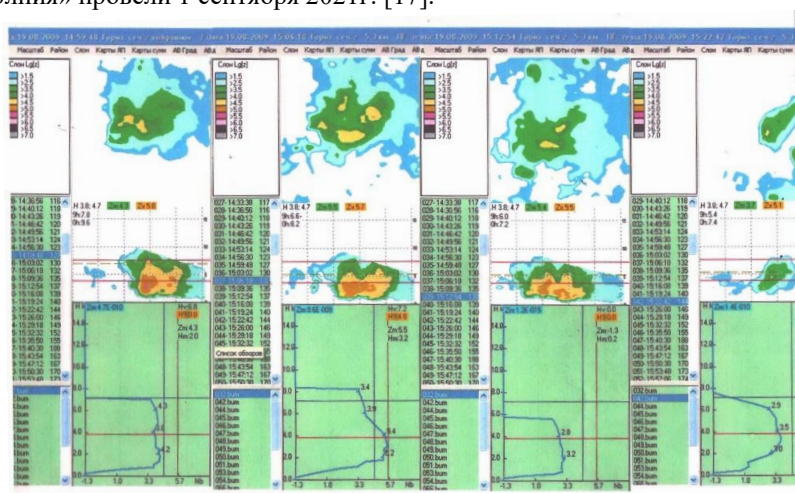


Рис. 3. АВ, проведенное на конвективную облачность инициированием ИМ

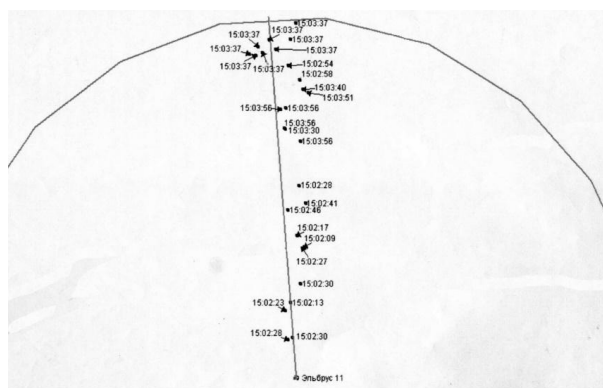


Рис. 4. Фрагмент регистрации грозопеленгатором LG8000 инициированных двумя ракетами ИМ и молний О-3

В день с воздействием был выбран объект воздействия (ОВ) 3 категории №5, который зародился в 15:20 вблизи селения Кёнделен и начал смещаться с скоростью около 20 км/час на защищаемую территорию – ЗТ КБР. При тенденции развития облака к усилению и росту града в нём, с 15:29 по 15:33 с пунктов воздействия (ПВ) вблизи селений Заюково и Верхний Куркужин тройками были запущены по 6 ракет с ТПП и 6 стандартных ракет с ПВ вблизи с. Кёнделен. Воздействие на облако №5 было продолжено с указанных ПВ с 15:35 по 15:42 запусками по 6 ракет с ТПП. В результате воздействия из ОВ №5 выпадали только ливневные осадки, В 15:47 это облако начало регенерировать и в районе г. Баксан появились новые гроззовые ячейки. В 16:02-16:04 с ПВ вблизи селения Кызбурун-2 была обработана передняя фронтальная область навеса радиоэха 6 ракетами с ТПП. Вблизи

облака №5 на высоте около 8 км в 16:05 локализовалась новая градоопасная ячейка №6. На этот ОВ 3 категории было запущено 6 ракет с ТПП с ПВ Кызбуруна-2. Всего на облако №6 израсходованы 10 ракет с ТПП. Выпадение града из ОВ 3 категории №5 и №6 не зафиксированы и это свидетельствует о том, что применение предлагаемого метода предотвращения града является обнадёживающим методом борьбы с градобитиями.

Различить возможные естественные молнии и искусственные на картине не представляется возможным. Однако дополнительный анализ времён запуска ракет и появления молний позволяет идентифицировать искусственные молнии (ИМ). Кроме того, кучность точек на северо-западном направлении от ПВ вблизи с. Заюково и в северном направлении вблизи с. Верхний Куркужин свидетельствует об инициировании ракетами с ТПП искусственных молний. Отсутствие такой кучности точек вблизи ПВ вблизи с. Кызбурун-2 на рис. 5, с которого также были запущены 12 ракет с ТПП, возможно, объясняется близостью датчиков грозопеленгатора, что может повлиять на идентификацию ИМ.

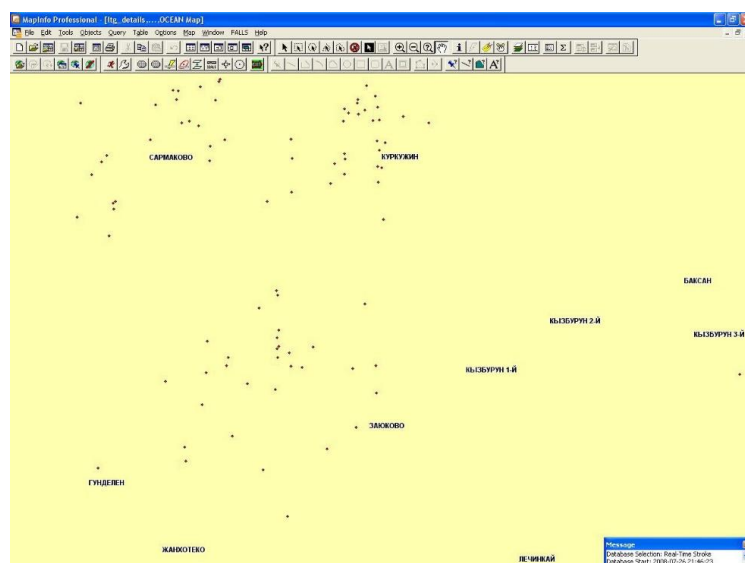


Рис. 5. Точки между селениями Кёндеден и Заюково и в районе селения Верхний Куркужин, соответствуют инициированным ИМ

По руководящему документу на грозоградовое облако с 13 ОВ 3 категории и 2 ОВ 2 категории с целью предотвращения града по принятому в противоградовых службах документам следовало бы потратить около 300 штатных противоградовых ракет.

Применение 35 противоградовых ракет с ТПП позволило сократить число ракет до 140 шт. Многочисленные ИМ, инициированные ракетами с ТПП, координаты которых фиксировались грозопеленгатором LG8000, способствовали предотвращению или уменьшению размеров града.

Заключение. Таким образом, физической основой предотвращения града с помощью комбинированного метода «реагент плюс искусственно инициируемые молнии» являются:

1. Создание конкурирующих с естественными зародышами града искусственных зародышей града:

а) на ГАЧ и СГАЧ заряжающихся в избыточном объёмном заряде, внедряемом в облачное пространство искусственно ИМ;

б) на частицах кристаллизующего реагента, вносимых в облако штатной противоградовой ракетой с ТПП.

2. Обрушение зон аккумуляции, образующихся между верхним + Р и нижним – N-зарядами, а также обрушение нижних +р зарядов, зависающих в электростатическом поле нижнего – N-заряда, которые происходит из-за ослабления электростатического поля искусственно инициируемыми молниями.

Кроме того, ракетами с ТПП, беспилотниками с проводниками в хвостовой части длиной в десятки метров можно защитить от молний и града различные объекты на земле: пусковые установки космических ракет, атомные электростанции, крупные электро-распределительные подстанции, ценные лесонасаждения, ограниченные площади – сады, табак, виноградники, теплицы и т.д. Так как инициируя молнии можно обрушить зоны аккумуляции – электроактивные зоны, в которых происходит рост частиц осадков в том числе и града, то искусственными молниями можно перераспределить осадки (по данным работ [18] каждый молниевый разряд приводит к выпадению 10^4 м^3 осадков).

Авторы выражают признательность профессору Куповых Г.В. за участие в обсуждении результатов исследований и подготовку статьи к публикации.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Базелян Э.М., Горин Б.Н., Левитов В.И. Физические и инженерные основы молниезащиты. – М.: Гидрометеиздат. 1978. – 202 с.
2. Петров Н.И., Петрова Г.Н. Физические механизмы развития молниевых разрядов между грозовым облаком и ионосферой // ЖТФ. – 1999. – Т. 69. – Вып. 4. – С. 134-137.
3. Хоргуани В.Г. Микрофизика зарождения и роста града. – М.: Гидрометеиздат, 1984. – 184 с.
4. Blevins L.L., Marwitz J.D. Visual observations of lightning in Some Great Plains hailstorms // Weather. – 1968. – 23. – P. 192-194.
5. Branick M.L., and Deswell C.A. III. An observation of the relationship between supercell structure and lightning ground strike polarity // Weather Forecast. – 1992. – 7. – P. 143-149.
6. Carey L.D., and Rutledge S.A. Electrical and multiparameter Radar observations of a severe hailstorm // J. Geophys. Res. – 1998. – 103. – P. 13,979-14,000.
7. Williams E.R., et al. The behavior of total lightning activity in Severe Florida thunderstorms // J Atmos. Res. – 1999. – 51. – P. 245-264.
8. Proctor D.E. Lighting and precipitation in Small Multicellular Thunderstorm // J. Geophys. Res. – 1983. – Vol. 88, n. C9. – P. 5421-5440.
9. Proctor D.E. Radar Observations of Lightning // J. Geophys. Res. – 1981. – Vol. 86, C12. – P. 12109-12114.
10. Iudin D.I., Rakov V.A., Syssoev A.A., Bulatov A.A., M. Hayakawa M. Formation of decimeter-scale, long-lived elevated ionic conductivity regions in thunderclouds // Climate and Atmospheric Science. – 2019. – P. 1-10.
11. Бейтуганов М.Н. Электрические разряды различных типов и молний как причина превращения частиц микропримеси атмосферы в зародыши града // Доклады Всероссийской конференции по физике облаков и активным воздействиям на гидрометеорологические процессы: Сб. научных трудов. Ч. 2. – Нальчик, 2017. – С. 249-256.
12. Райст П. Аэрозоли. Введение в теорию. – М.: Мир, 1987. – 277 с.
13. Дистлер Г.И. Кобзрева С.А. Дальнодействие активных центров кристаллических поверхностей // ДАН СССР. – 1967. – Т. 172, № 5. – С. 1069-1071.
14. Медалиев Х.Х., Сижажев С.М. Изучение степени грозоопасности облаков и определении времени и места возникновения канала молнии // Тр. Высокотгорного геофизического института. Вып. 29. Физика облаков и активных воздействий на градовые процессы. – Л.: Гидрометеиздат, 1975. – С. 3-11.
15. Soula S., Seity Y., F  ral L., Sauvageot H. Cloud-to-ground lightning activity in hail-bearing convective storms // Journal of Geophysical Research. – 2004. – 109 (D2). – P. D 2101.
16. Бейтуганов М.Н., Чочаев Х.Х. Предотвращение и прерывание града комбинированным методом «реагент плюс искусственно инициируемые молнии» // Тр. Военно-космической академии им. А.Ф. Можайского. Проблемы военно-прикладной геофизики и контроля состояния природной среды. – СПб., 2022. – Вып. 685. – С 19-27.
17. Бейтуганов М.Н. Метод предотвращения града инициированием искусственных молний // Сб. научных трудов 5 Российской конференции по атмосферному электричеству. – Владимир, 2003. – Т. 1. – С. 204-207.
18. Бейтуганов М.Н., Чочаев Х.Х. Опыты по предотвращению града инициированием искусственных молний с помощью противоградовых ракет с токопроводящей поверхностью // VII Всероссийская конференция по атмосферному электричеству: Сб. трудов. Нальчик 23-27 сентября 2019. – СПб., 2019. – С. 63-72.

19. Бейтуганов М.Н. Несмеянов П.А. Исследование возможности провоцирования электроразрядных явлений плазменными образованиями // Сб. трудов 6 Российской конференции по атмосферному электричеству. – Нижний Новгород, 2007. – С. 142-144.
20. Бейтуганов М.Н. Инициирование искусственных молний как метод предотвращения града // Известия Кабардино-Балкарского научного центра РАН. – 2013. – № 1 (51). – С. 38-43.

REFERENCES

1. Bazelyan E.M., Gorin B.N., Levitov V.I. Fizicheskie i inzhenernye osnovy molniezashchity [Physical and engineering fundamentals of lightning protection]. Moscow: Gidrometeoizdat, 1978, 202 p.
2. Petrov N.I., Petrova G.N. Fizicheskie mekhanizmy razvitiya molnievykh razryadov mezhdu grozovym oblakom i ionosferoy [Physical mechanisms of lightning discharges development between a thundercloud and the ionosphere], *ZhTF* [Journal of Technical Physics], 1999, Vol. 69, Issue 4, pp. 134-137.
3. Khorquani V.G. Mikrofizika zarozhdeniya i rosta grada [Microphysics of hail origin and growth]. Moscow: Gidrometeoizdat, 1984, 184 p.
4. Blevins L.L., Marwitz J.D. Visual observations of lightning in Some Great Plains hailstorms, *Weather*, 1968, 23, pp. 192-194.
5. Branick M.L., and Deswell C.A. III. An observation of the relationship between supercell structure and lightning ground strike polarity, *Weather Forecast*, 1992, 7, pp. 143-149.
6. Carey L.D., and Rutledge S.A. Electrical and multiparameter Radar observations of a severe hailstorm, *J. Geophys. Res.*, 1998, 103, pp. 13,979-14,000.
7. Williams E.R., et al. The behavior of total lightning activity in Severe Florida thunderstorms, *J Atmos. Res.*, 1999, 51, pp. 245-264.
8. Proctor D.E. Lightning and precipitation in Small Multicellular Thunderstorm, *J. Geophys. Res.*, 1983, Vol. 88, n. C9, pp. 5421-5440.
9. Proctor D.E. Radar Observations of Lightning, *J. Geophys. Res.*, 1981, Vol. 86, C12, pp. 12109-12114.
10. Iudin D.I., Rakov V.A., Syssoev A.A., Bulatov A.A., M. Hayakawa M. Formation of decimeter-scale, long-lived elevated ionic conductivity regions in thunderclouds, *Climate and Atmospheric Science*, 2019, pp. 1-10.
11. Beytuganov M.N. Elektricheskie razryady razlichnykh tipov i molnii kak prichina prevrashcheniya chastits mikroprimesi atmosfery v zarodyshi grada [Electrical discharges of various types and lightning as the cause of the transformation of atmospheric microparticles into hail nuclei], *Doklady Vserossiyskoy konferentsii po fizike oblakov i aktivnym vozdeystviyam na gidrometeorologicheskie protsessy: Sb. nauchnykh trudov* [Reports of the All-Russian Conference on Cloud Physics and Active Influences on Hydrometeorological Processes. Collection of Scientific Papers]. Part 2. Nal'chik, 2017, pp. 249-256.
12. Rayst P. Aerozoli. Vvedenie v teoriyu [Aerosols. Introduction to the theory]. Moscow: Mir, 1987, 277 p.
13. Distler G.I. Kobzeva S.A. Dal'nodeystvie aktivnykh tsentrov kristallicheskikh poverkhnostey [Long-Range action of active centers of crystalline surfaces], *DAN SSSR* [Proceedings of the Academy of Sciences of USSR], 1967, Vol. 172, No. 5, pp. 1069-1071.
14. Meddaliev Kh.Kh., Sizhazhev S.M. Izucheniye stepeni grozopasnosti oblakov i opredelenii vremeni i mesta vozniknoveniya kanala molnii [Study of the degree of thunderstorm danger of clouds and determination of the time and place of the lightning channel occurrence], *Tr. Vysokoqornogo geofizicheskogo instituta. Vyp. 29. Fizika oblakov i aktivnykh vozdeystviy na gradovye protsessy* [Proceedings of the High-Mountain Geophysical Institute. Issue 29. Physics of clouds and active influences on hail processes]. Leningrad: Gidrometeoizdat, 1975, pp. 3-11.
15. Soula S., Seity Y., Féral L., Sauvageot H. Cloud-to-ground lightning activity in hail-bearing convective storms, *Journal of Geophysical Research*, 2004, 109 (D2), pp. D 2101.
16. Beytuganov M.N., Chochaev Kh.Kh. Predotvrashchenie i preryvanie grada kombinirovannym metodom «reagent plus искусственно иницируемые молнии» [Prevention and interruption of hail by the combined method of "reagent plus artificially initiated lightning"], *Tr. Voenno-kosmicheskoy akademii im. A.F. Mozhayskogo. Problemy voenno-prikladnoy geofiziki i kontrolya sostoyaniya prirodnoy sredy* [Proceedings of the Military Space Academy named after A.F. Mozhaysky. Problems of Military Applied Geophysics and Control of the Natural Environment]. St. Petersburg, 2022, Issue 685, pp 19-27.
17. Beytuganov M.N. Metod predotvrashcheniya grada initsirovaniem iskusstvennykh molnii [The method of preventing hail by initiating artificial lightning], *Sb. nauchnykh trudov 5 Rossiyskoy konferentsii po atmosfernomu elektrichestvu* [Collection of scientific papers of the 5th Russian Conference on Atmospheric Electricity]. Vladimir, 2003, Vol. 1, pp. 204-207.
18. Beytuganov M.N., Chochaev Kh.Kh. Opyty po predotvrashcheniyu grada initsirovaniem iskusstvennykh molnii s pomoshch'yu protivogradovykh raket s tokoprovodyashchey poverkhnost'yu [Experiments on preventing hail by initiating artificial lightning using anti-hail rockets with a conductive surface], *VII Vserossiyskaya konferentsiya po atmosfernomu elektrichestvu: Sb. trudov. Nal'chik 23-27 sentyabrya 2019* [VII All-Russian Conference on Atmospheric Electricity. Collection of papers. Nalchik, September 23-27, 2019]. St. Petersburg, 2019, pp. 63-72.

19. *Beytuganov M.N. Nesmeyanov P.A. Issledovanie vozmozhnosti provotsirovaniya elektrorazryadnykh yavleniy plazmennymi obrazovaniyami* [Research on the possibility of provoking electric discharge phenomena by plasma formations], *Sb. trudov 6 Rossiyskoy konferentsii po atmosfernomu elektrichestvu* [Collection of papers of the 6th Russian Conference on Atmospheric Electricity]. Nizhniy Novgorod, 2007, pp. 142-144.
20. *Beytuganov M.N. Initsirovanie iskusstvennykh molniy kak metod predotvrashcheniya grada* [Initiation of artificial lightning as a method of preventing hail], *Izvestiya Kabardino-Balkarskogo nauchnogo tsentra RAN* [Izvestiya Kabardino-Balkarskogo Scientific Center of the Russian Academy of Sciences], 2013, No. 1 (51), pp. 38-43.

Бейтуганов Мусаби Ногманович – Северо-Кавказская военизированная служба, Росгидромет; e-mail: bmn3112m@mail.ru; г. Нальчик, Россия; тел.: +79289104874; начальник отдела активных воздействий СКВС; д.ф.-м.н.

Чочаев Хизир Хусейнович – Северо-Кавказская военизированная служба, Росгидромет; e-mail: gradskvs@rambler.ru; г. Нальчик, Россия; тел.: +79280819577; начальник Службы; к.ф.-м.н.

Beituganov Musabi Nogmanovich – North Caucasian Paramilitary Service, Roshydromet; e-mail: bmn3112m@mail.ru; Nalchik, Russia; phone: +79289104874; head of the Department of Active Effects of the CFS; dr. of phys. and math. sc.

Chochaev Khizir Huseynovich – North Caucasian Paramilitary Service, Roshydromet; e-mail: gradskvs@rambler.ru; Nalchik, Russia; phone: +79280819577; head of Service; cand. of phys. and math. sc.

УДК 551.594.1

DOI 10.18522/2311-3103-2026-4-180-190

О.В. Белоусова

ГЛОБАЛЬНЫЕ И ЛОКАЛЬНЫЕ ВАРИАЦИИ ЭЛЕКТРИЧЕСКОГО ПОЛЯ В ПРИЗЕМНОМ ЭЛЕКТРОДНОМ СЛОЕ

Представлена математическая модель структуры приземного электродного слоя атмосферы, учитывающая совместное действие глобальных и локальных генераторов электрического поля. Пространственно-временное моделирование основано на сопряжении гармонических тригонометрических функций суточного хода для локального турбулентного обмена и для плотности глобального полного тока. При моделировании поведения электрического поля использованы аналитические решения уравнений электродного эффекта в атмосфере. Предложенный подход заключается в пошаговой подстановке мгновенных значений плотности полного электрического тока и коэффициента турбулентной диффузии, полученных из уравнений их периодического изменения для конкретного часа суток, в аналитические формулы стационарного пространственного распределения характеристик турбулентного электродного слоя. Правомочность использования квазистационарного приближения строго обоснована существенным различием (более чем на два порядка) временных масштабов установления электрического равновесия среды и периода суточного хода глобального токового генератора. Установлено, что совмещенный режим моделирования адекватно воспроизводит синергетическое расширение размаха суточной кривой напряжённости (градиента потенциала) электрического поля и деформацию его профиля, характерную для летнего сезона в высокогорье. Достоверность полученных теоретических результатов подтверждена путем их сравнения с экспериментальными данными натурных измерений на высокогорной станции Пик Чегет (43° 16' с. ш. и 42° 30' в. д.), расположенной в Приэльбрусье на высоте 3040 м н. у. м. Анализ продемонстрировал хорошее количественное и хронологическое соответствие значений расчетных экстремумов наблюдаемой суточной вариации напряжённости электрического поля в летний сезон. Полученные результаты могут быть непосредственно использованы для повышения точности интерпретации данных геофизического мониторинга.

Математическое моделирование; приземный слой, атмосфера; электродный эффект; электрическое поле; аэроионы, ионизация; суточный ход.

O.V. Belousova

GLOBAL AND LOCAL VARIATIONS OF THE ELECTRIC FIELD
IN THE SURFACE ELECTRODE LAYER

The paper presents a mathematical model of the atmospheric surface electrode layer structure, taking into account the combined action of global and local electric field generators. Spatiotemporal modeling is based on the conjugation of harmonic trigonometric functions of the diurnal cycle for local turbulent exchange and for the global total current density. When modeling the behavior of the electric field, analytical solutions of the electrode effect equations in the atmosphere are used. The proposed approach consists of a step-by-step substitution of the total electric current density and the turbulent diffusion coefficient instantaneous values, obtained from the equations of their periodic variation for a specific hour of the day, into analytical formulas for the stationary spatial distribution of the turbulent electrode layer characteristics. The validity of using the quasi-stationary approximation is strictly substantiated by the significant difference (by more than two orders of magnitude) between the time scales of establishing electrical equilibrium in the medium and the period of the global current generator diurnal variation. It was established that the combined modeling mode adequately reproduces the synergistic expansion of the electric field strength (potential gradient) diurnal curve range and the deformation of its profile, characteristic of the summer season in the high mountains. The reliability of the obtained theoretical results was confirmed by their comparison with experimental data from in-situ measurements at the alpine station of Peak Cheget (430 16' N and 420 30' E), located in the Elbrus region at altitudes of 3040 m above sea level. The analysis demonstrated good quantitative and chronological agreement between the calculated extremes values and the observed diurnal variations in electric field during the summer season. The results obtained can be directly used to improve the accuracy of geophysical monitoring data interpretation.

Mathematical modeling; surface layer; atmosphere; electrode effect; electric field; aeroions; ionization; diurnal variation.

Введение. Исследование электродинамических процессов в приземном слое является одной из важных и практически значимых проблем физики атмосферы. Пространственно-временная структура электрического поля определяются влиянием глобальной электрической цепи (ГЭЦ) и электродным эффектом, интенсивность влияния которого зависит от метеорологических условий, в частности от степени турбулентного перемешивания [1–8]. Традиционные подходы к моделированию обычно рассматривают коэффициент турбулентной диффузии как стационарную функцию высоты, что упрощает вычисления, но не отражает реальную суточную динамику. Интенсивность турбулентного обмена претерпевает существенные изменения в течение суток под воздействием солнечной радиации и термической конвекции.

Аналитическое и численное моделирование электродинамических процессов, формирование пространственной структуры приземного электродного слоя в различных геофизических условиях выполнены в работах [9–13]. Экспериментальные [14–17] и теоретические [18–23] исследования структуры приземного электродного слоя формируют базу для математического моделирования. В работах [15, 16] используются методы численного моделирования электродинамических процессов в рамках теории электродного эффекта, а в работе [17] сочетаются методы комплексного математического моделирования и сопоставления с экспериментальными данными. Анализ суточного хода напряженности электрического поля атмосферы, представленный в работах [20, 21, 23], выполнен с использованием нестационарных моделей приземного электродного слоя с учетом действия глобальных и локальных факторов. Это позволяет объяснить наблюдаемые на практике аномалии в суточном ходе напряженности электрического поля, которые невозможно обосновать исключительно действием глобального генератора. В частности, наложение динамически меняющегося локального турбулентного переноса на унитарную вариацию приводит к деформации глобальной суточной вариации (кривой Карнеги) электрического поля – смещению временных экстремумов и генерации дополнительных пиков [8]. Математическое описание этих эффектов, подкрепленное натурными экспериментами, необходимо для проверки теоретических моделей и повышения точности интерпретации данных геофизического мониторинга.

В связи с этим актуальной задачей является детальное исследование влияния нестационарных турбулентных процессов на электрические характеристики приземного слоя.

В данной работе использован комбинированный подход к описанию электродинамической структуры приземного слоя атмосферы, синтезирующий нестационарные эмпирические закономерности и аналитические модели [18, 19].

Моделирование временного хода глобального токового генератора и суточный ход интенсивности турбулентного перемешивания реализуется аналогично подходу, примененному в работе [16], в виде тригонометрических функций, отражающих фазоамплитудные параметры соответствующих периодических изменений. Высотные профили электрических характеристик рассчитываются по полученным ранее аналитическим решениям математической модели стационарного турбулентного электродного эффекта, предложенной в работе [18], где в качестве базовой системы используются линеаризованные ионизационно-рекомбинационные уравнения и закон Ома.

Математическая модель. В условиях чистой атмосферы суточный ход электрического поля в приземном слое определяются, с одной стороны унитарной вариацией потенциала ионосферы, связанной в основном с колебаниями мощности глобального грозового генератора [1, 2, 5, 7]. С другой стороны, на величину и динамику электрических характеристик атмосферы вблизи поверхности земли влияет действие электродного эффекта, зависящего от метеорологических условий [5–8].

Математическое обоснование подстановки нестационарных временных зависимостей в аналитические профили стационарной модели базируется на квазистационарном приближении. Физическая правомерность такого шага обусловлена существенным различием временных масштабов протекающих процессов: время электрической релаксации приземной атмосферы (характерное время установления электрического равновесия в электродном слое) составляет порядка 100 секунд [6], а период суточных изменений интенсивности турбулентного обмена и глобального токового генератора (кривая Карнеги) составляет 24 часа. Поскольку временной масштаб изменения внешних факторов превышает на порядки время внутренней электродинамической релаксации среды, электрическая структура приземного слоя успевает подстраиваться под текущие значения турбулентности и полного тока.

Уравнение полного тока, описывающее суточный ход напряженности электрического поля в условиях совместного действия глобального и локального (турбулентного) токовых генераторов в приземном слое, имеет следующий вид [16, 17]:

$$\varepsilon_0 D(z, t) \frac{d^2 E}{dz^2} + \lambda(z) E = j(t), \quad (1)$$

где D – коэффициент турбулентной диффузии; λ – удельная электрическая проводимость атмосферы; j – плотность полного электрического тока; z – высота от земной поверхности, t – время; ε_0 – электрическая постоянная.

Следовательно, пространственное распределение напряженности электрического поля, концентраций аэроионов и объемного заряда в каждый конкретный момент времени с высокой точностью описывается аналитическими решениями стационарной модели, в которых параметры $j(t)$ и $D(t)$ выступают в качестве медленно меняющихся во времени квазипостоянных функций.

Функции изменения по времени плотности полного электрического тока (в соответствии с глобальной унитарной вариацией) и коэффициента турбулентного переноса (в слое 0,5–2 м) задавались следующим образом [16, 17]:

$$j(t) = j_0(1 - \alpha \sin(\omega t)), \quad (2)$$

$$D(t) = D_0 - 0,028 \sin(\omega(t - 3)) - 0,043 \cos(\omega(t - 3)), \quad (3)$$

где $\alpha = 0,2$ – безразмерный множитель, $t - BCB$ (всемирное скоординированное время по Гринвичу).

Нестационарная модель для определения суточных вариаций напряженности электрического поля в турбулентном приземном слое в условиях чистой атмосферы как следствие действия глобальных (2) и локальных (3) факторов, приведена в работах [16, 17]. Установлено, что в суточном ходе электрического поля в турбулентном приземном слое происходит запаздывание по времени (на 2–3 ч) наступления утреннего минимума и вечернего максимума по сравнению с унитарной вариацией [8]. При этом абсолютные величины амплитудных значений напряженности электрического поля относительно среднего значения возрастают от 20% до 40% при усилении турбулентного перемешивания [16, 17].

Для проведения моделирования используем решение уравнения (1) в квазистационарном виде для напряженности электрического поля из работы [18] с учетом выражений (2) и (3) для разных комбинаций зависимостей $D(t)$ и $j(t)$:

$$E(z, t) = \left(E_0 - \frac{j(t)}{\lambda_\infty} \right) e^{-\frac{(z-z_0)}{\sqrt{\frac{\lambda}{\varepsilon_0 D(t)}}}} + \frac{j(t)}{\lambda_\infty}, \quad (4)$$

здесь E_0 значение напряженности электрического поля у поверхности земли, λ_∞ асимптотическое значение удельной электрической проводимости. Зависимости $D(t)$ и $j(t)$ представим в виде набора дискретных значений с шагом 1 час.

Рассмотрим следующие режимы моделирования:

(а) в (4) плотность полного электрического тока принимается величиной постоянной в каждый момент времени $j(t) = \text{const} = j_0$ при переменном коэффициенте турбулентной диффузии $D(t)$;

(б) в (4) коэффициент турбулентной диффузии принимается постоянным $D(t) = \text{const} = D_0$ при переменной плотности тока $j(t)$;

(в) в общем случае выражения для $D(t)$ и $j(t)$ сохраняют исходный вид (2) и (3) в выражении (4).

Проведем расчеты электрического поля, применяя выражения (4), (5) и (6) с заданными параметрами, характерными для высокогорных условий [8]:

$$b_1 = 1,4 \cdot 10^{-4} \text{ м}^2 \text{ с}^{-1} \text{ В}^{-1}, \quad b_2 = 1,4 \cdot 10^{-4} \text{ м}^2 \text{ с}^{-1} \text{ В}^{-1}, \quad \alpha = 1,6 \cdot 10^{-12} \text{ м}^3 \text{ с}^{-1}, \quad e = 1,6 \cdot 10^{-19} \text{ Кл}, \\ q = 20 \cdot 10^6 \text{ м}^{-3} \text{ с}^{-1}, \quad E_0 = -500 \text{ В} \cdot \text{м}^{-1}, \quad j_0 = 22 \cdot 10^{-12} \text{ А} \cdot \text{м}^{-2}, \quad \varepsilon_0 = 8,85 \cdot 10^{-12} \text{ Ф} \cdot \text{м}^{-1}, \\ \lambda_\infty = 15,8 \cdot 10^{-14} \text{ См} \cdot \text{м}^{-1}, \quad D_0 = 0,053 \text{ м}^2 \cdot \text{с}^{-1}.$$

Используем полученные ранее аналитические выражения для концентраций $n_{1,2}$ полярных аэроионов [18]:

$$n_1(z) = \frac{q}{\alpha n_\infty} (1 - e^{k_{12}(z-z_0)}), \quad k_{12} = \frac{b_1 E_0 - \sqrt{4D_0 \alpha n_\infty + (b_1 E_0)^2}}{2D_0}, \\ n_2(z) = \frac{q}{\alpha n_\infty} (1 - e^{k_{22}(z-z_0)}), \quad k_{22} = \frac{-b_2 E_0 - \sqrt{4D_0 \alpha n_\infty + (b_2 E_0)^2}}{2D_0}, \quad (5)$$

где $b_{1,2}$ – подвижности полярных аэроионов, α – коэффициент рекомбинации, q – скорость ионизации воздуха, e – элементарный заряд, n_∞ асимптотическое значение концентраций аэроионов. В случаях (а) и (в) в выражения (5) вместо параметра D_0 подставлялось соответствующее значение $D(t)$ в каждом временном интервале.

Удельная электрическая проводимость и ее асимптотическое значение рассчитывались для каждой комбинации параметров (а, б, в) моделирования на соответствующей высоте по следующим аналитическим выражениям [18]:

$$\lambda(z) = e(b_1 n_1(z) + b_2 n_2(z))$$

$$\lambda_{\infty} = e n_{\infty} (b_1 + b_2), \quad (6)$$

где $n_{\infty} = \sqrt{q/\alpha}$ – асимптотическое значение концентрации полярных ионов.

Глобальная унитарная вариация напряженности электрического поля представлена в относительных единицах как отношение абсолютных значений с шагом дискретизации 1 час к среднесуточному по формуле:

$$E_{\text{отн}}^{\text{унит}} = E_{\text{отн}}^{\text{унит}} / E_{\text{ср.сут.}}^{\text{унит}};$$

$$E_{\text{ср.сут.}} = \frac{j(t)}{\lambda}. \quad (7)$$

Расчеты выполнены на трех высотах: $z = 0,5 \text{ м}$, $z = 1 \text{ м}$ и $z = 2 \text{ м}$. Для каждого уровня проведен вычислительный эксперимент во всех режимах моделирования. В выражения (4) и (5), в зависимости от выбранного режима моделирования, подставлены либо значения $D(t)$ и $j(t)$, рассчитанные по (2) и (3) с шагом дискретизации 1 час, либо постоянные значения, заданные в параметрах моделирования.

Результаты. Относительные значения напряженности электрического поля $E_{\text{отн}} = E(t) / E(t)_{\text{ср.сут.}}$ (отношение часовых значений к среднесуточному по соответствующему режиму моделирования) на высотах 0,5 м, 1,0 м и 2,0 м, приведены в табл. 1. На рис. 1-3 приведен суточный ход напряженности электрического поля при различных режимах моделирования на указанных высотах.

На основе сравнительного анализа данных моделирования суточного хода электрического поля ($E_{\text{отн}}$) на высотах 0,5 м, 1,0 м и 2,0 м и глобальной унитарной вариации ($E_{\text{отн}}^{\text{унит}}$) установлено, что изменение высоты расчётной точки (наблюдения) существенно влияет на амплитуду исследуемых процессов. А именно, увеличение высоты точки наблюдения приводит к существенному увеличению размаха суточных колебаний. Наибольшая амплитуда регистрируется на высоте 2,0 м в режимах, учитывающих изменяющийся во времени коэффициент турбулентной диффузии $D(t)$.

В режиме постоянного коэффициента турбулентной диффузии D_0 при изменяющейся плотности тока $j(t)$ (кривая 3) на всех высотах отмечается наименьшая амплитуда суточных изменений со сглаженными, слабо выраженными экстремумами.

Режим динамической турбулентной диффузии $D(t)$ при постоянной плотности тока j_0 (кривая 2) формирует глубокий утренний минимум строго в 5 ч ВСВ, достигая 60% на высоте 0,5 м и 52% на высоте 2,0 м. Высокий вечерний максимум этой кривой регистрируется в интервале 17–18 ч ВСВ на высоте 0,5 м 112% и смещается чётко на 17 ч ВСВ 127% с ростом высоты до 2,0 м.

Совместный учёт динамических параметров $D(t)$ и $j(t)$ (кривая 4) сохраняет временные положения экстремумов (минимум в 5 ч ВСВ, максимум в 17 ч ВСВ), но существенно увеличивает размах колебаний. На высоте 2,0 м это приводит к наиболее сильному падению относительной напряжённости поля утром (до 43%) и её максимальному росту в вечернее время (до 134%).

Таблица 1

Суточный ход электрического поля $E_{отн}$ при различных режимах моделирования

$t, \text{ч}$ $z, \text{м}$		0-1	1-2	2-3	3-4	4-5	5-6	6-7	7-8	8-9	9-10	10-11	11-12
(a)	0,5	1,05	1,01	0,95	0,86	0,74	0,60	0,68	0,82	0,92	0,99	1,03	1,06
	1,0	1,05	0,98	0,89	0,77	0,62	0,49	0,55	0,71	0,85	0,95	1,02	1,08
	2,0	1,02	0,93	0,82	0,69	0,57	0,52	0,54	0,64	0,77	0,88	0,98	1,07
(б)	0,5	1,00	0,99	0,97	0,96	0,96	0,96	0,96	0,96	0,96	0,98	0,99	1,00
	1,0	0,99	0,97	0,96	0,95	0,95	0,94	0,94	0,94	0,95	0,96	0,98	0,99
	2,0	0,99	0,97	0,95	0,94	0,93	0,93	0,93	0,93	0,93	0,94	0,96	0,98
(в)	0,5	1,05	1,01	0,95	0,85	0,71	0,56	0,64	0,80	0,91	0,98	1,03	1,07
	1,0	1,06	0,98	0,88	0,74	0,57	0,42	0,49	0,67	0,82	0,93	1,02	1,08
	2,0	1,03	0,92	0,79	0,63	0,49	0,43	0,45	0,55	0,70	0,84	0,96	1,06
$E_{отн}^{унит}$		1,00	0,95	0,90	0,86	0,83	0,81	0,80	0,81	0,83	0,86	0,90	0,95

$t, \text{ч}$ $z, \text{м}$		12-13	13-14	14-15	15-16	16-17	17-18	18-19	19-20	20-21	21-22	22-23	23-24
(a)	0,5	1,08	1,10	1,11	1,11	1,12	1,12	1,12	1,11	1,11	1,10	1,09	1,07
	1,0	1,12	1,15	1,17	1,19	1,20	1,20	1,20	1,19	1,18	1,16	1,14	1,10
	2,0	1,13	1,18	1,22	1,25	1,27	1,28	1,27	1,26	1,24	1,20	1,16	1,10
(б)	0,5	1,01	1,02	1,03	1,03	1,03	1,04	1,04	1,03	1,03	1,03	1,02	1,01
	1,0	1,01	1,03	1,04	1,05	1,05	1,06	1,06	1,06	1,05	1,04	1,03	1,01
	2,0	1,00	1,03	1,05	1,06	1,08	1,08	1,08	1,08	1,07	1,05	1,03	1,01
(в)	0,5	1,09	1,11	1,12	1,12	1,13	1,13	1,13	1,13	1,12	1,11	1,10	1,08
	1,0	1,13	1,17	1,19	1,21	1,22	1,23	1,23	1,22	1,21	1,19	1,16	1,12
	2,0	1,15	1,21	1,26	1,30	1,32	1,33	1,33	1,32	1,29	1,25	1,20	1,12
$E_{отн}^{унит}$		1,00	1,05	1,10	1,14	1,17	1,19	1,20	1,19	1,17	1,14	1,10	1,05

Заметное расхождение кривых 2 и 4 на всех рассматриваемых высотах показывает, что вклад временных вариаций плотности электрического тока $j(t)$ в общую изменчивость суточного хода является весьма существенным. Наложение вариаций тока $j(t)$ на динамическую диффузию $D(t)$ (кривая 4) приводит к дополнительному сильному увеличению, как утреннего минимума, так и вечернего максимума. Таким образом, в то время как временной ход коэффициента турбулентной диффузии $D(t)$ выступает в качестве основного локального фактора, формирующего структуру и фазу суточной вариации электрического поля, изменчивость плотности тока $j(t)$ выполняет функцию амплитудного модулятора. Эти два фактора действуют синергетически, определяя результирующую локальную динамику электрического поля в приземном электродном слое.

Глобальная унитарная вариация электрического поля (кривая 1) демонстрирует экстремумы в 6 ч ВСВ (минимум) и 18 ч ВСВ (максимум). Модели, учитывающие динамический коэффициент турбулентной диффузии $D(t)$ (кривые 2 и 4), показывают систематическое опережение утреннего экстремума на 1 час (сдвиг на 5 ч ВСВ), что обусловлено локальными приземными метеорологическими процессами. Вечерний максимум модели

на всех высотах практически совпадает с фазой глобальной унитарной вариации (17–18 ч ВСВ). При этом расчетный профиль при полном учёте динамических параметров $D(t)$ и $j(t)$ (кривая 4) на высоте 2,0 м наиболее сильно отклоняется от глобальной унитарной вариации, значительно превышая её по размаху: колебания составляют от 43% до 133% против 80% и 120% соответственно.

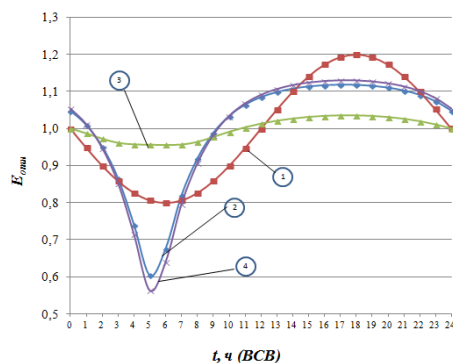


Рис. 1. Суточный ход электрического поля при различных режимах моделирования на высоте 0,5 м: 1 – глобальная унитарная вариация, 2 – режим $D(t)$ и j_0 , 3 – режим D_0 и $j(t)$, 4 – режим $D(t)$ и $j(t)$

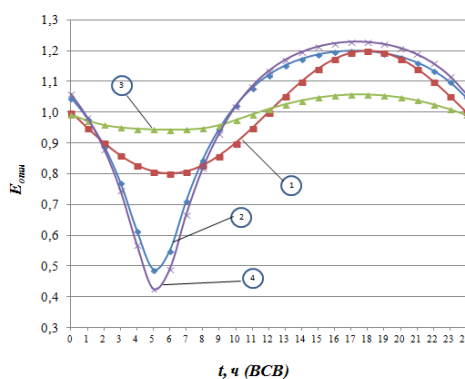


Рис. 2. Суточный ход электрического поля при различных режимах моделирования на высоте 1 м: 1 – глобальная унитарная вариация, 2 – режим $D(t)$ и j_0 , 3 – режим D_0 и $j(t)$, 4 – режим $D(t)$ и $j(t)$

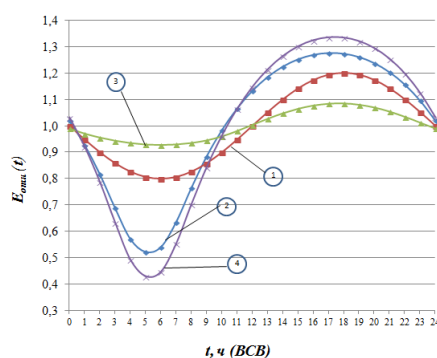


Рис. 3. Суточный ход электрического поля при различных режимах моделирования на высоте 2 м: 1 – глобальная унитарная вариация, 2 – режим $D(t)$ и j_0 , 3 – режим D_0 и $j(t)$, 4 – режим $D(t)$ и $j(t)$

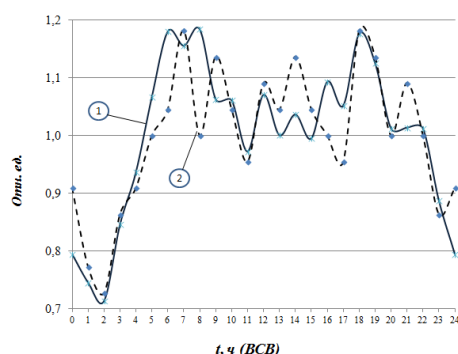


Рис. 4. Суточный ход электрических характеристик атмосферы на Пике Чегет (летний период): 1 – $E_{отн}$, 2 – $j_{отн}$

На рис. 4 приведен суточный ход напряженности (градиента потенциала) электрического поля $E_{отн}$ и плотности полного электрического тока $j_{отн}$ по данным натурных наблюдений [8] на высокогорной станции Пик Чегет ($43^{\circ} 16'$ с. ш. и $42^{\circ} 30'$ в. д.), расположенной в Приэльбрусье на высоте 3040 м н. у. м. Относительная плотность тока рассчитывалась как $j_{отн} = j(t) / j(t)_{ср.сут.}$ (отношение часовых значений к среднесуточному).

При сопоставлении результатов моделирования и экспериментальных данных [8] выявлено совпадение по фазе суточных колебаний. Расчётные минимумы для режимов, учитывающих динамическую диффузию (кривые 2 и 4 на рис. 1 и 3 при 5 ч ВСКВ) совпадают по фазе с экспериментально зафиксированным суточным минимумом на Пике Чегет, приходящимся на 2 ч ВСКВ. Возникающий временной сдвиг (около 3 часов) находится в рамках допустимой погрешности и обусловлен суточной динамикой локального ветрового режима высокогорной станции.

Расчётные кривые, учитывающие динамическую диффузию (кривые 2 и 4 на рис. 1 и 3), формируют протяжённую волну повышенных значений в интервале 11–21 ч ВСКВ с пологим пиком в районе 17–18 ч ВСКВ. Это демонстрирует точное временное совпадение с главным вечерним максимумом натурных измерений на Пике Чегет (16–19 ч ВСКВ с абсолютным пиком в 18 ч ВСКВ), где значение напряженности электрического поля достигает своего пика в 118% (в 6, 8 и 18 ч ВСКВ). Отсутствие в модели выраженного дневного пика (фиксируемого в эксперименте в 06–11 ч ВСКВ) физически обосновано принятыми в расчётах условиями устойчивой стратификации приземного слоя, при которых, согласно данным [17], дополнительный максимум нивелируется.

Анализ показывает, что полученная в ходе моделирования амплитуда колебаний поля качественно согласуется с реальными измерениями, отражая сильное влияние локальных факторов по мере удаления от поверхности земли. В утреннем минимуме (5 ч ВСКВ) на высоте 0,5 м расчётное относительное значение падает до 60% (кривая 2) и 56% (кривая 4) от среднего уровня, а на высоте 2,0 м – до 52% (кривая 2) и 43% (кривая 4). В вечернем максимуме (17 ч ВСКВ) значения электрического поля поле повышаются до 112% (кривая 2) и 113% (кривая 4) на высоте 0,5 м, а на высоте 2,0 м – до 127% (кривая 2) и 134% (кривая 4). Столь значительный размах суточных колебаний на двух метрах (от 43% до 133% от среднесуточного значения) свидетельствует о резком усилении влияния турбулентных процессов при удалении от подстилающей поверхности.

Полученный в ходе моделирования эффект увеличения утреннего минимума поля с ростом высоты (значение относительной напряженности падает с 56 % на высоте 0,5 м до 43% на высоте 2,0 м для комплексного режима) качественно подтверждает выводы работы [17], где отмечено, что усиление турбулентного перемешивания приводит к росту амплитудных значений вариаций электрического поля. Анализ результатов моделирования показывает, что определяющим локальным фактором является суточная изменчивость

коэффициента турбулентной диффузии $D(t)$, которая формирует саму структуру и фазу суточного хода поля за счёт перераспределения объемных зарядов. В то же время, явное расхождение кривых 2 и 4 доказывает существенный синергетический вклад глобального фактора – временных вариаций плотности тока $j(t)$. Накладываясь на локальную турбулентность, вариации тока выступают в роли амплитудного модулятора, который значительно усиливает экстремумы и увеличивает общий размах колебаний по мере удаления от земной поверхности.

Таким образом, комплексное временное, амплитудное и пространственное сопоставление результатов моделирования и натурных измерений позволяют сделать вывод об эффективности предложенного подхода. Проведенный анализ подтверждает правомерность применения квазистационарной модели, а также применимость полученных аналитических решений для расчёта и прогнозирования поведения электрического поля в турбулентном приземном электродном слое.

Заключение. В работе проведено моделирование суточной динамики электрических характеристик приземного слоя атмосферы. Предложенный подход основан на квазистационарном сопряжении временных гармонических зависимостей локальных и глобальных факторов с аналитическими решениями уравнений электродного эффекта. Физическая правомерность методики расчета обоснована существенным различием временных масштабов установления электрического равновесия среды и периода суточного хода внешних генераторов. Разработанная расчетная схема позволила выявить определяющую роль временного хода турбулентного обмена, приводящего к значительному расширению амплитудного диапазона и деформации профиля суточных колебаний электрического поля по сравнению с глобальной унитарной вариацией. Достоверность полученных теоретических результатов подтверждена путем сравнения с экспериментальными данными натурных измерений в высокогорных условиях, продемонстрировав временное совпадение расчетных экстремумов с ходом суточной вариации электрического поля в летний период.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. *Mauchly S.J.* Studies in atmosphere electricity based on observations made on the Carnegie (1915-1921) // *Researches of the Department of Terrestrial Magnetism*. Washington: Carnegie Institution. – 1926. – P. 385-424.
2. *Harrison R.G.* The Carnegie curve // *Surveys in Geophysics*. – 2013. – Vol. 34, No. 2. – P. 209-232.
3. *Adzhiev A. Kh., Kupovykh G.V.* Measurements of the Atmospheric Electric Field under High-Mountain Conditions in the vicinity of Mt. Elbrus // *Izvestiya, Atmospheric and Oceanic Physics*. – 2015. – Vol. 51, No. 6. – P. 633-638.
4. *Kupovykh G.V., Kudrinskaya T.V., Gritsov V.V.* The atmosphere electrical characteristics' monitoring as an element of technosphere safety // *IOP Conference Series: Materials Science and Engineering*. – 2020. – 913 052041. – 6 p.
5. *Морозов В.Н.* Атмосферное электричество // *Атмосфера. Справочник (справочные данные, модели)*. – Л.: Гидрометеиздат, 1991. – С. 394-408.
6. *Куповых Г.В., Морозов В.Н., Шварц Я.М.* Теория электродного эффекта в атмосфере. – Таганрог: Изд-во ТРТУ, 1998. – 123 с.
7. *Морозов В.Н., Куповых Г.В.* Математическое моделирование глобальной атмосферной электрической цепи и электричества приземного слоя: монография. – СПб.: Астерион, 2017. – 307 с.
8. *Куповых Г.В.* Электродинамические процессы в приземном слое атмосферы. – Таганрог: Изд-во ТТИ ЮФУ, 2009. – 114 с.
9. *Куповых Г.В., Морозов В.Н.* Турбулентный электродный эффект в приземном слое // *Известия высших учебных заведений. Северо-Кавказский регион. Естественные науки. Приложение № 3*. – 2003. – С. 55-62.
10. *Редин А.А., Клово А.Г., Куповых Г.В.* Электродинамическая модель атмосферного приземного слоя // *Известия ЮФУ. Технические науки*. – 2009. – № 8. – С. 93-106.
11. *Kupovykh G., Redin A., Boldyreff A.* Modeling of ionization-recombination processes in the atmospheric surface layer // *Journal of Electrostatics*. – 2013. – Vol. 71. – P. 305-311.
12. *Кудринская Т.В., Куповых Г.В., Редин А.А.* Исследования ионизационного состояния приземного слоя атмосферы в разных геофизических условиях // *Метеорология и гидрология*. – 2018. – № 4. – С. 77-85.

13. Свидельский С.С., Литвинова В.С., Куповых Г.В., Клово А.Г. Формирование структуры атмосферного электродного слоя // Известия ЮФУ. Технические науки. – 2020. – № 5. – С. 130-141.
14. Kupovykh G.V., Timoshenko D.V., Klovo A. G., Kudrinskaya T.V. Electrodynamics models in atmospheric surface layer // IOP Conference Series: Materials Science and Engineering. – 2019. – Vol. 698 044034. – 8 p.
15. Kupovykh G.V., Timoshenko D.V., Kudrinskaya T.V., Klovo A.G. Modeling of the atmospheric electric field local variations in the turbulent surface layer // Journal of Physics: IOP Conference Series. VIII All-Russian Conference on Atmospheric Electricity. – 2020. – 1604 012003. – 5 p.
16. Timoshenko D.V., Kupovykh G.V., Kudrinskaya T.V. Daily dynamics of the electric field in the turbulent surface air layer under the action of local factors // Atmospheric and Oceanic Optics. – 2025. – Vol. 38, No. 1. – P. 77-82.
17. Adzhiev A.Kh., Klovo A.G., Kudrinskaya T.V., Kupovykh G.V., Timoshenko D.V. Diurnal Variations of the Electric Field in the Atmospheric Boundary Layer // Izvestiya, Atmospheric and Oceanic Physics. – 2021. – Vol. 57, No. 4. – P. 397-405.
18. Белоусова О.В., Куповых Г.В., Клово А.Г., Гривцов В.В. Результаты моделирования электродинамической структуры турбулентного приземного слоя // Известия ЮФУ. Технические науки. – 2022. – № 4 (228). – С. 245253.
19. Куповых Г.В., Белоусова О.В., Клово А.Г. Приближенные аналитические решения уравнений электродного эффекта в атмосфере // Сб. докладов XXVIII Всероссийской открытой научной конференции «Распространение радиоволн». – Йошкар-Ола, 2023. – С. 544-548.
20. Куповых Г.В., Тимошенко Д.В., Кудринская Т.В., Клово А.Г., Белоусова О.В. Локальные вариации электрического поля в турбулентном приземном слое // Матер. IX Всероссийской научной конференции по атмосферному электричеству. – СПб.: ВКА им. А.Ф. Можайского, 2023. – С. 497-503.
21. Тимошенко Д.В., Куповых Г.В. Исследование глобальных и локальных вариаций электрических характеристик турбулентного приземного слоя атмосферы // Тр. ГГО им. А.И. Воейкова. – СПб., 2025. – Вып. 619. – С. 115-128.
22. Тимошенко Д.В., Куповых Г.В., Белоусова О.В. Анализ электродинамической модели приземного слоя атмосферы // Известия высших учебных заведений. Северо-Кавказский регион. Естественные науки. – 2024. – № 4. – С. 104-112.
23. Kupovykh G.V., Timoshenko D.V., Klovo A.G., Kudrinskaya T.V. Electrode effect on the daily variation in the atmospheric electric field in the surface air layer // Atmospheric and Oceanic Optics. – 2024. – Vol. 37, No. 1. – P. 94-98.

REFERENCES

1. Mauchly S.J. Studies in atmosphere electricity based on observations made on the Carnegie (1915-1921), Researches of the Department of Terrestrial Magnetism. Washington: Carnegie Institution, 1926, pp. 385-424.
2. Harrison R.G. The Carnegie curve, *Surveys in Geophysics*, 2013, Vol. 34, No. 2, pp. 209-232.
3. Adzhiev A. Kh, Kupovykh G.V. Measurements of the Atmospheric Electric Field under High-Mountain Conditions in the vicinity of Mt. Elbrus, *Izvestiya, Atmospheric and Oceanic Physics*, 2015, Vol. 51, No. 6, pp. 633-638.
4. Kupovykh G.V., Kudrinskaya T.V., Grivtsov V.V. The atmosphere electrical characteristics' monitoring as an element of technosphere safety, *IOP Conference Series: Materials Science and Engineering*, 2020, 913 052041, 6 p.
5. Morozov V.N. Atmosfernoe elektrichestvo [Atmospheric electricity], *Atmosfera. Spravochnik (spravochnye dannye, modeli)* [Atmosphere. Handbook (reference data, models)]. Leningrad: Gidrometeoizdat, 1991, pp. 394-408.
6. Kupovykh G.V., Morozov V.N., Shvarts Ya.M. Teoriya elektrodnoy efekta v atmosfere [Theory of the Electrode Effect in the Atmosphere]. Taganrog: Izd-vo TRTU, 1998, 123 p.
7. Morozov V.N., Kupovykh G.V. Matematicheskoe modelirovanie global'noy atmosfernoy elektricheskoy tsepi i elektrichestva prizemnogo sloya: monografiya [Mathematical modeling of the global atmospheric electric circuit and surface layer electricity: monograph]. Saint Petersburg: Asterion, 2017, 307 p.
8. Kupovykh G.V. Elektrodinamicheskie protsessy v prizemnom sloe atmosfery [Electrodynamics processes in the surface layer of the atmosphere]. Taganrog: Izd-vo TTI YuFU, 2009, 114 p.
9. Kupovykh G.V., Morozov V.N. Turbulentnyy elektrodnyy effekt v prizemnom sloe [Turbulent electrode effect in the surface layer], *Izvestiya vysshikh uchebnykh zavedeniy. Severo-Kavkazskiy region. Estestvennye nauki. Prilozhenie* [Bulletin of Higher Educational Institutions. North Caucasus Region. Natural Sciences. App. No. 3], 2003, pp. 55-62.
10. Redin A.A., Klovo A.G., Kupovykh G.V. Elektrodinamicheskaya model' atmosfer'nogo prizemnogo sloya [Electrodynamics model of the atmospheric surface layer], *Izvestiya YuFU. Tekhnicheskie nauki* [Izvestiya SFedU. Engineering Sciences], 2009, No. 8, pp. 93-106.

11. Kupovykh G., Redin A., Boldyreff A. Modeling of ionization-recombination processes in the atmospheric surface layer, *Journal of Electrostatics*, 2013, Vol. 71, pp. 305-311.
12. Kudrinskaya T.V., Kupovykh G.V., Redin A.A. Issledovaniya ionizatsionnogo sostoyaniya prizemnogo sloya atmosfery v raznykh geofizicheskikh usloviyakh [Studies of the ionization state of the atmospheric surface layer under various geophysical conditions], *Meteorologiya i gidrologiya* [Meteorology and Hydrology], 2018, No. 4, pp. 77-85.
13. Svidel'skiy S.S., Litvinova V.S., Kupovykh G.V., Klovo A.G. Formirovanie struktury atmosfernogo elektrodnoogo sloya [Formation of the structure of the atmospheric electrode layer], *Izvestiya YuFU. Tekhnicheskie nauki* [Izvestiya SFedU. Engineering Sciences], 2020, No. 5, pp. 130-141.
14. Kupovykh G.V., Timoshenko D.V., Klovo A. G., Kudrinskaya T.V. Electrodynamic processes models in atmospheric surface layer, *IOP Conference Series: Materials Science and Engineering*, 2019, Vol. 698 044034, 8 p.
15. Kupovykh G.V., Timoshenko D.V., Kudrinskaya T.V., Klovo A.G. Modeling of the atmospheric electric field local variations in the turbulent surface layer, *Journal of Physics: IOP Conference Series. VIII All-Russian Conference on Atmospheric Electricity*, 2020, 1604 012003, 5 p.
16. Timoshenko D.V., Kupovykh G.V., Kudrinskaya T.V. Daily dynamics of the electric field in the turbulent surface air layer under the action of local factors, *Atmospheric and Oceanic Optics*, 2025, Vol. 38, No. 1, pp. 77-82.
17. Adzhiev A.Kh., Klovo A.G., Kudrinskaya T.V., Kupovykh G.V., Timoshenko D.V. Diurnal Variations of the Electric Field in the Atmospheric Boundary Layer, *Izvestiya, Atmospheric and Oceanic Physics*, 2021, Vol. 57, No. 4, pp. 397-405.
18. Belousova O.V., Kupovykh G.V., Klovo A.G., Grivtsov V.V. Rezul'taty modelirovaniya elektrodinamicheskoy struktury turbulentnogo prizemnogo sloya [Results of simulation of electrodynamic structure of turbulent surface layer], *Izvestiya YuFU. Tekhnicheskie nauki* [Izvestiya SFedU. Engineering Sciences], 2022, No. 4 (228), pp. 245253.
19. Kupovykh G.V., Belousova O.V., Klovo A.G. Priblizhennyye analiticheskie resheniya uravneniy elektrodnoogo effekta v atmosfere [Approximate analytical solutions of the electrode effect equations in the atmosphere], *Sb. dokladov XXVIII Vserossiyskoy otkrytoy nauchnoy konferentsii «Rasprostraneniye radiovoln»* [Proceedings of the XXVIII All-Russian Open Scientific Conference "Radio Wave Propagation"], Yoshkar-Ola, 2023, pp. 544-548.
20. Kupovykh G.V., Timoshenko D.V., Kudrinskaya T.V., Klovo A.G., Belousova O.V. Lokal'nye variatsii elektricheskogo polya v turbulentnom prizemnom sloe [Local variations of the electric field in a turbulent surface layer], *Mater. IX Vserossiyskoy nauchnoy konferentsii po atmosfernomu elektrichestvu* [Proceedings of the IX All-Russian Scientific Conference on Atmospheric Electricity]. Saint Petersburg: VKA im. A.F. Mozhayskogo, 2023, pp. 497-503.
21. Timoshenko D.V., Kupovykh G.V. Issledovanie global'nykh i lokal'nykh variatsiy elektricheskikh kharakteristik turbulentnogo prizemnogo sloya atmosfery [Investigation of global and local variations of electrical characteristics of the turbulent surface layer of the atmosphere], *Tr. GGO im. A.I. Voeykova* [Proceedings of the Voeikov Main Geophysical Observatory]. Saint Petersburg, 2025, Issue 619, pp. 115-128.
22. Timoshenko D.V., Kupovykh G.V., Belousova O.V. Analiz elektrodinamicheskoy modeli prizemnogo sloya atmosfery [Analysis of the electrodynamic model of the atmospheric surface layer], *Izvestiya vysshikh uchebnykh zavedeniy. Severo-Kavkazskiy region. Estestvennye nauki* [Bulletin of Higher Educational Institutions. North Caucasus Region. Natural Sciences], 2024, No. 4, pp. 104-112.
23. Kupovykh G.V., Timoshenko D.V., Klovo A.G., Kudrinskaya T.V. Electrode effect on the daily variation in the atmospheric electric field in the surface air layer, *Atmospheric and Oceanic Optics*, 2024, Vol. 37, No. 1, pp. 94-98.

Белусова Ольга Васильевна – Южный федеральный университет; e-mail: obelousova@sfedu.ru; г. Таганрог, Россия; тел.: +79281727171.

Belousova Olga Vasil'evna – Southern Federal University; e-mail: obelousova@sfedu.ru; Taganrog, Russia; phone: +79281727171.

И.С. Бурлаченко, Г.М. Чечин

РЕЗОНАНСНОЕ ВОЗБУЖДЕНИЕ БУШЕЙ НЕЛИНЕЙНЫХ НОРМАЛЬНЫХ МОД В ДВУМЕРНЫХ РЕШЁТКАХ

Рассматривается проблема резонансного возбуждения бушей в двумерных решётках с дискретной симметрией. Каждый буш представляет собой совокупность колебательных мод, которая сохраняется во времени и определяет некоторое точное решение нелинейных уравнений движения, которое нельзя получить в рамках теории возмущений. Актуальность данного исследования, заключается в том, что указанная схема возбуждения бушей, в отличие от прошлых математических моделей направленных на возбуждения бушей нелинейных нормальных мод может быть реализована в рамках физического эксперимента. Вышеуказанные ансамбли колебательных мод рассматриваются в рамках теории Чечина и Сахненко, основанной на теоретико-групповых методах исследования, что позволяет с равным успехом использовать любые модели потенциалов межатомного взаимодействия. Что в свою очередь – позволяет рассматривать динамические системы множества масштабов от молекулярных колебаний до движения макроскопических кристаллических структур. Цель работы: представить математическую модель возбуждения устойчивых нелинейных динамических объектов в двумерных молекулярных структурах пригодную как для вычислительного, так и для физического эксперимента. Предоставление результатов проведённого вычислительного эксперимента с описанием обнаруженных динамических объектов. В отличие от традиционных методов, теория бушей обеспечивает твёрдый теоретический аппарат, основанный на теории групп, который преодолевает ограничения теории возмущений, предлагая результаты исключительной надёжности. При этом свойства симметрии, используемые в теории бушей, способны давать значительные преимущества для численного моделирования, заметно снижая вычислительную сложность. На данный момент указанная проблематика исследуется многими авторами по всему миру. Стоит, однако, отметить, что работы, в которых обсуждаются проблемы, возбуждения динамических объектов, близких к понятию бушей, относятся только к одномерным случаям и что теоретико-групповые методы в них, за весьма редким исключением, не рассматриваются.

Нелинейные нормальные моды; моды Розенберга; пространственные группы симметрии; вычислительная физика; подгруппы пространственных групп симметрии; инвариантные множества; теоретико-групповые правила отбора; Гамильтоновы системы; буши нелинейных нормальных мод; колебательное движение.

I.S. Burlachenko, G.M. Chechin

RESONANT EXCITATION OF BUSH NONLINEAR NORMAL MODES IN TWO-DIMENSIONAL LATTICES

This paper addresses the problem of resonant excitation of bushes in two-dimensional lattices with discrete symmetry. A bush is defined as a set of vibrational modes that persists over time and constitutes an exact solution to the nonlinear equations of motion, which cannot be obtained within the framework of perturbation theory. The relevance of this study lies in the fact that the proposed scheme for bush excitation, in contrast to previous mathematical models aimed at exciting bushes of nonlinear normal modes, can be implemented in a physical experiment. The above-mentioned ensembles of vibrational modes are considered within the framework of the theory of Chechin and Sakhnenko, based on group-theoretic research methods, which makes it possible to use any models of interatomic interaction potentials with equal success. This, in turn, allows us to consider dynamic systems of multiple scales, from molecular vibrations to the movement of macroscopic crystal structures. The aim of the work is to present a mathematical model of excitation of stable nonlinear dynamic objects in two-dimensional molecular structures suitable for both computational and physical experiments. Providing the results of a computational experiment with a description of the detected dynamic objects. Unlike traditional methods, Bush theory provides a solid theoretical framework based on group theory that overcomes the limitations of perturbation theory, offering results of exceptional reliability. At the same time, the symmetry properties used in Bush theory can provide significant advantages for numerical modeling, significantly reducing computational complexity. Now, this issue is being investigated by many authors around the world. It is worth noting, however, that the works in which the problems of excitation of dynamic objects close to the concept of bushes are discussed relate only to one-dimensional cases and that group-theoretic methods are not considered in them, with very rare exceptions.

Nonlinear normal modes; Rosenberg modes; space symmetry groups, computational physics; subgroups of space symmetry groups; invariant manifolds; group-theoretical selection rules; Hamiltonian systems; bushes of nonlinear normal modes; oscillatory motion.

Введение. Ставшая классической работа Ферми, Пасты, Улама и Цингоу опубликованная 1955 г. [1] положила начало многочисленным исследованиям межмодовых взаимодействий в нелинейных структурах. Тогда, однако, не был поставлен вопрос об определении правил отбора для передачи возбуждения между нелинейными модами различной симметрии. Данные правила отбора обеспечивают возможность возбуждения устойчивых ансамблей однотипных мод, а набор мод, взаимодействующих внутри каждого такого ансамбля, со временем не меняется. Вышеуказанные группы рассматриваются в рамках теории так называемых «бушей» развитой в работах Чечина и Сахненко [2–4], в рамках которой эти динамические режимы называются бушами нелинейных нормальных мод. В этих работах впервые была поставлена и решена задача о теоретико-групповых методах исследования и передачи возбуждения между модами *разной симметрии*. Ввиду теоретико-групповой основы теории бушей – она имеет широкие границы применимости и не зависит от типа межатомных взаимодействий в системе. На данный момент указанная проблематика исследуется многими авторами по всему миру.

Стоит, однако, сразу отметить, что работы, в которых обсуждаются проблемы, возбуждения динамических объектов, близких к понятию бушей, относятся только к одномерным случаям и что теоретико-групповые методы в них не рассматриваются. Исключением является статья Боба Ринка [5], в которой содержится обсуждение теоретико-групповых вопросов, но опять-таки только для одномерного случая.

Прежде всего, стоит отметить статью Харольда Стоукса и Дориана Хэтча [6], где в случае цепочки ФПУ были обнаружены долгоживущие динамические объекты, которые являлись линейными комбинациями некоторого числа линейных колебательных мод. Эта работа авторов была основана на анализе свойств самих динамических уравнений широко-известной модели Ферми, Пасты и Улама, которая была предложена этими авторами при исследовании механической системы грузиков, связанных пружинками с нелинейными добавками к обычному закону Гука. Как известно, эта работа является именно той работой, которая и породила современную нелинейную динамику. Совместно с ними опубликован ряд работ, объединяющие оба подхода [7].

Целый ряд работ таких авторов как Йошимура, Шинохара продолжили эту традицию по исследованию цепочек ФПУ на основе специфики уравнений движения этой модели. Стоит ещё раз подчеркнуть, что единственной работой, нарушившей эту традицию, является уже упоминавшаяся работа Боба Ринка [5]. Но в ней опять-таки исследование ограничено лишь одномерным случаем и не была выявлена фундаментальная роль подгрупп группы симметрии рассматриваемых динамических систем для исследования их динамики.

Отдельно нужно выделить работы, в которых используются первопринципные расчёты на основе методов типа Хартри-Фока и метода функционала плотности, в частности работы Доя, Нишимуры и их японских коллег [8, 9].

В заключение стоит подчеркнуть, что динамические объекты, в той или иной степени аналогичные бушам колебательных мод, представляют существенный интерес для современной физики.

Однако есть один принципиальный момент, связанный с теорией бушей. Дело в том, что могут существовать объекты, не связанные только с механическими движениями широко изучаемых систем. Могут существовать и буши **немеханической природы**, например буши **магнитных мод** и разные их комбинации с колебательными модами. В этом отношении открывается широкая перспектива по части дальнейшего развития теории бушей нелинейных нормальных мод.

В рамках данной работы рассматривается математическая модель, в которой предполагается, что колебания в системе возбуждаются путём воздействия на один из незакреплённых атомов решётки внешней периодической силой. Поскольку атом, на который действует эта внешняя сила, передаёт возбуждение на связанные с ним атомы – то в случае наличия резонансной частоты – возбуждение передаётся по всей структуре, что и приводит к возбуждению некоторой группы нелинейных нормальных мод. Такая группа может соответствовать некоторому бушу, сохраняющему свою структуру в течение эволюции системы во времени.

Теория бушей нелинейных нормальных мод. Постановка задачи. В 1930 году Вигнер впервые использовал теоретико-групповые методы для изучения *малых атомных колебаний* молекул и кристаллов в своей работе [10]. Он показал, что нормальные моды, которые вводятся в гармоническом приближении для систем с дискретной симметрией, могут быть классифицированы по неприводимым представлениям группы симметрии этих равновесных систем, независимо от их физической природы.

Все моды, связанные с данным неприводимым представлением, имеют одинаковую частоту, а соответствующие им собственные векторы матрицы силовых постоянных могут рассматриваться как базисные векторы этого неприводимого представления. Возможность такой классификации основана на *линейности* динамических уравнений, полученных при использовании гармонического приближения.

Линейность уравнения Шредингера позволяет провести аналогичную классификацию квантовых состояний (теорема Вигнера) [11].

В 90-х годах прошлого века в Ростовской исследовательской группе под руководством Г.М. Чечина была разработана теория бушей нелинейных нормальных мод, с целью изучения колебаний атомов в системах с дискретной симметрией при *произвольных амплитудах* колебаний [2–4].

Существуют различные концепции, определяющие понятие нелинейных нормальных мод. Теория бушей использует концепцию таких мод, восходящих к определению, данному Розенбергом. Моды Розенберга оказались весьма плодотворной и удобной концепцией – широко используемой в разнообразных современных инженерных расчётах [12, 13].

Подобно теории Вигнера, эта теория основана на применении аппарата неприводимых представлений групп точечной и пространственной симметрии. Теория бушей нелинейных нормальных мод (ННМ) основана на исследовании правил отбора для передачи возбуждения между модами различной симметрии.

Буши нелинейных нормальных мод, представляют собой *точные* динамические объекты в нелинейных системах с дискретной симметрией, которые нельзя получить ни в каком порядке обычной теории возмущений.

Присутствие точечной или пространственной симметрии в нелинейной системе обуславливает наличие правил отбора (ПО), регулирующих процесс передачи возбуждения между нормальными модами данной системы [4–6].

Правила отбора могут быть получены на основе метода полного конденсата первичных и вторичных параметров порядка, который был разработан в рамках теории фазовых переходов исследовательской группой под руководством Г.М. Чечина и В.П. Сахненко в НИИФ РГУ.

В классической механике симметрия системы материальных точек определяется набором их скоростей и координат. Однако для последующего анализа симметричного состояния системы достаточно учитывать только пространственную группу симметрии её геометрической конфигурации.

Предположим, что группа G_0 описывает систему в состоянии равновесия. При возникновении малых колебаний в этой системе *симметрия понижается* и будет характеризоваться некоторой подгруппой G группы G_0 :

$$G \subset G_0.$$

Совокупность смещений всех N частиц рассматриваемой системы из положения равновесия может быть представлена в виде вектора:

$$\vec{X}(t) = \{\delta x_1(t), \dots, \delta x_{3N}(t)\}. \quad (1)$$

Данный вектор можно разложить по полному набору базисных функций всех неприводимых представлений группы G_0 :

$$\vec{X}(t) = \sum_{l=1}^{3N} \mu_l(t) \varphi_l, \quad (2)$$

здесь φ_l – базисная функция неприводимого представления, которая определяет направления смещений для всех атомов системы, а $\mu_l(t)$ – задаёт временную зависимость этих смещений.

В дальнейшем, согласно данному представлению, под колебательной модой будем понимать линейную комбинацию базисных функций, определяемую набором коэффициентов $\mu_i(t)$.

Предположим, что в начальный момент времени в системе возбуждена определенная корневая мода, принадлежащая неприводимому представлению $\Gamma_j^{(k)}$. Здесь и далее под корневой модой понимается мода, которая возбуждается в начальный момент, в то время как другие моды буша, которые вовлечены в колебательный процесс из-за межмодовых нелинейных взаимодействий называются *вторичными модами* этого буша.

Рассмотрим теоретико-групповой метод для определения всех вторичных мод, которые могут быть возбуждены в процессе динамики рассматриваемой системы.

Учитывая инвариантность вектора $\vec{X}(f)$ относительно всех элементов g группы G , определяющей симметрию корневой моды, для каждого неприводимого представления Γ_j группы G_0 справедливо соотношение:

$$M_j(g)\vec{C}_j = \vec{C}_j, \quad g \in G. \quad (3)$$

Здесь $M_j(g)$ – матрица представления Γ_j размерности n_j , соответствующая элементу g , а компоненты C_{ji} вектора $\vec{C}_j = (C_{j1}, \dots, C_{jn_j})$ – коэффициенты разложения $\vec{X}(t)$ по базисным функциям этого представления.

Решение $\vec{C}_j$ уравнения (3) для каждого НП является стационарным вектором, который не изменяется при воздействии на него любых матриц этого представления для $g \in G$. Этот вектор, зависящий от S_j произвольных параметров ($0 \leq S_j \leq n$) выделяет некоторую колебательную моду.

Пусть, $\vec{C}^{(k)}$ обозначает стационарный вектор, определяющий корневую моду представления $\Gamma^{(k)}$. Множество всех ненулевых векторов $\vec{C}_j$, получаемых путем решения системы (3) для каждого НП группы G_0 , будет называться бушем стационарных векторов, который соответствует корневому вектору $\vec{C}^{(k)}$. В теории структурных фазовых переходов этот набор называется полным конденсатом параметров порядка. Из метода нахождения буша мод с использованием системы (3) ясно, что группа симметрии G' любой вторичной моды не ниже группы симметрии G корневой моды ($G' \geq G$). Критерием же для полноты набора мод, входящих в рассматриваемый буш с симметрией G является:

$$\sum_j n_j S_j = index \quad (4)$$

Здесь n_j обозначает размерность НП Γ_j , S_j – число произвольных параметров, описывающих стационарный вектор $\vec{C}_j$, $index$ – индекс подгруппы G в группе G_0 .

Существенно, что обычно достаточно большое число НП группы G_0 не дает вклада в данный буш (для них $\vec{C}_j = 0$). Следовательно, соответствующие моды не могут быть возбуждены в результате взаимодействия с корневой модой при любых ангармонических воздействиях, что указывает на строгие правила отбора для передачи возбуждения между модами различной симметрии.

Как известно, матричное представление группы симметрии использует матрицы, – которые представляют операторы симметрии. В рамках матричных представлений каждому элементу группы симметрии сопоставляется матрица, которая действует на скаляры, векторы или псевдовекторы, описывающие систему. Называют эти представления – соответственно: перестановочными, механическими или спинowymi.

Общее число механических степеней свободы μ , для данного буша, вытекающее из формулы (2), определяется как $m = \sum_j P_j S_j$, где P_j – количество раз, которое Γ_j входит в механическое представление рассматриваемой системы, причем динамические уравнения для них можно написать в виде:

$$\ddot{\mu}_l = -\frac{\partial U}{\partial \mu_l}, \quad (5)$$

где $l = 1, \dots, m$, а $U = U(\mu_1, \dots, \mu_m)$ – потенциальная энергия системы, выраженная через переменные μ_l .

Предположив аналитичность функции $U = U(\mu_1, \dots, \mu_m)$ можно представить её в форме бесконечного степенного ряда по всем аргументам. Аргументы в разложении – могут быть представлены в виде инвариантных комбинаций относительно прямой суммы НП, соответствующих данному бушу.

Разложения функции $U(\mu_1, \mu_2)$ для различных кристаллических структур часто имеют одинаковый вид, что позволяет выделять соответствующие им буши в отдельные классы универсальности.

Простейший класс двумерных бушей определяется функцией

$$U(\mu_1, \mu_2) = \sum_{k=0}^{\infty} \mu_1^k F_k(\mu_2), \quad (6)$$

где $F_k(\mu_2)$ суть произвольные четные функции от μ_2 .

В контексте анализа малых колебаний рассмотрение можно ограничить небольшим количеством первых ангармонических членов в выражении (6). Для минимизации количества произвольных констант используется соответствующая замена переменных. В результате, уравнения (5), определяющие функцию $U(\mu_1, \mu_2)$ из (6), могут быть представлены следующим образом:

$$\ddot{\mu}_1 + \mu_1 = \mu_2^2 + \dots, \quad (7a)$$

$$\ddot{\mu}_2 + \omega^2 \mu_2 = 2\mu_1 \mu_2 + \dots, \quad (7b)$$

Уравнения (7) являются динамическими уравнениями двух связанных нелинейных осцилляторов, соответствующих модам $\mu_1(t)$ и $\mu_2(t)$.

Из структуры данных уравнений следует, что активация корневой моды $\mu_2(t)$ неизбежно вызывает возбуждение вторичной моды $\mu_1(t)$ поскольку система не допускает решения, при котором $\mu_2(t) \neq 0$ и одновременно $\mu_1(t) \equiv 0$. В то же время возбуждение $\mu_1(t)$, вообще говоря, не приводит к появлению корневой моды $\mu_2(t)$ так как система допускает решение, при котором $\mu_1(t) \neq 0$ и $\mu_2(t) \equiv 0$. Подобное подчинение мод для бушей большей размерности приводит к созданию иерархии взаимодействия между ними.

Правила отбора, определяющие моды определённого буша, базируются на основе общих симметричных соображений. Тем не менее, эти правила могут быть спонтанно нарушены под влиянием процессов аналогичных параметрическому резонансу. В результате, могут быть возбуждены моды, которые не являются частью первоначального буша с симметрией G , что приводит к его расширению и одновременному снижению симметрии до подгруппы $G' \subset G$.

В этом случае мы работаем в области неустойчивых решений и имеем дело с некоторым динамическим аналогом фазового перехода. В таком переходе мгновенная симметрия колеблющейся системы спонтанно снижается, что влечёт за собой расширение первоначального буша.

Теория бушей и резонанс. В этой работе была представлена математическая модель, которая описывает систему, состоящую из нескольких одинаковых грузов, соединенных пружинами. Растяжение таких пружин вызывает возникновение упругой силы, которая подчиняется закону Гука с добавлением нелинейных членов второго (модель FPU- α) или третьего (модель FPU- β) порядка. Далее вместо грузиков будут рассматриваться атомы или частицами, а вместо нелинейных пружин – потенциалы межатомного взаимодействия.

Данная элементарная механическая модель оказала значительное воздействие на современную теорию нелинейной динамики, а также послужила основой для открытия таких феноменов, как солитоны [14] и динамический хаос [15, 16].

В последующем, процесс возбуждения моды в системе для каждой ФПУ-подобной модели как правило включает в себя проведение следующей математической процедуры: сначала рассчитывается вектор начальных смещений, который является собственным вектором матрицы силовых постоянных системы; затем устанавливаются маленькие начальные смещения атомов в соответствии с этим вектором; начальные скорости устанавливаются равными нулю; задаются граничные условия (чаще всего периодические). После этого выполняется решение задачи Коши для системы нелинейных уравнений движения с использованием определенного численного метода на некотором временном интервале.

Присутствие ангармонических членов в разложении потенциальной энергии приводит к возникновению слабого взаимодействия между нормальными модами, что способствует возникновению определенного упорядоченного набора нелинейных нормальных мод.

Описанная выше процедура возбуждения колебательных бушей довольно проста для выполнения в вычислительных экспериментах, но ее невозможно реализовать в экспериментах физических.

Предлагаемый метод возбуждения колебательных бушей основан на резонансных воздействиях на изучаемую систему и, в отличие от метода, описанного выше, может быть осуществлён в рамках физических экспериментов.

Резонансное возбуждение бушей нелинейных нормальных мод в двумерных решётках. Основные результаты. Рассмотрим двумерную квадратную решётку, все межатомные взаимодействия которой описываются потенциалом Леннарда-Джонса и атомы которой оказывают прямое влияние лишь на атомы первой «координационной сферы». Закрепим все граничные атомы системы и будем считать массы всех атомов равными друг другу.

Размер решётки положим равным 3×3 подвижных атомов. Начнём поочерёдно прикладывать к каждому атому системы внешнюю периодическую силу равную $0.01 \cos(\omega t)$ и имеющую направление $[1,0]$. Интегрирование будем проводить в течении $t=4000$. Здесь и далее, нумерация атомов ведётся с учётом граничных, неподвижных атомов.

Как показало моделирование, энергетические пики резонансных частот весьма различны по своей высоте, что приводит к разным амплитудам возбуждаемых колебаний.

К примеру, вот так (табл. 1–3) выглядит процесс накачки энергии при разных частотах, для случаев, когда периодическая сила приложена к атому $(2,2)$.

Таблица 1

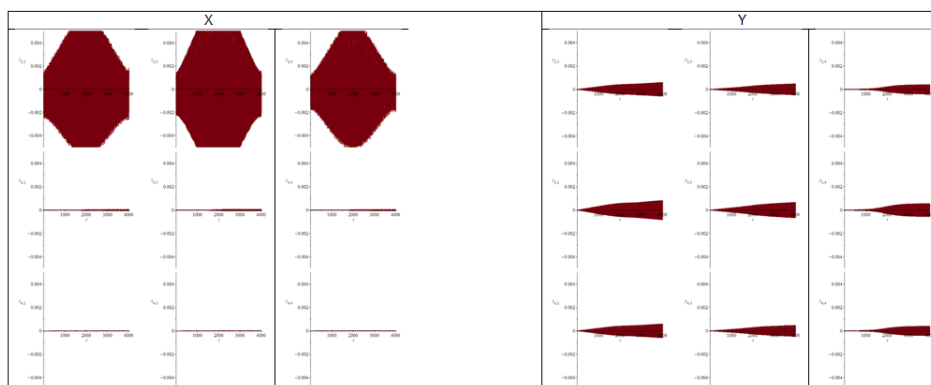
Резонансное возбуждение нормальной моды с частотой $\omega = 1.0225$ 

Таблица 2

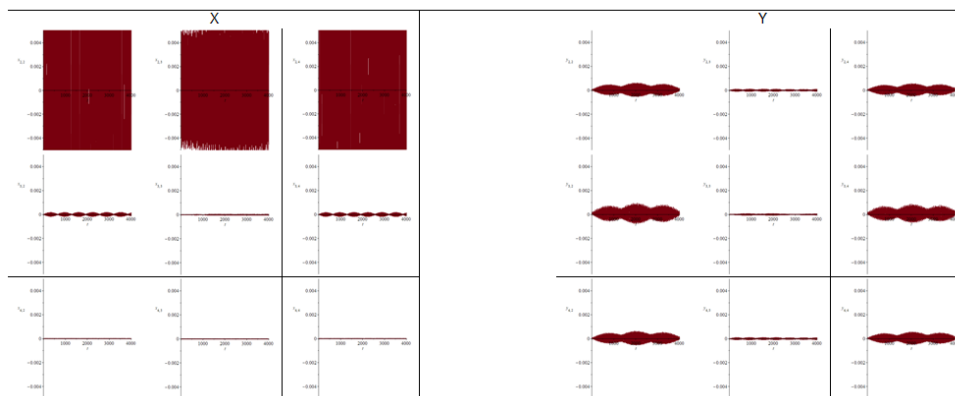
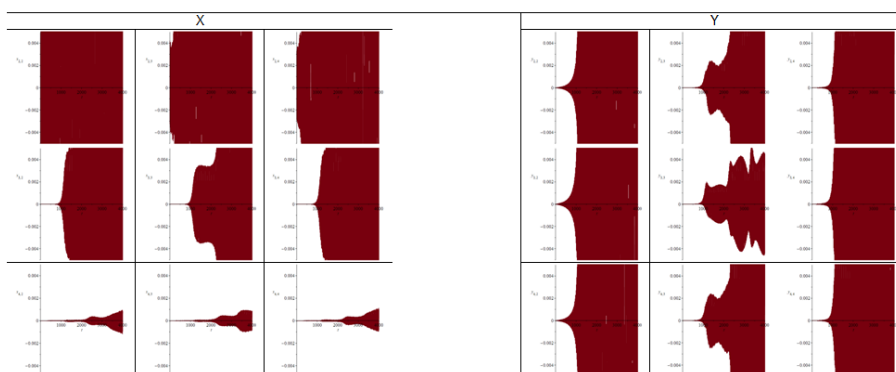
Резонансное возбуждение нормальной моды с частотой $\omega = 3.882$ 

Таблица 3

Резонансное возбуждение нормальной моды с частотой $\omega = 5.831$



Тут и далее, первая таблица иллюстрирует колебания атомов от положения равновесия вдоль оси X, вторая – вдоль оси Y.

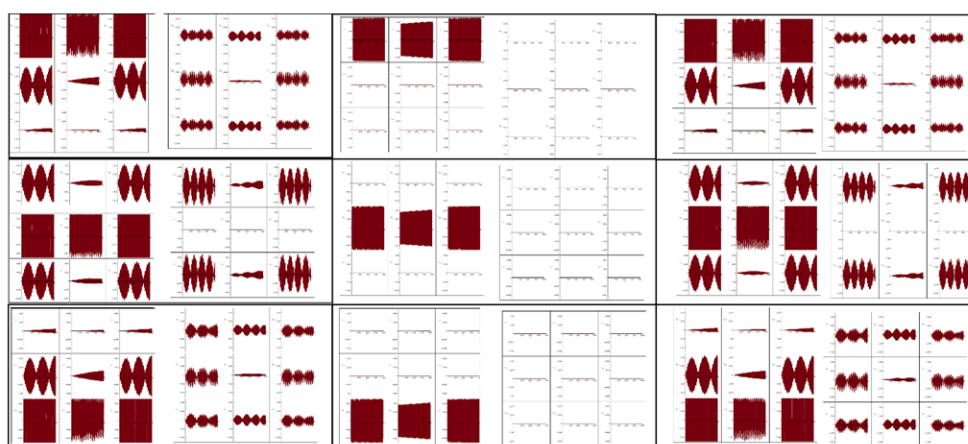
В табл. 1–4 возбуждение производится в плоской квадратной атомной решётке Леннарда-Джонса 5x5 и закреплёнными краями. В табл. 1–3 – резонансное возбуждение производится путём приложения внешней периодической силы в направлении [1,0] к атому (2,2).

При приложении внешней периодической силы, с не резонансной частотой, возбуждение по системе передаваться всё же будет, однако исключительно вдоль направления оси этой силы.

Точно так же хорошо видна и Z-структура нормальных мод (Z-структура – расположение всех нулевых элементов паттерна атомных смещений нелинейной нормальной моды). К примеру, если увеличить масштаб табл. 2 и подействовать $\omega = 3.882$ на каждый подвижный атом системы (табл. 4), то можно будет увидеть, что при прикладывании внешней периодической силы к атомам (2,3), (3,3), (4,3) возбуждение в системе передаётся лишь вдоль оси этой силы.

Таблица 4

Резонансное возбуждение нормальной моды с частотой $\omega = 3.882$ в, проведённое путём приложения внешней периодической силы в направлении [1,0] поочерёдно к каждому подвижному атому решётки



Заключение. Данная работа – посвящена вопроса возбуждения бушей нелинейных нормальных мод, теория которых развивается в ЮФУ и УрФУ несколько последних десятилетий [17–34].

Каждый буш представляет собой совокупность колебательных мод, которая сохраняется во времени и определяет некоторое точное решение нелинейных уравнений движения, которые нельзя получить в рамках теории возмущений.

В работе передоложена математическая модель позволяющая произвести возбуждение всех возможных бушей в квадратной двумерной решётке. Предполагается, что указанная схема возбуждения бушей мод могут быть использованы для реального возбуждения бушей в физических системах с дискретной симметрией

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. *Fermi E., Pasta J., Ulam S.M.* Studies of nonlinear problems I // Technical Report LA-1940, LANL. – May 1955. – P. 978-988.
2. *Chechin G.M., Sakhnenko V.P.* Interactions between normal modes in nonlinear dynamical systems with discrete symmetry // Exact results Phys. D. – 1998. – 117, 43.
3. *Sakhnenko V.P., Chechin G.M.* Symmetrical selection rules in nonlinear dynamics of atomic systems // Phys. Dokl. – 1993. – 38, 219.
4. *Chechin G.M., Ryabov D.S.* Introduction to the theory of bushes of nonlinear normal modes for studying large-amplitude atomic vibrations in systems with discrete symmetry // Lett. Mater. – 2020. – 10 (4). – P. 523-534.
5. *Rink B.* Symmetry and resonance in periodic FPU chains // Physica D. – 2001. – Vol. 218, No. 3. – P. 665-668.
6. *Stokes H.T., Hatch D.M., Campbell B.J.* ISOTROPY Software Suite. – <https://iso.byu.edu/>.
7. *Stokes H.T., Smith A.D., Hatch D.M., Chechin G.M., Sakhnenko V.P.* Group theoretical methods applied to non-linear dynamics in crystals // International journal of non-linear mechanics. – 1998. – 35 (3). – P. 497-513.
8. *Yoshimura K., Kolchinsky A., Dechant A.* Modulational instability of zone boundary mode in nonlinear lattices // Phys. Rev. Research. – 2023. – 5, 013017. – <https://doi.org/10.1103/PhysRevResearch.5.013017>.
9. *Nakamura S., Shinohara S., et al.* Nonlinear Scar Modes Generated by Self-Organization of Resonance Modes // Proc. NOLTA. – Catania, 2023. – P. 682-682. – DOI:10.34385/proc.76.D1L-34.
10. *Wigner E.* Über die elastischen Eigenschwingungen symmetrischer systeme // Nachricht. Akad. Wiss. Göttingen, Math.-Phys. Kl. – Berlin, 1930.
11. *Petrashen M.I., Trifonov E.D.* Applications of group theory in quantum mechanics. – Cambridge, MIT Press, 1969.
12. *Аврамов К.В., Михлин Ю.В.* Нелинейная динамика упругих систем // Модели, методы, явления. – М.: Ижевск: Институт компьютерных исследований, 2019. – 715 с.
13. *Корзникова Е.А., Бокий Д.И., Дмитриев С.В.* Нелинейная динамика гексагональной решетки. – СПб.: Изд-во «Наукоемкие технологии», 2018. – 114 с.
14. *Zabusky N.J., Kruskal M.D.* Interaction of “solitons” in a collisionless plasma and the recurrence of initial states // Phys. Rev. Lett. – 1965. – 15 (6). – P. 240-243.
15. *Khasseh R., Fazio R., Ruffo S., & Russomanno A.* Many-body synchronization in a classical hamiltonian system // Physical Review Letters. – 2019. – 123 (18), 184301. – <https://doi.org/10.1103/PhysRevLett.123.184301>.
16. *Trengginas E.P., Sutantyo et al.* Chaotic behaviour of modified hamiltonian Peyrard-Bishop-Dauxois model on DNA system // J. Phys.: Conf. Ser. – 1245 012070. – DOI 10.1088/1742-6596/1245/1/012070.
17. *Lobzenko I.P., Chechin G.M., Bezuglova G.S. et al.* Ab initio simulation of gap discrete breathers in strained grapheme // Phys. Solid State. – 2016. – 58. – P. 633-639. – <https://doi.org/10.1134/S1063783416030203>.
18. *Ryabov D.S., Chechin G.M., Naumov E.K. et al.* One-component delocalized nonlinear vibrational modes of square lattices // Nonlinear Dynamics. – 2023. – 111. – P. 8135-8153. – <https://doi.org/10.1007/s11071-023-08264-6>.
19. *Korznikova E.A., Shcherbinin S.A., Ryabov D.S., Chechin G.M., Ekomasov E.G., Barani E., Zhou K., Dmitriev S.V.* Delocalized nonlinear vibrational modes in graphene: second harmonic generation and negative pressure // Physica Status Solidi (b). – 2019. – 256, 1800061. – <https://doi.org/10.1002/pssb.201800061>.
20. *Ryabov D.S., Chechin G.M., Upadhyaya A. et al.* Delocalized nonlinear vibrational modes of triangular lattices // Nonlinear Dynamics. – 2020. – 102. – P. 2793-2810. – <https://doi.org/10.1007/s11071-020-06015-5>.
21. *Shcherbinin S.A., Semenova M.N., Semenov A.S. et al.* Dynamics of a three-component delocalized nonlinear vibrational mode in graphene // Phys. Solid State. – 2019. – 61. – P. 2139-2144. – <https://doi.org/10.1134/S1063783419110313>.
22. *Abdullina D.U., Semagina A.S., Ryabov D.S., Chechin G.M., Korznikova E.A., Baimova J.A., & Dmitriev S.V.* Stability of in-plane delocalized vibrational modes in triangular Morse lattice // IOP Conference Series: Materials Science and Engineering. – 2018. – 447 (1), 012060. – <https://doi.org/10.1088/1757-899X/447/1/012060>.

23. Chechin G.M., & Lapina V.S. Discrete breathers of new type in monoatomic chains // arXiv. – 2018. – <https://doi.org/10.48550/arXiv.1809.08901>.
24. Chechin G.M., & Lapina V.S. Universal bi-structures of strained monoatomic chains // Computational Materials Science. – 2020. – 177, 109578. – <https://doi.org/10.1016/j.commatsci.2019.109578>.
25. Chechin G., Ryabov D., Shcherbinin S. Large-amplitude in-plane atomic vibrations in strained graphene monolayer: bushes of nonlinear normal modes // Letters on Materials. – 2017. – 7 (4). – P. 367-372. – DOI: 10.17580/ism.2017.04.07.
26. Баимова Ю.А., Щербинин С.А., Чечин Г.М., Дмитриев С.В. Об устойчивости одномерных бушей нелинейных колебательных мод в графене // Letters on Materials. – 2018. – 8 (1). – P. 43-48. – <https://doi.org/10.22226/2410-3535-2018-1-43-48>.
27. Shcherbinin S.A., Krylova K.A., Chechin G.M., Soboleva E.G., Dmitriev S.V. Delocalized nonlinear vibrational modes in fcc metals // Communications in Nonlinear Science and Numerical Simulation. – 2021. – Vol. 105. – P. 106083. – DOI: 10.1016/j.cnsns.2021.106083.
28. Кедров А.А., Щербинин С.А. Исследование устойчивости колебательной π -моды в возмущенной цепочке Тоды // Известия высших учебных заведений. Прикладная нелинейная динамика. – 2026. – Т. 34, № 2. – С. 206-222. – DOI: 10.18500/0869-6632-003202. – EDN: NRVWDG.
29. Ryabov D.S., Kosarev I.V., Xiong D., Kudreyko A.A., Dmitriev S.V. Delocalized Nonlinear Vibrational Modes in Bcc Lattice for Testing and Improving Interatomic Potentials // Computers, Materials and Continua. – 2025. – 82 (3). – P. 3797-3820. – <https://doi.org/10.32604/cmc.2025.062079>.
30. Ryabov D.S., Bezuglova G.S., Korznikova E.A., & Dmitriev S.V. Testing interatomic potentials for binary alloys using exact solutions to the equations of atomic motion // Computational Materials Science. – 2024. – 242, 113054. – <https://doi.org/10.1016/j.commatsci.2024.113054>.
31. Косарев И.В., Абдуллина Д.В., Дмитриев С.В. Делокализованные нелинейные колебательные моды ОЦК решетки // Известия Алтайского государственного университета. – 2025. – № 1 (141). – С. 58-66. – DOI: 10.14258/izvasu(2025)1-07.
32. Chechin G., Ryabov D. Exact solutions of nonlinear dynamical equations for large-amplitude atomic vibrations in arbitrary monoatomic chains with fixed ends // arXiv:2109.09101v1. – 2021. – <https://doi.org/10.1016/j.cnsns.2023.107176>.
33. Семенов А.С., Рябов Д.С., Чечин Г.М., Дмитриев С.В. Локализованные и делокализованные нелинейные колебательные моды треугольной решетки // Комплексный анализ, математическая физика и нелинейные уравнения. – 2021. – P. 64-65.
34. Chechin G., Ryabov D. Properties of large-amplitudes vibrations in dynamical systems with discrete symmetry. Geometrical aspects // arXiv:2005.13343v1. – 2020. – <https://arxiv.org/abs/2005.13343v1>.

REFERENCES

1. Fermi E., Pasta J., Ulam S.M. Studies of nonlinear problems I, *Technical Report LA-1940*, LANL, May 1955, pp. 978-988.
2. Chechin G.M., Sakhnenko V.P. Interactions between normal modes in nonlinear dynamical systems with discrete symmetry, *Exact results Phys. D*, 1998, 117, 43.
3. Sakhnenko V.P., Chechin G.M. Symmetrical selection rules in nonlinear dynamics of atomic systems, *Phys. Dokl.*, 1993, 38, 219.
4. Chechin G.M., Ryabov D.S. Introduction to the theory of bushes of nonlinear normal modes for studying large-amplitude atomic vibrations in systems with discrete symmetry, *Lett. Mater.*, 2020, 10 (4), pp. 523-534.
5. Rink B. Symmetry and resonance in periodic FPU chains, *Physica D*, 2001, Vol. 218, No. 3, pp. 665-668.
6. Stokes H.T., Hatch D.M., Campbell B.J. ISOTROPY Software Suite. Available at: <https://iso.byu.edu/>.
7. Stokes H.T., Smith A.D., Hatch D.M., Chechin G.M., Sakhnenko V.P. Group theoretical methods applied to non-linear dynamics in crystals, *International journal of non-linear mechanics*, 1998, 35 (3), pp. 497-513.
8. Yoshimura K., Kolchinsky A., Dechant A. Modulational instability of zone boundary mode in nonlinear lattices, *Phys. Rev. Research*, 2023, 5, 013017. Available at: <https://doi.org/10.1103/PhysRevResearch.5.013017>.
9. Nakamura S., Shinohara S., et al. Nonlinear Scar Modes Generated by Self-Organization of Resonance Modes, *Proc. NOLTA*. Catania, 2023, pp. 682-682. DOI: 10.34385/proc.76.D1L-34.
10. Wigner E. Über die elastischen Eigenschwingungen symmetrischer systeme, *Nachricht. Akad. Wiss. Göttingen, Math.-Phys. Kl.* Berlin, 1930.
11. Petrashen M.I., Trifonov E.D. Applications of group theory in quantum mechanics. Cambridge, MIT Press, 1969.

12. Avramov K.V., Mikhlin Yu.V. Nelineynaya dinamika uprugikh sistem [Nonlinear dynamics of elastic systems], *Modeli, metody, yavleniya* [Models, Methods, Phenomena]. Moscow: Izhevsk: Institut komp'yuternykh issledovaniy, 2019, 715 p.
13. Korznikova E.A., Bokiy D.I., Dmitriev S.V. Nelineynaya dinamika geksgonal'noy reshetki [Nonlinear dynamics of a hexagonal lattice]. Saint Petersburg: Izd-vo «Naukoemkie tekhnologii», 2018, 114 p.
14. Zabusky N.J., Kruskal M.D. Interaction of “solitons” in a collisionless plasma and the recurrence of initial states, *Phys. Rev. Lett.*, 1965, 15 (6), pp. 240-243.
15. Khasseh R., Fazio R., Ruffo S., & Russomanno A. Many-Body Synchronization in a Classical Hamiltonian system, *Physical Review Letters*, 2019, 123 (18), 184301. Available at: <https://doi.org/10.1103/PhysRevLett.123.184301>.
16. Trengginas E.P. Sutantyo et al. Chaotic behaviour of modified hamiltonian Peyrard-Bishop-Dauxois model on DNA system, *J. Phys.: Conf. Ser.*, 1245 012070. DOI 10.1088/1742-6596/1245/1/012070.
17. Lobzenko I.P., Chechin G.M., Bezuglova G.S. et al. Ab initio simulation of gap discrete breathers in strained grapheme, *Phys. Solid State*, 2016, 58, pp. 633-639. Available at: <https://doi.org/10.1134/S1063783416030203>.
18. Ryabov D.S., Chechin G.M., Naumov E.K. et al. One-component delocalized nonlinear vibrational modes of square lattices, *Nonlinear Dynamics*, 2023, 111, pp. 8135-8153. Available at: <https://doi.org/10.1007/s11071-023-08264-6>.
19. Korznikova E.A., Shcherbinin S.A., Ryabov D.S., Chechin G.M., Ekomasov E.G., Barani E., Zhou K., Dmitriev S.V. Delocalized nonlinear vibrational modes in graphene: second harmonic generation and negative pressure, *Physica Status Solidi (b)*, 2019, 256, 1800061. Available at: <https://doi.org/10.1002/pssb.201800061>.
20. Ryabov D.S., Chechin G.M., Upadhyaya A. et al. Delocalized nonlinear vibrational modes of triangular lattices, *Nonlinear Dynamics*, 2020, 102, pp. 2793-2810. Available at: <https://doi.org/10.1007/s11071-020-06015-5>.
21. Shcherbinin S.A., Semenova M.N., Semenov A.S. et al. Dynamics of a three-component delocalized nonlinear vibrational mode in grapheme, *Phys. Solid State*, 2019, 61, pp. 2139-2144. Available at: <https://doi.org/10.1134/S1063783419110313>.
22. Abdullina D.U., Semagina A.S., Ryabov D.S., Chechin G.M., Korznikova E.A., Baimova J.A., & Dmitriev S.V. Stability of in-plane delocalized vibrational modes in triangular Morse lattice, *IOP Conference Series: Materials Science and Engineering*, 2018, 447 (1), 012060. Available at: <https://doi.org/10.1088/1757-899X/447/1/012060>.
23. Chechin G.M., & Lapina V.S. Discrete breathers of new type in monoatomic chains, *arXiv*, 2018. Available at: <https://doi.org/10.48550/arXiv.1809.08901>.
24. Chechin G.M., & Lapina V.S. Universal bi-structures of strained monoatomic chains, *Computational Materials Science*, 2020, 177, 109578. Available at: <https://doi.org/10.1016/j.commatsci.2019.109578>.
25. Chechin G., Ryabov D., Shcherbinin S. Large-amplitude in-plane atomic vibrations in strained graphene monolayer: bushes of nonlinear normal modes, *Letters on Materials*, 2017. – 7 (4), pp. 367-372. DOI: 10.17580/ism.2017.04.07.
26. Baimova Yu.A., Shcherbinin S.A., Chechin G.M., Dmitriev S.V. Ob ustoychivosti odnomernykh bushey nelineynykh kolebatel'nykh mod v grafene [On the stability of one-dimensional bushes of nonlinear vibrational modes in graphene], *Letters on Materials*, 2018, 8 (1), pp. 43-48. Available at: <https://doi.org/10.22226/2410-3535-2018-1-43-48>.
27. Shcherbinin S.A., Krylova K.A., Chechin G.M., Soboleva E.G., Dmitriev S.V. Delocalized nonlinear vibrational modes in fcc metals, *Communications in Nonlinear Science and Numerical Simulation*, 2021, Vol. 105, pp. 106083. DOI: 10.1016/j.cnsns.2021.106083.
28. Kedrov A.A., Shcherbinin S.A. Issledovanie ustoychivosti kolebatel'noy π -mody v vozmushchennoy tsepoche Tody [Study of stability of oscillatory π -mode in perturbed Toda chain], *Izvestiya vysshikh uchebnykh zavedeniy. Prikladnaya nelineynaya dinamika* [Bulletin of higher educational institutions. Applied nonlinear dynamics], 2026, Vol. 34, No. 2, pp. 206-222. DOI: 10.18500/0869-6632-003202. EDN: NRVWDG.
29. Ryabov D.S., Kosarev I.V., Xiong D., Kudreyko A.A., Dmitriev S.V. Delocalized Nonlinear Vibrational Modes in Bcc Lattice for Testing and Improving Interatomic Potentials, *Computers, Materials and Continua*, 2025, 82 (3), pp. 3797-3820. Available at: <https://doi.org/10.32604/cmc.2025.062079>.
30. Ryabov D.S., Bezuglova G.S., Korznikova E.A., & Dmitriev S.V. Testing interatomic potentials for binary alloys using exact solutions to the equations of atomic motion, *Computational Materials Science*, 2024, 242, 113054. Available at: <https://doi.org/10.1016/j.commatsci.2024.113054>.
31. Kosarev I.V., Abdullina D.U., Dmitriev S.V. Delokalizovannyye nelineynyye kolebatel'nyye mody OTSK reshetki [Delocalized nonlinear vibrational modes of the bcc lattice], *Izvestiya Altayskogo gosudarstvennogo universiteta* [Bulletin of the Altai State University], 2025, No. 1 (141), pp. 58-66. DOI: 10.14258/izvasu(2025)1-07.

32. Chechin G., Ryabov D. Exact solutions of nonlinear dynamical equations for large-amplitude atomic vibrations in arbitrary monoatomic chains with fixed ends, *arXiv:2109.09101v1*, 2021. Available at: <https://doi.org/10.1016/j.cnsns.2023.107176>.
33. Semenov A.S., Ryabov D.S., Chechin G.M., Dmitriev S.V. Lokalizovannye i delokalizovannye nelineynye kolebatel'nye mody treugol'noy reshetki [Localized and delocalized nonlinear vibrational modes of a triangular lattice], *Kompleksnyy analiz, matematicheskaya fizika i nelineynye uravneniya* [Complex Analysis, Mathematical Physics and Nonlinear Equations], 2021, pp. 64-65.
34. Chechin G., Ryabov D. Properties of large-amplitudes vibrations in dynamical systems with discrete symmetry. Geometrical aspects, *arXiv:2005.13343v1*, 2020. Available at: <https://arxiv.org/abs/2005.13343v1>.

Бурлаченко Иван Сергеевич – Южный федеральный университет; e-mail: iburlachenko@sfedu.ru; г. Таганрог, Россия; тел.: +79054395264; кафедра физико-математических основ инженерного образования; ассистент.

Чечин Георгий Михайлович – НИИ Физики, Южный федеральный университет; e-mail: gmchechin@sfedu.ru; г. Ростов-на-Дону, Россия; тел.: +79054315081; с.н.с.; к.ф.-м.н.

Burlachenko Ivan Sergeevich – Southern Federal University; e-mail: iburlachenko@sfedu.ru; Taganrog, Russia; phone: +79054395264; the Department of Physical and Mathematical Foundations of Engineering Education; assistant.

Chechin Georgy Mikhailovich – Research Institute of Physics, Southern Federal University; e-mail: gmchechin@sfedu.ru; Rostov-on-Don, Russia, phone: +79054315081; senior researcher; cand. phys. and math. sc.

УДК 003.26

DOI 10.18522/2311-3103-2026-4-201-212

Ж.-М.Н. Дакуо, С.В. Беззатеев, В.В. Давыдов, И.Д. Иогансон, А.Ф. Хуцаева

МОДЕЛЬ СХЕМЫ ЭЛЕКТРОННОГО ГОЛОСОВАНИЯ С ВОЗМОЖНОСТЬЮ ИЗМЕНЕНИЯ РЕШЕНИЯ УЧАСТНИКОМ ДО МОМЕНТА ОКОНЧАНИЯ ГОЛОСОВАНИЯ

Существующие протоколы электронного голосования, как правило, обладают архитектурой, в которой отправка бюллетеня является необратимым действием. Такой подход ограничивает удобство использования системы, поскольку участник не может исправить случайную ошибку или изменить свое решение в течение периода голосования без прохождения повторной процедуры аутентификации либо аннулирования ранее поданного бюллетеня. Целью настоящей работы является разработка теоретической модели криптографического протокола электронного голосования, обеспечивающей нативную возможность многократного изменения решения для повышения удобства использования при сохранении анонимности и корректности подсчета. Архитектура предлагаемой модели основана на совместном использовании схемы обязательств с потайным входом для реализации свойства изменения решения, проверяемой случайной функции (VRF) для формирования бюллетеней, протокола Privacy Pass для анонимной авторизации участников и смешанных сетей (Mix-Nets), обеспечивающих сокрытие связи между пользователем и передаваемыми сообщениями. Разработан протокол электронного голосования, позволяющий участнику неоднократно изменять свое решение без повторного обращения к центру регистрации после получения права на участие в голосовании, что уменьшает количество интерактивных действий и повышает удобство использования системы. Проведен теоретический анализ стойкости предложенной модели, подтверждающий выполнение основных и ряда расширенных свойств, включая, полноту, устойчивость, приемлемость, проверяемость, конфиденциальность участника, одноразовость, справедливость, атомарную и универсальную проверяемость, а также возможность изменения решения. Благодаря модульной архитектуре модель допускает применение различных криптографических примитивов, включая постквантовые алгоритмы, без изменения общей логики функционирования протокола, что делает предложенный подход перспективной основой для построения защищенных систем электронного голосования, стойких к атакам квантового компьютера.

Электронное голосование; схема обязательств с потайным входом; изменение решения; проверяемая случайная функция; универсальная проверяемость; Privacy Pass; анонимность.

Z.-M.N. Dakuo, S.V. Bezzateev, V.V. Davydov, I.D. Ioganson, A.F. Khutsaeva

MODEL OF AN ELECTRONIC VOTING SCHEME WITH THE POSSIBILITY OF CHANGING THE PARTICIPANT'S DECISION BEFORE THE END OF VOTING

Existing electronic voting protocols often feature an architecture where ballot submission is an irreversible action. This significantly reduces system usability, depriving the participant of the right to correct an accidental error or change their opinion during lengthy decision-making procedures. The aim of the work is to develop a model of a cryptographic protocol that provides a native capability for multiple decision changes to enhance usability while maintaining strict anonymity and counting correctness. The architecture is based on the integration of a trapdoor commitment scheme to implement choice flexibility, a verifiable random function (VRF) for ballot generation, the Privacy Pass protocol for anonymous authorization, and mix-networks. An electronic voting protocol has been developed that implements the possibility for a participant to change their decision multiple times. The fulfillment of basic and some extended security properties, including, but not limited to, participant confidentiality and universal verifiability, has been theoretically proven. It is shown that, unlike existing analogs, changing a decision in the proposed model does not require repeating a complex authentication procedure or key revocation, which minimizes the number of interactive user actions. The proposed solution aligns the user experience of secure voting with the standards of modern digital services by providing a choice editing function, thereby increasing the convenience of the voting procedure for participants. The modular structure of the model allows for the application of any robust cryptographic algorithms, including post-quantum ones.

Electronic voting; trapdoor commitment scheme; revisability; verifiable random function; universal verifiability; Privacy Pass; anonymity.

Введение. В условиях цифровой трансформации общества системы электронного голосования (ЭГ) становятся широко используемым инструментом для дистанционного принятия коллективных решений. Данные технологии находят применение не только в государственных избирательных процессах, но и в корпоративном управлении, деятельности акционерных обществ и децентрализованных автономных организаций. Однако переход к цифровому формату требует реализации сложных криптографических протоколов, способных обеспечить уровень доверия и безопасности, сопоставимый с традиционным голосованием.

Проектирование систем ЭГ требует соблюдения комплекса обязательных условий. В современной теории [1–3] эти требования классифицируются на три группы: **основные, расширенные**, а также **эксплуатационные**, отражающие практическую и удобство использования системы.

К основным требованиям относят:

- ◆ **Полнота (Completeness):** все действительные решения участников подсчитываются правильно.
- ◆ **Устойчивость (Soundness):** недействительные решения участников не учитываются.
- ◆ **Приемлемость (Eligibility):** только легитимные участники могут принимать участие в голосовании.
- ◆ **Проверяемость (Verifiability):** никто не может фальсифицировать результаты голосования.
- ◆ **Конфиденциальность участника (Privacy):** решение отдельного участника хранится в тайне.
- ◆ **Одноразовость (Unreusability):** решение каждого участника учитывается единожды.
- ◆ **Справедливость (Fairness):** ничто не должно влиять на голосование, т. е. никакие промежуточные результаты никому не могут быть известны.

К расширенным требованиям относят:

- ◆ **Надежность (Robustness):** никакая коалиция участников или организаторов разумного размера не может сорвать процедуру голосования.
- ◆ **Атомарная проверяемость (Atomic verifiability):** отдельные участники могут убедиться, что их решения были подсчитаны должным образом.

- ♦ *Универсальная проверяемость (Universal verifiability)*: любой может проверить результаты подсчета.
 - ♦ *Отсутствие квитанций (Receipt-freeness)*: ни один участник не может составить квитанцию, подтверждающую содержание его бюллетеня.
 - ♦ *Декларируемость (Declarability)*: можно проверить, проголосовал ли конкретный участник.
 - ♦ *Непринуждаемость (Uncoercibility)*: невозможность злоумышленника определить, соответствует ли участник, подвергнутый принуждению, его требованиям.
 - ♦ *Изменения решения (Revisability)*: возможность изменения участником своего решения в течение периода голосования.
- К свойствам практичности относят:*
- ♦ *Гибкость (Flexibility)*: допускается множество форматов бюллетеней.
 - ♦ *Эффективность (Efficiency)*: процедура голосования использует разумное количество ресурсов.
 - ♦ *Мобильность (Mobility)*: участники могут принимать решение из любого места.
 - ♦ *Удобство (Convenience)*: процедура голосования состоит из одного раунда общения с организаторами.

Развитие протоколов электронного голосования началось в 1981 году с работы Дэвида Чаума [4], предложившего использование смешанных сетей (Mix-Nets) для обеспечения анонимности каналов связи. В 1992 году А. Фудзиока, Т. Окамото и К. Ота представили протокол [5] с использованием слепых цифровых подписей, что позволило эффективно разделить процессы регистрации и подачи голоса.

В последующие десятилетия научное сообщество активно разрабатывало альтернативные теоретические модели, направленные на решение узкоспециализированных задач. Так, схема MIZAN [6] стала одной из первых попыток реализовать свойство изменения решения в рамках традиционной PKI-инфраструктуры. С развитием технологий распределенного реестра появились протоколы на базе технологии блокчейн [7, 8], предлагающие децентрализованное хранение бюллетеней для повышения доверия к системе без единого администратора. Особое внимание уделяется защите от квантовых угроз: в частности, предлагаются усовершенствованные смешанные сети на основе криптографии на решетках [9]. Параллельно с разработкой архитектур развивается и теоретический базис: исследователи формализуют понятия проверяемости [10], создавая строгие критерии для оценки стойкости протоколов. В отечественной науке также существуют оригинальные подходы, например схема защищенного голосования с использованием «слепых посредников» [11].

В настоящее время системы ЭГ активно эксплуатируются на государственном уровне. Первой страной, внедрившей общенациональное интернет-голосование, стала Эстония в 2005 году [12]. В Российской Федерации технологии дистанционного электронного голосования (ДЭГ) применяются с 2019 года, а в 2024 году масштабно использовались в ходе выборов Президента РФ [13].

Несмотря на широкое распространение, значительная часть существующих протоколов обладает существенным ограничением: в таких системах изменение решения участником либо технически невозможно, либо требует сложной процедуры повторной аутентификации и аннулирования предыдущего бюллетеня, что снижает эффективность системы. При этом свойство изменения решения является критически важным для защиты от принуждения. Если протокол позволяет многократно перезаписывать голос, то участник, находящийся под давлением злоумышленника, может отдать требуемый голос, а затем, оказавшись в безопасности, изменить его на истинный. Отсутствие удобного механизма смены решения в существующих моделях является актуальной научной проблемой.

Отсутствие механизма простой процедуры изменения решения не учитывает специфику современных цифровых процедур. Во-первых, существенную роль играет временной фактор: электронные процедуры часто растянуты во времени и могут продолжаться несколько дней. За этот период внешние обстоятельства могут измениться, или

участник может получить новую существенную информацию, влияющую на его позицию. Во-вторых, нельзя исключать влияние человеческого фактора – участник может совершить случайную ошибку. В системе, не предусматривающей возможность коррективки решения, такая ошибка становится фатальной.

Целью данной работы является разработка модели протокола электронного голосования, обеспечивающего нативную возможность изменения решения участником без компрометации анонимности. В основе предлагаемой модели лежит семейство криптографических схем обязательства с потайным входом [14].

Следует отметить, что работа носит теоретический характер и посвящена построению алгоритмической модели взаимодействия участников. Предложенная архитектура является универсальной: выбор конкретных криптографических примитивов может варьироваться в зависимости от требований реализации, в том числе с учетом необходимости защиты от квантовых угроз.

Предварительные сведения. В данном разделе введем формальные определения криптографических примитивов, используемых для построения предлагаемой модели протокола.

Симметричная схема шифрования представляет собой набор алгоритмов, позволяющих двум сторонам, обладающим общим секретным ключом, обмениваться конфиденциальной информацией. В контексте отечественных стандартов в качестве базового примитива может использоваться блочный шифр «Кузнечик», регламентированный ГОСТ Р 34.12-2015 [15].

Формально симметричная криптосистема Π_{sym} определяется как кортеж из трех полиномиальных алгоритмов ($KeyGen, Enc, Dec$):

1. $KeyGen(1^\lambda) \rightarrow k$: алгоритм, который принимает на вход параметр безопасности λ и возвращает секретный ключ k .
2. $Enc(k, m) \rightarrow c$: алгоритм, который принимает на вход секретный ключ k и сообщение $m \in \mathcal{M}$ где $\mathcal{M}$ – пространство сообщений, и возвращает шифртекст c .
3. $Dec(k, c) \rightarrow m$: алгоритм, который принимает на вход секретный ключ k и шифртекст c . Он возвращает сообщение m или ошибку $\perp$, если шифртекст некорректен.

Для корректности схемы требуется, чтобы для любого ключа k , выработанного $KeyGen(1^\lambda)$, и любого сообщения $m \in \mathcal{M}$ выполнялось равенство:

$$Dec(k, Enc(k, m)) = m.$$

Асимметричные криптосистемы позволяют осуществлять шифрование и расшифрование данных с использованием пары математически связанных ключей: открытого и закрытого. Стойкость таких систем основывается на вычислительной сложности решения определенных математических задач. Классическим примером асимметричной криптосистемы шифрования является работа Ривеста, Шамира и Адельмана [16].

Формально асимметричная криптосистема Π_{asym} задается кортежем из трех алгоритмов ($KeyGen, Enc, Dec$):

1. $KeyGen(1^\lambda) \rightarrow (pk, sk)$: вероятностный алгоритм, который по параметру безопасности λ вырабатывает пару ключей: открытый ключ pk и закрытый ключ sk .
2. $Enc(pk, m) \rightarrow c$: вероятностный алгоритм, который принимает на вход открытый ключ pk и сообщение m , возвращая шифртекст c .
3. $Dec(sk, c) \rightarrow m$: детерминированный алгоритм, который использует закрытый ключ sk , и возвращает исходное сообщение m или ошибку $\perp$.

Условие корректности аналогично симметричному случаю: для любой пары ключей (pk, sk) и любого сообщения m :

$$Dec(sk, Enc(pk, m)) = m.$$

В предлагаемой модели асимметричное шифрование используется для обеспечения конфиденциальности каналов связи между участниками и серверами (например, при передаче бюллетеней через Mix-Nets), а также применяется при процедуре аутентификации сторон.

Проверяемая случайная функция (Verifiable Random Function, VRF) – это криптографический примитив, который позволяет владельцу секретного ключа вычислять псевдослучайное значение для заданного входного параметра и генерировать общедоступное доказательство корректности этого вычисления [17]. В контексте предлагаемой модели данный примитив используется для формирования анонимного бюллетеня с возможностью последующего аудита.

Формально схема VRF определяется как кортеж из трех полиномиальных алгоритмов ($VRF.keyGen, VRF.eval, VRF.verify$):

1. $VRF.keyGen(1^\lambda) \rightarrow (vpk, vsk)$. Вероятностный алгоритм, принимающий параметр безопасности λ и возвращающий пару ключей: открытый pk и закрытый sk .
2. $VRF.eval(vsk, m) \rightarrow (v, proof)$. алгоритм, состоящий из пары функций F_1 и F_2 , принимающий на вход сообщение m и закрытый ключ vsk . Алгоритм возвращает двоичные строки $v = F_1(vsk, m)$ и $proof = F_2(vsk, m)$.
3. $VRF.verify(vpk, m, v, proof) \rightarrow 1/0$: детерминированный алгоритм, принимающий открытый ключ vpk , сообщение m , псевдослучайное значение v и доказательство $proof$. Возвращает либо 1 (проверка успешна), либо 0 (проверка неуспешна).

Для криптосистемы VRF должны выполняться следующие свойства:

- ♦ *Корректность (Correctness)*: для любой пары ключей (vpk, vsk) и любого сообщения m , если $(v, proof) \leftarrow VRF.eval(vsk, m)$, то проверка $VRF.verify(vpk, m, v, proof)$ всегда возвращает 1.

- ♦ *Уникальная доказуемость (Unique Provability)*: для любого открытого ключа vpk и любого сообщения m вычислительно невозможно найти такие значения $v_1 \neq v_2$ и доказательства $proof_1, proof_2$, что $VRF.verify(vpk, m, v_1, proof_1) = 1$ и $VRF.verify(vpk, m, v_2, proof_2) = 1$.

- ♦ *Псевдослучайность (Pseudorandomness)*: для любого вероятностного полиномиального злоумышленника, не владеющего секретным ключом vsk , и для любого сообщения m , для которого злоумышленник не запрашивал значение $VRF.eval(vsk, m)$, выходное значение v вычислительно неотличимо от равномерно случайного значения.

Схема обязательств с потайным входом (Trapdoor Commitment, TC) – это криптосистема, являющаяся модификацией классической схемы обязательств, которая позволяет зафиксировать значение, а впоследствии изменить его при наличии секретного ключа, не меняя при этом самого обязательства [14].

Пусть $\mathcal{M}$ – пространство сообщений, $\mathcal{R}$ – пространство случайностей, $\mathcal{K}$ – пространство ключей и $\mathcal{C}$ – пространство обязательств.

Тогда схема обязательств с потайным входом TC определяется как кортеж из пяти полиномиальных алгоритмов:

1. $TC.parGen(1^\lambda) \rightarrow pp$: вероятностный алгоритм, принимающий на вход параметр безопасности λ и возвращающий набор открытых параметров pp .
2. $TC.keyGen(pp) \rightarrow (cpk, csk)$: алгоритм, принимающий параметры pp и возвращающий пару ключей: открытый ключ $cpk \in \mathcal{K}_{pk}$ и закрытый ключ $csk \in \mathcal{K}_{sk}$.
3. $TC.tCom(cpk, m) \rightarrow (c, r)$: алгоритм принимает открытый ключ cpk и сообщение $m \in \mathcal{M}$. Возвращает обязательство $c \in \mathcal{C}$ и случайность $r \in \mathcal{R}$.
4. $TC.tVer(cpk, c, m, r) \rightarrow 1/0$: детерминированный алгоритм, который проверяет соответствие обязательства c паре (m, r) . Возвращает 1, если $c = TC.tCom(cpk, m)$, и 0 в противном случае.
5. $TC.tOpen(csk, cpk, c, m') \rightarrow r'$: алгоритм принимает закрытый ключ csk , открытый ключ cpk , текущее обязательство c и новое сообщение $m' \in \mathcal{M}$ ($m' \neq m$). Алгоритм вычисляет и возвращает новое значение случайности $r' \in \mathcal{R}$, такое что проверка для нового сообщения проходит успешно:

$$TC.tVer(cpk, c, m', r') = 1.$$

Для схемы обязательств с потайным входом должны выполняться следующие свойства безопасности:

♦ *Соккрытие (Hiding)*: по значению обязательства c вычислительно невозможно извлечь информацию о сообщении m без знания r .

♦ *Связывание (Binding)*: для стороны, не владеющей секретным ключом csk , вычислительно трудно найти две различные пары (m, r) и (m', r') , соответствующие одному и тому же обязательству c .

Протокол Privacy Pass [18] представляет собой механизм аутентификации, который позволяет пользователю доказать право доступа к серверу без идентификации пользователя сервером. Кроме клиента (Client) и сервера (Origin) в данном протоколе также участвуют аттестатор (Attester) и издатель (Issuer). Аттестатор выполняет роль проверяющего, то есть он отвечает за аутентификацию клиента в системе. Далее, если аутентификация была пройдена успешно, то издатель и клиент совместно генерируют заданное количество токенов для клиента. Эти токены затем могут быть использованы клиентом для получения анонимного доступа к серверу. Следует отметить, что сервер, аттестатор и издатель могут быть одной сущностью.

В основе работы протокола лежит криптографический примитив, называемый забывчивой псевдослучайной функцией (Oblivious PseudoRandom Function, OPRF) [19]. Данный примитив выполняется двумя сторонами, назовем их отправитель и получатель, и позволяет им вычислить значение псевдослучайной функции от их секретных входных данных, таким образом, что получатель получает результат вычисления, не раскрывая своего входа, а отправитель не узнаёт вычисленного значения. Другими словами, пусть получатель имеет некоторое секретное значение X , отправитель секретное значение Y . В результате выполнения протокола получатель совместно с отправителем вычислит значение $z = \text{OPRF}(x, y)$. Использование OPRF в протоколе Privacy Pass позволяет реализовать механизм выдачи анонимных криптографических токенов, обеспечивающий невозможность связывания этапа их получения с последующим использованием при обращении к серверу.

Смешанные сети (Mix-Network, Mix-Net) представляют собой криптографический протокол, выполняемый набором mix -серверов, для обеспечения анонимности группы пользователей. Основная цель протокола заключается в разрыве связи между входящим и исходящим трафиком на каждом узле, делая невозможным для наблюдателя, не контролирующего все узлы сети, отследить, кто с кем общается [20–22].

Как правило, архитектура протокола предполагает последовательность независимых mix -серверов, совместно исполняющих протокол. Далее приводится краткое описание выполняемых шагов каждого из участников протокола.

Пусть n пользователей желают передать сообщения $m_1, \dots, m_n$ соответственно получателю (R) с помощью mix -серверов $\text{Mix}_1, \text{Mix}_2$:

1. Каждый пользователь с помощью открытого ключа R шифрует сообщение $c^{(i)} = \text{Enc}_R(m_i)$, а затем повторно шифрует $c^{(i)}$ с помощью открытых ключей mix -серверов $\text{Mix}_1, \text{Mix}_2$, вкладывая в отправляемые сообщения путь до следующего узла (add):

$$c_{\text{Mix}_2}^{(i)} = \text{Enc}_{\text{Mix}_2}(c^{(i)}, \text{add}_R),$$

$$c_{\text{Mix}_1}^{(i)} = \text{Enc}_{\text{Mix}_1}(c_{\text{Mix}_2}^{(i)}, \text{add}_{\text{Mix}_2}).$$

2. Каждый mix -сервер принимает список закрытых текстов от предыдущего узла, расшифровывает с помощью своего закрытого ключа, применяет секретную перестановку к порядку элементов и передает полученный результат следующему узлу или конечному получателю.

Таким образом, если хотя бы один сервер в цепочке честно выполняет перестановку и не хранит логи соответствия входов и выходов, анонимность участников сохраняется.

Предлагаемое решение. В данной работе предлагается теоретическая модель протокола электронного голосования с возможностью изменения решения участника до момента окончания голосования (E-Voting with Vote Changing, EVVC). Главной особенностью данного протокола является то, что каждый участник имеет возможность произвольное количество раз изменять свое решение в течение периода голосования, но при этом в итоговом подсчёте учитывается только его последнее решение.

Важно подчеркнуть, что представленная схема является алгоритмической моделью взаимодействия сторон и описывает логику работы протокола, абстрагируясь от конкретных программно-аппаратных реализаций.

В схеме принимают участие следующие субъекты:

1. *Участник (User)*: пользователь системы, обладающий правом принятия решения, и желающий его зафиксировать.
2. *Центр регистрации (Registration Center, RC)*: узел системы, проводящий аутентификацию участников и выдающий права на участие (бюллетени/тикеты) валидным пользователям.
3. *Центр подсчета (Counting Center, CC)*: узел системы, проводящий сбор зашифрованных решений, их проверку, подсчет и публикацию итоговых результатов.
4. *Смешанные сети (Mix-Nets)*: промежуточные узлы, осуществляющие анонимизацию участников при передаче сообщений центру подсчета.

Основное преимущество предлагаемой модели перед современными аналогами заключается в том, что участник может изменить свое решение без повторного обращения к организаторам (Центру регистрации). Другим преимуществом является минимизация количества взаимодействий между сторонами протокола.

Ниже представлено формальное описание необходимых алгоритмов и протоколов:

♦ $RC.setup(1^\lambda) \triangleq VRF.keyGen(1^\lambda) \rightarrow (vsk, vpk)$ (Центр регистрации): алгоритм вырабатывает ключевую пару центра регистрации (закрытый vsk и открытый vpk ключи для VRF).

♦ $User.setup(1^\lambda) \triangleq TC.parGen(1^\lambda) \rightarrow pp \& TC.keyGen(pp) \rightarrow (cpk, csk)$ (Участник): Алгоритм вырабатывает параметры схемы обязательств и ключевую пару участника (открытый cpk и закрытый csk ключи).

♦ $User.decide(decision, cpk) \triangleq TC.tCom(decision, cpk) \rightarrow (C_{decision}, r)$ (Участник): данный алгоритм принимает на вход решение участника $decision$ и его открытый ключ cpk , а на выходе возвращает обязательство $C_{decision}$, которое фиксирует выбор участника, и некоторое случайное число r .

♦ $UserRC.PrivacyPass(ID) \rightarrow token$ (Участник $\leftrightarrow$ Центр регистрации): данный интерактивный протокол основан на механизме Privacy Pass и предназначен для анонимного подтверждения права участника на получение бюллетеня без раскрытия его личности центру подсчета и внешним наблюдателям. В рамках протокола участник проходит предварительную проверку, подтверждающую его право на участие в голосовании, после чего получает криптографический токен $token$, удостоверяющий это право. Центр регистрации обеспечивает, что каждому участнику может быть выдано не более одного такого токена, при этом полученный токен не содержит информации, позволяющей связать его с личностью участника при последующем использовании.

♦ $UserRC.TicketReceive(cpk \parallel C_{decision}, token) \rightarrow (ticket, proof)$ (Участник $\leftrightarrow$ Центр регистрации): данный интерактивный протокол отвечает за создание подтвержденного бюллетеня (тикета). Используя ранее полученный анонимный токен $token$, участник передает центру регистрации свой псевдоним (открытый ключ cpk) и замаскированное решение (обязательство $C_{decision}$). Центр регистрации, верифицировав токен, формирует индивидуальный тикет с помощью алгоритма проверяемой случайной функции:

$$VRF.eval(vsk, cpk \parallel C_{decision}) \rightarrow (ticket, proof).$$

После этого центр регистрации возвращает полученный тикет $ticket$ и доказательство $proof$ обратно через анонимный канал:

♦ $UserCC.SendTicket(ticket, proof, C_{decision}, cpk)$ (Участник $\rightarrow$ Центр подсчета): данный протокол при помощи Mix-Nets передает информацию о решении участника (набор данных $m = \{ticket, proof, C_{decision}, cpk\}$ центру подсчета, скрывая личность отправителя.

♦ $CC.TicketVerify(vpk, C_{decision} \parallel cpk, ticket, proof) \triangleq VRF.verify(vpk, C_{decision} \parallel cpk, ticket, proof) \rightarrow 1/0$ (Центр подсчета): данный алгоритм отвечает за проверку входящих данных на подлинность. Для этого на вход центр

подсчета принимает набор $\{ticket, proof, C_{decision}, cpk, vpk\}$ и возвращает 1 (данные подлинны) или 0. Далее в открытый доступ публикуются $ticket, proof, cpk$ и $C_{decision}$, чтобы каждый мог проверить их правильность.

♦ $User.DecisionChange(csk, cpk, C_{decision}, decision') \triangleq$

$TC.tOpen(csk, cpk, C_{decision}, decision') \rightarrow r'$ (Участник): при желании участник имеет возможность изменить свое решение $decision$. Если данный протокол не применялся, то $decision' = decision$ и $r' = r$. Если же изменение необходимо, участник использует свой секретный ключ csk для вычисления новой случайности r' для нового решения $decision'$, соответствующей старому обязательству $C_{decision}$.

♦ $UserCC.DecisionReveal(r', decision', ticket)$ (Участник $\rightarrow$ Центр подсчета): данный протокол при помощи Mix-Network передает информацию для раскрытия решения $m = \{r', decision', ticket\}$ центру подсчета, скрывая личность участника.

♦ $CC.Reveal(r', decision', ticket, cpk, C_{decision}) \rightarrow 1/0$ (Центр подсчета): центр подсчета ищет по уникальному значению $ticket$ соответствующие ему данные $\{C_{decision}, cpk\}$ и проверяет, что решение валидно при помощи алгоритма:

$$TC.tVer(cpk, C_{decision}, decision', r') \rightarrow 1/0.$$

Центр подсчитывает все полученные решения и публикует результаты $\{r', decision', ticket, cpk, C_{decision}\}$, чтобы любой желающий мог их проверить.

♦ $Verify(cpk, C_{decision}, decision', r') = TC.tVer(cpk, C_{decision}, decision', r') \rightarrow 1/0$ (Любой желающий): данный алгоритм позволяет любому внешнему наблюдателю проверить честность и правильность учета решений на основе опубликованных данных.

Для протокола EVVC должны выполняться следующие свойства:

1. Полнота (Completeness): все действительные решения участников подсчитываются правильно.
2. Устойчивость (Soundness): недействительные решения участников не учитываются.
3. Конфиденциальность участника (Voter's Privacy): решение отдельного участника хранится в тайне.
4. Одноразовость (Unreusability): решение каждого участника учитывается единожды.
5. Приемлемость (Eligibility): принимать участие могут только те, кому разрешено голосовать.
6. Справедливость (Fairness): ничто не должно влиять на голосование, т. е. никакие промежуточные результаты никому не могут быть известны.
7. Проверяемость (Verifiability): никто не может фальсифицировать результаты голосования.
8. Атомарная проверяемость (Atomic verifiability): отдельные участники могут убедиться, что их решения были подсчитаны должным образом
9. Универсальная проверяемость (Universal verifiability): любой может проверить результаты подсчета.
10. Изменение решения (Revisability): возможность изменения участником своего решения в течение периода голосования.

Стойкость предлагаемой модели. В данном разделе мы приводим краткое обоснование стойкости предлагаемого решения, в котором намечаем основные пути доказательства свойств. Пусть все используемые криптографические примитивы (VRF, Trapdoor Commitment, Privacy Pass, Mix-Nets и алгоритмы шифрования) являются криптографически стойкими. Также предполагается модель нарушителя, в которой Центр подсчета (CC) и Центр регистрации (RC) не вступают в сговор и не могут обмениваться закрытой информацией друг с другом.

Далее покажем, что при выполнении данных допущений для протокола выполняются все заявленные свойства безопасности:

Полнота (Completeness). Данное свойство выполняется, так как алгоритмы $CC.TicketVerify$ и $CC.Reveal$ публикуют информацию в открытом доступе (на публичной доске). Каждый желающий может проверить наличие своего уникального тикета, а также провести независимую проверку правильности подсчета с помощью алгоритма $Verify(TC.tVer)$.

Устойчивость (Soundness). Выполнение свойства гарантируется корректностью используемой схемы обязательств с потайным входом. Если на вход проверки будут поданы некорректные параметры r' , $decision'$ или cpk , протокол при проверке соответствия обязательству $C_{decision}$ выдаст ошибку. Таким образом, недействительные решения технически не могут быть учтены в итоговом результате.

Конфиденциальность пользователя (User's Privacy). Данное свойство обеспечивается свойством «сокрытия» (hiding) схемы обязательств с потайным входом: сторона, знающая обязательство $C_{decision}$, не может извлечь из него информацию о решении $decision$. Анонимность личности пользователя гарантируется использованием протокола Privacy Pass при взаимодействии с Центром регистрации (разрыв связи ID-токен) и использованием Mix-Nets при взаимодействии с Центром подсчета (разрыв связи IP-сообщение).

Одноразовость (Unreusability). Свойство выполняется за счет того, что на этапе аутентификации Центр регистрации при помощи протокола $User \leftrightarrow RC.PrivacyPass$ фиксирует факт выдачи токена для каждого идентификатора ID . Пользователь технически не может получить валидный токен повторно и, следовательно, не может отправить более одного учитываемого голоса (даже если он меняет решение, учитывается только последняя валидная пара).

Приемлемость (Eligibility). Выполняется, так как каждому пользователю необходимо пройти строгую аутентификацию (проверку права на участие) перед получением токена. Проверка осуществляется Центром регистрации в момент выполнения протокола $PrivacyPass$.

Справедливость (Fairness). Свойство обеспечивается стойкостью используемых примитивов. Никакой злоумышленник (включая организаторов), обладающий современными вычислительными мощностями, не может узнать содержание решений до момента их официального раскрытия (публикации ключей раскрытия или самих значений), что исключает влияние промежуточных итогов на мнение пользователей.

Проверяемость (Verifiability). Свойство выполняется при предположении об отсутствии сговора между RC и CC. Центр регистрации не может сгенерировать больше валидных тикетов, чем реальных пользователей (так как $PrivacyPass$ требует взаимодействия), а Центр подсчета публикует все полученные данные $\{ticket, proof, C_{decision}, cpk\}$ в открытом доступе. Стойкость VRF гарантирует, что тикеты не могут быть подделаны.

Атомарная проверяемость (Atomic verifiability). После завершения процедуры Центр подсчета публикует полный набор данных $\{decision', r', cpk, C_{decision}, ticket\}$. Любой пользователь может найти свой тикет на «доске» и криптографически убедиться, что его решение учтено верно.

Универсальная проверяемость (Universal verifiability). Поскольку все данные $\{decision', r', cpk, C_{decision}, ticket\}$ для всех пользователей публикуются открыто, любой внешний наблюдатель может перепроверить математическую корректность каждого голоса и итоговой суммы.

Возможность изменения решения (Revisability). Пользователь, зная свой секретный ключ csk , может изменить решение $decision$ на $decision'$ даже после отправки обязательства. С помощью алгоритма $TC.tOpen(csk, cpk, C_{decision}, decision') \rightarrow r'$ он вычисляет новую случайность r' , которая делает новое решение валидным для старого обязательства $C_{decision}$. Центр подсчета примет это обновление, проверив его через $TC.tVer$, и учтет последнее пришедшее валидное сообщение.

Таким образом, проведенный анализ показывает, что предлагаемая модель EVVC удовлетворяет всем ключевым требованиям, предъявляемым к защищенным системам электронного голосования. Использование комбинации криптографических примитивов (VRF, Trapdoor Commitment, Privacy Pass и Mix-nets) позволяет достичь баланса между строгой анонимностью участников и прозрачностью процесса подсчета (универсальной проверяемостью). При этом, в отличие от большинства классических подходов, модель успешно реализует свойство изменения решения, обеспечивая пользователю гибкость в принятии решений и дополнительную защиту от внешнего давления.

Заключение. В данной работе была разработана и формально описана теоретическая модель протокола электронного голосования, отличительной особенностью которой является свойство изменения решения участником. Предложенная архитектура позволяет пользователю многократно изменять свое решение в течение отведенного времени без необходимости повторного прохождения процедуры аутентификации или взаимодействия с администраторами системы, что существенно повышает удобство использования.

Проведенный анализ стойкости показал, что модель удовлетворяет всем критически важным требованиям, предъявляемым к системам электронного голосования, включая полную анонимность пользователей, невозможность фальсификации результатов, а также обеспечивая универсальную проверяемость итогов любым внешним наблюдателем.

Предложенная модель характеризуется криптографической гибкостью. Благодаря модульной архитектуре в качестве базовых «строительных блоков» протокола могут быть использованы как классические, так и постквантовые криптографические примитивы без изменения общей логики взаимодействия сторон. Переход к использованию постквантовых криптографических стандартов позволит гарантировать неизменность и тайну принятых решений даже при появлении полномасштабных квантовых компьютеров.

Работа выполнена в рамках государственного задания (проект FSER-2025-0003).

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Lee B., Kim K. Receipt-free electronic voting scheme with a tamper-resistant randomizer // International conference on information security and cryptology. – Berlin, Heidelberg: Springer Berlin Heidelberg, 2002. – P. 389-406.
2. Gritzalis D.A. Principles and requirements for a secure e-voting system // Computers & Security. – 2002. – Vol. 21, No. 6. – P. 539-556.
3. Sampigethaya K., Poovendran R. A survey on mix networks and their secure applications // Proceedings of the IEEE. – 2006. – Vol. 94, No. 12. – P. 2142-2181.
4. Chaum D.L. Untraceable electronic mail, return addresses, and digital pseudonyms // Communications of the ACM. – 1981. – Vol. 24, No. 2. – P. 84-90.
5. Fujioka A., Okamoto T., Ohta K. A Practical Secret Voting Scheme for Large Scale Elections // Advances in Cryptology – AUSCRYPT '92. LNCS. Vol. 718. – Springer, 1993. – P. 244-251.
6. Mohammadi M. K., Yaghoobi M. MIZAN: A secure E-voting schema with vote changeability // 2010 International Conference on Information Society. – IEEE, 2010. – P. 1-5.
7. Maddaloni F. Voting Verification Mechanism for a Distributed Ledger-based Remote Electronic Voting System. – University of Zurich, 2021.
8. Yi L., Wang Q. An E-voting Protocol Based on Blockchain // IACR Cryptology ePrint Archive. – 2022. – Report 2022/1132.
9. Farzaliyev V., Willemson J., Kaasik J. K. Improved Lattice-Based Mix-Nets for Electronic Voting // IET Information Security. – 2022. – Vol. 16, Issue 6. – P. 423-438.
10. Cortier V., Galindo D., Küsters R., Truderung T. SoK: Verifiability Notions for E-Voting Protocols // 2016 IEEE Symposium on Security and Privacy (SP). – IEEE, 2016. – P. 779-798.
11. Бабенко Л.К., Писарев И.А., Макаревич О.Б. Защищенное электронное голосование с использованием слепых посредников // Известия ЮФУ. Технические науки. – 2017. – № 6 (191). – С. 26-36.
12. Heiberg S., Willemson J. Verifiable internet voting in Estonia // 2014 IEEE 2nd International Workshop on Security and Privacy in Big Data (BigSecurity). – IEEE, 2014. – P. 1-8.
13. Федеральный закон от 12.06.2002 № 67-ФЗ (ред. от 23.05.2025) «Об основных гарантиях избирательных прав и права на участие в референдуме граждан Российской Федерации».
14. Беззатеев С.В., Давыдов В.В., Дакуо Ж.Н., Иогансон И.Д., Хуцаева А.Ф. Постквантовая схема обязательств с потайным входом // Прикладная дискретная математика. Приложение. – 2025. – № 18. – С. 226-229.
15. ГОСТ Р 34.12-2015. Информационная технология. Криптографическая защита информации. Блочные шифры. – Введ. 2016-01-01. – М.: Стандартинформ, 2015.
16. Rivest R.L., Shamir A., Adleman L. A method for obtaining digital signatures and public-key cryptosystems // Communications of the ACM. – 1978. – Vol. 21, No. 2. – P. 120-126.
17. Дакуо Ж.Н. WaveVRF: постквантовая проверяемая псевдослучайная функция, основанная на кодах, исправляющих ошибки // Научно-технический вестник информационных технологий, механики и оптики. – 2025. – Т. 25, № 1. – С. 174-177.

18. Davidson A., et al. Privacy Pass: Bypassing Internet Challenges Anonymously // *Proceedings on Privacy Enhancing Technologies*. – 2018.
19. Jarecki S., Kiayias A., Krawczyk H. Round-optimal password-protected secret sharing and TPAKE in the password-only model. In: P. Sarkar, T. Iwata (eds.) // *ASIACRYPT 2014, Part II, LNCS*. Vol. 8874. – P. 233-253. – Springer, Heidelberg, December 2014.
20. Sampigethaya K., Poovendran R. A survey on mix networks and their secure applications // *Proceedings of the IEEE*. – 2007. – Vol. 94, No. 12. – P. 2142-2181.
21. Furukawa J., Mori K., Sako K. An implementation of a mix-net based network voting scheme and its use in a private organization // *Towards Trustworthy Elections: New Directions in Electronic Voting*. – Springer, 2010. – P. 141-154.
22. Oluranti J., Olajide F., Ayo C. Network forensics tools in a mixed-network environment and the adoption of e-voting system in developing countries // *International Journal of Pharmacy and Technology*. – 2016. – Vol. 8, No. 4. – P. 23115-23118.

REFERENCES

1. Lee B., Kim K. Receipt-free electronic voting scheme with a tamper-resistant randomizer, *International conference on information security and cryptology*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2002, pp. 389-406.
2. Gritzalis D.A. Principles and requirements for a secure e-voting system, *Computers & Security*, 2002, Vol. 21, No. 6, pp. 539-556.
3. Sampigethaya K., Poovendran R. A survey on mix networks and their secure applications, *Proceedings of the IEEE*, 2006, Vol. 94, No. 12, pp. 2142-2181.
4. Chaum D.L. Untraceable electronic mail, return addresses, and digital pseudonyms, *Communications of the ACM*, 1981, Vol. 24, No. 2, pp. 84-90.
5. Fujioka A., Okamoto T., Ohta K. A Practical Secret Voting Scheme for Large Scale Elections, *Advances in Cryptology – AUSCRYPT '92. LNCS*. Vol. 718. Springer, 1993, pp. 244-251.
6. Mohammadi M. K., Yaghoobi M. MIZAN: A secure E-voting schema with vote changeability, *2010 International Conference on Information Society*. IEEE, 2010, pp. 1-5.
7. Maddaloni F. Voting Verification Mechanism for a Distributed Ledger-based Remote Electronic Voting System. University of Zurich, 2021.
8. Yi L., Wang Q. An E-voting Protocol Based on Blockchain, *IACR Cryptology ePrint Archive*, 2022, Report 2022/1132.
9. Farzaliyev V., Willemson J., Kaasik J. K. Improved Lattice-Based Mix-Nets for Electronic Voting, *IET Information Security*, 2022, Vol. 16, Issue 6, pp. 423-438.
10. Cortier V., Galindo D., Küsters R., Truderung T. SoK: Verifiability Notions for E-Voting Protocols, *2016 IEEE Symposium on Security and Privacy (SP)*. IEEE, 2016, pp. 779-798.
11. Babenko L.K., Pisarev I.A., Makarevich O.B. Zashchishchennoe elektronnoe golosovanie s ispol'zovaniem slepykh posrednikov [Secure electronic voting using blind intermediaries], *Izvestiya YuFU. Tekhnicheskie nauki* [Izvestiya SFedU. Engineering Sciences], 2017, No. 6 (191), pp. 26-36.
12. Heiberg S., Willemson J. Verifiable internet voting in Estonia, *2014 IEEE 2nd International Workshop on Security and Privacy in Big Data (BigSecurity)*. IEEE, 2014, pp. 1-8.
13. Federal'nyy zakon ot 12.06.2002 № 67-FZ (red. ot 23.05.2025) «Ob osnovnykh garantiyakh izbiratel'nykh prav i prava na uchastie v referendumе grazhdan Rossiyskoy Federatsii» [13. Federal Law of 12.06.2002 No. 67-FZ (as amended on 23.05.2025) “On the Basic Guarantees of Electoral Rights and the Right to Participate in Referendums of Citizens of the Russian Federation”].
14. Bezzateev S.V., Davydov V.V., Dakuo Zh.N., Ioganson I.D., Khutsaeva A.F. Postkvantovaya skhema obyazatel'stv s potaynym vkhodom [Post-quantum commitment scheme with a trapdoor], *Prikladnaya diskretnaya matematika. Prilozhenie* [Applied Discrete Mathematics. Supplement], 2025, No. 18, pp. 226-229.
15. GOST R 34.12-2015. Informatsionnaya tekhnologiya. Kriptograficheskaya zashchita informatsii. Blochnye shifry [GOST R 34.12-2015. Information technology. Cryptographic protection of information. Block ciphers]. Introduced 2016-01-01. Moscow: Standartinform, 2015.
16. Rivest R.L., Shamir A., Adleman L. A method for obtaining digital signatures and public-key cryptosystems, *Communications of the ACM*, 1978, Vol. 21, No. 2, pp. 120-126.
17. Dakuo Zh.N. WaveVRF: postkvantovaya proverayemaya psevdosluchaynaya funktsiya, osnovannaya na kodakh, ispravlyayushchikh oshibki [WaveVRF: Post-quantum verifiable pseudorandom function based on error-correcting codes], *Nauchno-tekhnicheskiy vestnik informatsionnykh tekhnologiy, mekhaniki i optiki* [Scientific and Technical Bulletin of Information Technologies, Mechanics and Optics], 2025, Vol. 25, No. 1, pp. 174-177.
18. Davidson A., et al. Privacy Pass: Bypassing Internet Challenges Anonymously, *Proceedings on Privacy Enhancing Technologies*, 2018.

19. Jarecki S., Kiayias A., Krawczyk H. Round-optimal password-protected secret sharing and TPAKE in the password-only model, In: P. Sarkar, T. Iwata (eds.), *ASIACRYPT 2014, Part II, LNCS. Vol. 8874*, pp. 233-253. Springer, Heidelberg, December 2014.
20. Sampigethaya K., Poovendran R. A survey on mix networks and their secure applications, *Proceedings of the IEEE*, 2007, Vol. 94, No. 12, pp. 2142-2181.
21. Furukawa J., Mori K., Sako K. An implementation of a mix-net based network voting scheme and its use in a private organization, *Towards Trustworthy Elections: New Directions in Electronic Voting*. Springer, 2010, pp. 141-154.
22. Oluranti J., Olajide F., Ayo C. Network forensics tools in a mixed-network environment and the adoption of e-voting system in developing countries, *International Journal of Pharmacy and Technology*, 2016, Vol. 8, No. 4, pp. 23115-23118.

Дакуо Жан-Мишель Никодэмович – Национальный исследовательский университет ИТМО; e-mail: jeandakuo@mail.ru; г. Санкт-Петербург, Россия; тел.: +79217863122; факультет безопасности информационных технологий; аспирант.

Беззатеев Сергей Валентинович – Государственный университет аэрокосмического приборостроения; e-mail: bsv@guar.ru; г. Санкт-Петербург, Россия; тел.: +79045170951; кафедра информационной безопасности; зав. кафедрой.

Давыдов Вадим Валерьевич – ООО «КуАпп»; e-mail: vadimdavydov@outlook.com; г. Москва, Россия; криптограф-исследователь.

Иогансон Иван Дмитриевич – Национальный исследовательский университет ИТМО; e-mail: ivan.ioganson@yandex.ru; г. Санкт-Петербург, Россия; тел.: +79052278519; факультет безопасности информационных технологий; м.н.с.

Хуцаева Алтана Феликсовна – Государственный университет аэрокосмического приборостроения; e-mail: afkhutsaeva@itmo.ru; г. Санкт-Петербург, Россия; кафедра информационной безопасности; ассистент.

Dakuo Zhan-Mishel Nikodemovich – ITMO University; e-mail: jeandakuo@mail.ru; St. Petersburg, Russia; phone: +79217863122; Faculty of Secure Information Technologies; graduate student.

Bezzateev Sergey Valentinovich – State University of Aerospace Instrumentation; e-mail: bsv@guar.ru; St. Petersburg, Russia; phone: +79045170951; the Department of Information Security; head of the Department.

Davydov Vadim Valerievich – QApp LLC, Cryptographer-Researcher; e-mail: vadimdavydov@outlook.com; Moscow, Russia; cryptographer researcher.

Ioganson Ivan Dmitrievich – ITMO University; e-mail: ivan.ioganson@yandex.ru; St. Petersburg, Russia; phone: +79052278519; Faculty of Secure Information Technologies; junior scientific researcher.

Khutsaeva Altana Felixovna – State University of Aerospace Instrumentation; e-mail: afkhutsaeva@itmo.ru; St. Petersburg, Russia; the Department of Information Security; assistant.

УДК 681.5

DOI 10.18522/2311-3103-2026-4-212-226

А.Н. Каркищенко

КОЛИЧЕСТВЕННЫЕ ОЦЕНКИ ПОГРЕШНОСТИ ПОДДЕРЖАНИЯ СТРОЯ ПОДВОДНЫХ АППАРАТОВ ПРИ ЗАДЕРЖКАХ НАВИГАЦИОННЫХ ДАННЫХ

Изучается влияние задержек в получении навигационной информации на поддержание строя подводных аппаратов в режиме «лидер-ведомые». Исследование опирается на предположение о градиентном принципе управления строем в процессе движения. Сложность задачи обусловлена неизвестностью траектории лидера, неполным знанием и нестационарностью среды, ошибками измерений. Полученные результаты не зависят от конфигурации аппаратов в строю. Показано, что достаточно исследовать влияние навигационных задержек на движение одного аппарата. Основные предположения: траектория движения лидера неизвестна; все взаимодействия между лидером и ведомыми аппаратами синхронизированы; лидер с определенной частотой передает свои координаты другим аппаратам; смещения ведомого аппарата от его места в строю могут

происходить в любом пространственном направлении. Предполагается, что между получением данных о положении лидера ведомый аппарат движется по прямой линии в направлении, выбор которого можно сделать многими способами. Рассматривается один из вариантов выбора такого направления и строится математическая модель движения. Ввиду того, что момент поступления информации неизвестен, применяется нелинейный закон для описания движения ведомого аппарата, учитывающий расстояние до целевой точки. Найдены оценки уклонения аппарата от целевой позиции в строю в каждый момент времени, а также усредненные оценки его уклонения на отрезке времени между получением навигационной информации и на всей траектории. Предложена и обоснована верхняя оценка интегрального уклонения, позволяющая дать простую интерпретацию факторов, влияющих на уклонение и нарушение строя.

Подводный аппарат; лидер-ведомые; поддержание строя; уклонение от позиции; верхние оценки.

A.N. Karkishchenko

QUANTITATIVE ESTIMATES OF THE ERROR IN MAINTAINING THE FORMATION OF UNDERWATER VEHICLES WITH DELAYS IN NAVIGATION DATA

The influence of delays in receiving navigation information on maintaining a formation of underwater vehicles in the leader-followers mode is studied. The study is based on the assumption of a gradient principle of formation control during movement. The complexity of the problem is due to the unknown trajectory of the leader, incomplete knowledge and non-stationarity of the environment, and measurement errors. The obtained results are independent of the configuration of vehicles in the formation. It is shown that it is sufficient to study the influence of navigation delays on the motion of a single vehicle. Key assumptions: the trajectory of the leader is unknown; all interactions between the leader and follower vehicles are synchronized; the leader transmits its coordinates to the other vehicles at a certain frequency; displacements of the follower from its position in the formation can occur in any spatial direction. It is assumed that between receiving data on the leader's position, the follower vehicle moves in a straight line in a direction that can be chosen in many ways. One option for choosing such a direction is considered, and a mathematical model of the motion is constructed. Since the moment of information receipt is unknown, a nonlinear law is used to describe the follower's motion, taking into account the distance to the target point. Estimates of the vehicle's deviation from the target position in the formation at each moment in time are found, as well as average estimates of its deviation over the time interval between receiving navigation information and over the entire trajectory. An upper bound for the integral deviation is proposed and substantiated, allowing for a simple interpretation of the factors influencing deviation and formation breakdown.

Underwater vehicle; leader-followers; formation maintenance; position evasion; upper estimates.

Введение. Развитие подводной робототехники делает особенно актуальным групповое применение автономных необитаемых подводных аппаратов (АНПА) для наблюдения и мониторинга морских акваторий, исследования рельефа дна, инспекции искусственных и природных объектов, а также решения поисковых и специальных прикладных задач [1]. Такие комплексы позволяют повысить пространственный охват наблюдаемой области, обеспечить функциональное распределение ролей между аппаратами и повысить общую устойчивость выполнения миссии по сравнению с использованием одиночного подводного аппарата [2, 3]. Вместе с тем расширение практики группового применения АНПА сопровождается возрастанием требований к точности навигации, устойчивости управления и согласованности движения аппаратов в общей системе координат [2, 4].

Одной из ключевых задач группового управления является поддержание заданного строя при движении вдоль требуемой траектории [4, 5]. Под строем АНПА понимается пространственно организованная группа подводных аппаратов, совместно функционирующих для достижения общей цели [3]. При этом управление группой предполагает не только удержание предписанной геометрической конфигурации, но и планирование траектории всей группы и отдельных аппаратов, согласование скоростей и направлений движения, а также адаптацию к внешним ограничениям и возмущениям среды [3, 4, 6].

Специфика подводной среды существенно усложняет решение задачи формирования и поддержания строя. В отличие от наземных и воздушных робототехнических систем, для АНПА характерны отсутствие прямого доступа к сигналам навигационных спутниковых систем, ограниченные возможности радиосвязи, зависимость качества навигации от гидроакустических измерений, а также влияние течений, рельефа дна, неоднородности среды и подвижных препятствий [3, 7]. Именно поэтому задачи управления строем АНПА тесно связаны с проблемами неполноты информации, запаздывания измерений и ограниченной пропускной способности каналов связи [3].

В настоящее время в задачах управления строем подводных аппаратов выделяются несколько базовых подходов: архитектура «лидер–ведомые» [8], ситуационный подход [9], методы виртуальной структуры [10], а также распределённые и схемы кооперативного управления [3, 4]. Подход «лидер–ведомые» удобен из-за простоты организации движения группы и задания строя относительно аппарата-лидера, однако он чувствителен к ошибкам навигации лидера и к нарушениям каналов передачи информации [3, 11, 12]. Подход виртуальной структуры позволяет описывать группу как единый пространственный объект с фиксированной геометрией [13], однако распределённые схемы обеспечивают большую гибкость и живучесть группы при отказах отдельных аппаратов [4, 14].

Значительное число работ посвящено использованию методов искусственных потенциалов, градиентных алгоритмов и их модификаций для локальной навигации и формирования строя мобильных объектов, включая подводные аппараты [15–18]. Данные методы позволяют единообразно учитывать притяжение к целевой конфигурации, взаимное пространственное согласование аппаратов и уклонение от препятствий [6, 13, 19].

Обзоры последних лет показывают, что исследования в области формирования строя АНПА развиваются по нескольким взаимосвязанным направлениям: обеспечение устойчивости геометрии строя, координация движения при ограничениях связи, согласование траекторий в трёхмерном пространстве и повышение робастности по отношению к неопределённостям модели и внешним возмущениям [3, 4]. При этом всё большее внимание уделяется не только формированию законов управления, но и оценке влияния сетевых и навигационных ограничений, включая задержки передачи данных и дискретность обмена сообщениями на качество поддержания строя [11, 20].

Особое значение имеют работы, в которых рассматривается управление строем нескольких АНПА при ограничениях на коммуникацию и запаздывании информации. В частности, известно, что задержки и пропуски данных приводят к ухудшению точности взаимного позиционирования и могут вызывать заметное искажение желаемой конфигурации группы [11]. Для подводных систем эта проблема особенно существенна, поскольку задержка возникает не только в линиях связи, но и на уровне получения и обработки навигационной информации, что непосредственно влияет на ошибки поддержания строя вдоль всей траектории движения [3, 7].

Следует отметить то обстоятельство, что во многих работах основное внимание уделяется формированию управления группой АНПА, тогда как количественная оценка погрешности поддержания строя на участках траектории изучена существенно слабее [3, 4]. Между тем именно такие оценки представляют практический интерес при проектировании групп АНПА, поскольку позволяют установить допустимые требования к навигационным средствам, частоте обновления данных и скорости движения [3, 7]. Для неоднородных по динамике групп данная проблема становится ещё более сложной и значимой [20].

В ряде современных публикаций рассматриваются адаптивные, робастные и интеллектуальные методы формирования строя, включая сочетание потенциальных полей, виртуальных структур и различных методов обучения [11]. Эти исследования демонстрируют расширение возможностей группового управления в сложной обстановке, однако не снимают вопроса об оценках точности поддержания строя при наличии навигационных задержек. Поэтому задача построения количественных оценок погрешности поддержания строя, зависящих от величины задержек и скорости движения аппаратов, сохраняет самостоятельную теоретическую и прикладную значимость [3].

В настоящей работе в рамках подхода «лидер–ведомый» изучается влияние задержек навигационных данных на поддержание строя АНПА. Исследование опирается на градиентный принцип формирования строя в процессе группового управления движением подводных аппаратов и продолжает развитие подхода, основанного на вероятностной природе потенциального поля [21]. В отличие от работ, в которых основное внимание сосредоточено на конструктивном синтезе управления [11, 13, 14, 19, 22], здесь акцент делается на анализе точности поддержания строя и на получении количественных оценок погрешности, применимых к различным пространственным конфигурациям расположения аппаратов.

Представляемые результаты ориентированы на выявление зависимости точности поддержания строя от величины задержек навигационной информации и параметров движения группы. Такой анализ имеет важное значение для обоснования требований к навигационному и информационному обеспечению, а также режимам функционирования группы АНПА при выполнении практических миссий в подводной среде [3, 7].

Основные предположения и постановка задачи. Будем рассматривать подход к управлению группой подводных аппаратов, основанный на принципе «лидер–ведомые». Под *строем* будем понимать виртуальную пространственную структуру, образованную точками – позициями подводных аппаратов, которая полностью определяется положением и ориентацией АНПА-лидера, а также относительными позициями ведомых аппаратов. При этом считается, что лидер группы движется по заданной неизвестной траектории с почти постоянной скоростью v_L , которая лишь на некоторых участках может незначительно меняться. Относительно него формируется реальный строй. Ведомые АНПА имеют назначенные им заранее позиции относительно лидера в виртуальном строе, которые они должны по возможности максимально точно занимать в процессе движения. Будем считать, что динамические и кинематические характеристики лидера и остальных АНПА одинаковы. При этом известны максимальная $v_{\max}$ и минимальная $v_{\min}$ скорости подводных аппаратов, которые могут ими достигаться при движении в неподвижной воде. Ставится задача исследовать влияние частоты обновления навигационных данных на точность поддержания строя.

Основные предположения состоят в следующем:

- ♦ лидер и ведомые аппараты оснащены средствами обнаружения препятствий, что позволяет им двигаться независимо;
- ♦ все взаимодействия между лидером и ведомыми АНПА синхронизированы во времени, т.е. все подводные аппараты получают информацию о текущем положении лидера одним пакетом и обновляют свои параметры пространственного состояния почти одновременно;
- ♦ лидер с определенной частотой передает собственные координаты другим АНПА. Дальность передачи считаем достаточной для ее получения аппаратами, входящими в состав группы;
- ♦ вызванное неблагоприятными факторами смещение ведомого АНПА от его места в строе может происходить в любом пространственном направлении и на любое расстояние в пределах радиуса информационной связи с лидером.

Далее оценку уклонения ведомого АНПА от его позиции в строе будем проводить для одного аппарата, так как уклонение для всей группы легко агрегируется с помощью любой усредненной метрики в пространстве, размерность которого определяется числом ведомых подводных аппаратов в группе.

Место в строе, которое должно быть занято аппаратом в строе, будем называть *позицией*. Считаем, что АНПА в точности занимает свою позицию в строе, если координаты его центра тяжести совпадают с координатами назначенного места в строе, а ориентация его продольной оси сонаправлена с осью лидера.

Введем обозначения для величин, характеризующих пространственное положение произвольного АНПА в некоторой неподвижной системе координат $OXYZ$ в момент времени t :

$P^L(t) = (x^L(t), y^L(t), z^L(t))$ – текущие абсолютные координаты позиции в строе;

$P(t) = (x(t), y(t), z(t))$ – текущие абсолютные координаты АНПА;

$\|P^L(t) - P(t)\| = \sqrt{(x^L(t) - x(t))^2 + (y^L(t) - y(t))^2 + (z^L(t) - z(t))^2}$ – евклидово расстояние (уклонение) между позицией в строе и положением АНПА в момент t . Интегральную оценку уклонения на всей траектории за период $[0, T]$ будем характеризовать с помощью усредненного евклидова расстояния (уклонения)

$$d = \frac{1}{T} \int_0^T \|P^L(t) - P(t)\| dt.$$

Траектория $P^L(t)$ позиции в строе полностью определяется траекторией движения лидера и по условию ведомым АНПА априори неизвестна. Траектории ведомых аппаратов также зависят от траектории лидера и определяются только в процессе движения. Отмеченные выше особенности гидроакустической связи приводят к задержкам и нерегулярности поступления информации о координатах положения лидера. Из этого следует, что аналитические выражения для $P^L(t)$ и $P(t)$, описывающие траектории позиции в строе и остальных аппаратов, отсутствуют. Доступными оказываются только дискретно и в общем случае нерегулярно поступающие сообщения о координатах лидера. Подобные ограничения на поступление навигационных данных ограничивают возможности эффективного поддержания строя и обеспечение безопасности группы в целом.

Будем предполагать, что данные о нахождении лидера поступают в моменты времени $t_0, t_1, \dots, t_n$, в которые ведомые АНПА получают информацию о меняющихся координатах $P^L(t_0), P^L(t_1), \dots, P^L(t_n)$ позиции в строе. Ведомые аппараты обладают также информацией о координатах своего собственного положения. Ввиду этого для усредненного евклидова расстояния можно записать:

$$d = \frac{1}{T} \sum_{k=0}^{n-1} \int_{t_k}^{t_{k+1}} \|P^L(t) - P(t)\| dt + \frac{1}{T} \int_{t_n}^T \|P^L(t) - P(t)\| dt.$$

Следует отметить еще раз, что на каждом из интервалов функции $P^L(t)$ и $P(t)$ определяются в общем случае разными выражениями. Далее введем простое допущение, что на каждом участке времени $[t_k, t_{k+1})$ АНПА движется по прямой линии от точки $P(t_k)$, в которой он находится в момент t_k , в направлении, выбор которого можно сделать многими способами. Ниже рассмотрим одну из таких возможностей. При этом за координаты точки, в которой ведомый АНПА окажется в момент времени t_{k+1} , следует принять $P(t_{k+1})$. Эта точка будет отправной для следующего прямолинейного участка движения. Понятно, что положение точки $P(t_{k+1})$ зависит от направления и скорости движения подводного аппарата вдоль выбранной прямолинейной траектории.

Задача состоит в получении интегральной характеристики уклонения аппарата от его позиции на $[0, T]$. Выражение должно зависеть от параметра Δt , определяющего частоту обновления данных. Для этого необходимо получить выражения, задающие траектории движения АНПА на каждом участке времени.

Закон движения ведомого подводного аппарата. Предположим, что как только в момент времени t_k становится известно положение лидера $P^L(t_k)$ в строе, ведомый АНПА начинает прямолинейное движение в эту точку из точки $P(t_k)$ с постоянной на промежутке $[t_k, t_{k+1})$ скоростью v_k . Это движение описывается уравнением: $P(t) = P(t_k) + a_k^0 v_k (t - t_k)$, $t \in [t_k, t_{k+1})$. В этом выражении a_k^0 – единичный направляющий

вектор прямолинейной траектории (рис. 1), в качестве которого можно взять вектор $a_k^0 = \frac{P^L(t_k) - P(t_k)}{\|P^L(t_k) - P(t_k)\|}$. Считаем, что в начальный момент времени значения $P^L(t_0)$ и $P(t_0)$ известны. Отметим еще раз, что на каждом временном отрезке скорость ведомых подводных аппаратов постоянная, но на каждом отрезке она может быть своя.

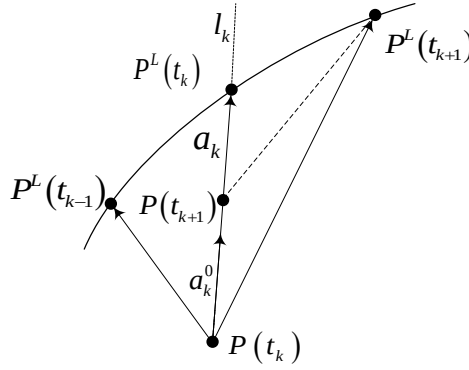


Рис. 1. Траектория движения АНПА

Координаты позиции полностью определяются координатами лидера. Поэтому будем иногда для простоты обозначать через $P^L(t)$ и координаты лидера. В момент t_{k+1} АНПА будет находиться в точке $P(t_{k+1}) = P(t_k) + a_k^0 v_k (t_{k+1} - t_k)$. В это время позиция в строе будет иметь координаты $P^L(t_{k+1})$, а расстояние между позицией в строе и АНПА будет равно:

$$\|P^L(t_{k+1}) - P(t_{k+1})\| = \|P^L(t_{k+1}) - P(t_k) - a_k^0 v_k (t_{k+1} - t_k)\|.$$

Условие, при котором АНПА после движения на отрезке сблизится со своей позицией в строе, выражается неравенством $\|P^L(t_k) - P(t_k)\| > \|P^L(t_{k+1}) - P(t_{k+1})\|$. Надо, чтобы с ростом k это расстояние как можно быстрее убывало, и соответственно, ведомый АНПА как можно быстрее «догонял» свою позицию.

Пусть момент t_{k+1} известен и $t_{k+1} - t_k = \Delta t_k$. Скорость v_k^* равномерного движения, при которой аппарат к моменту t_{k+1} попадет в точку $P^L(t_k)$ должна быть равна $v_k^* = \frac{P^L(t_k) - P(t_k)}{\Delta t_k}$. Однако, учитывая ограничения $v_{\min} \leq v_k \leq v_{\max}$, скорость v_k на отрезке времени $[t_k, t_{k+1})$ и положение АНПА в момент t_{k+1} должны определяться следующим образом:

$$v_k = \begin{cases} v_{\max}, & \text{если } v_k^* \geq v_{\max}, \\ v_k^*, & \text{если } v_{\min} < v_k^* < v_{\max}, \\ v_{\min}, & \text{если } v_{\min} \geq v_k^*. \end{cases} \quad P(t_{k+1}) = \begin{cases} P(t_k) + a_k^0 v_{\max} \Delta t_k, & \text{если } v_k^* \geq v_{\max}, \\ P^L(t_k), & \text{если } v_{\min} < v_k^* < v_{\max}, \\ P(t_k) + a_k^0 v_{\min} \Delta t_k, & \text{если } v_{\min} \geq v_k^*. \end{cases}$$

В первом случае аппарат к моменту t_{k+1} , очевидно, не достигнет точки $P^L(t_k)$, в то время как в третьем случае «проскочит» ее.

Однако точный момент времени t_{k+1} , как правило, неизвестен. В этом случае следует использовать не постоянную скорость, а ввести некоторый нелинейный закон $v(t)$ изменения скорости ведомого АНПА, при котором он вначале максимально быстро дви-

жется по направлению к точке $P^L(t_k)$, а затем по мере приближения к ней так же быстро, но плавно уменьшает скорость вплоть до $v_{\min}$. В этом случае закон движения АНПА будет описываться выражением

$$P(t) = P(t_k) + a_k^0 \int_{t_k}^t v(\tau) d\tau = P(t_k) + a_k^0 s(t), \quad t \in [t_k, t_{k+1}), \quad (1)$$

где a_k^0 имеет тот же смысл, что и выше, а $s(t)$ – путь, пройденный ведомым аппаратом на k -м участке движения к моменту t . В качестве такого закона целесообразно взять выражение, которое было получено и использовалось в [13]:

$$v(t) = v(\rho(t)) = v_{\max} - (v_{\max} - v_{\min}) e^{-\gamma \rho(t)}, \quad t \in [t_k, t_{k+1}), \quad (2)$$

где γ – параметр, влияющий на скорость снижения скорости при приближении АНПА к своей позиции $P^L(t_k)$, $\rho(t)$ – расстояние от АНПА до целевой позиции в строе $P^L(t_k)$ в произвольный момент времени t . При достаточно большом расстоянии до целевой точки и большом значении параметра γ аппарат будет стартовать от начальной точки $P(t_k)$ со скоростью близкой к $v_{\max}$, экспоненциально уменьшающейся вплоть до $v_{\min}$. Еще раз укажем, что значение t_{k+1} неизвестно.

Найдем явное выражение для $s(t)$ на участке $[t_k, t)$. Подводный аппарат движется прямолинейно из точки $P(t_k)$ в точку $P^L(t_k)$. Как следует из (2), по мере приближения АНПА к целевой позиции его скорость уменьшается до $v_{\min}$. Введем обозначение $W_1(t_k) = P^L(t_k) - P(t_k)$. Тогда $\rho(t) = \|W_1(t_k)\| - s(t)$ (рис. 2).

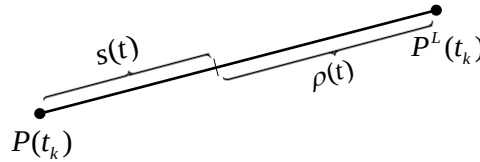


Рис. 2. Расстояния $\rho(t)$ и $s(t)$

Поскольку $v(t) = \frac{ds}{dt}$, то можем записать $\frac{ds}{dt} = v_{\max} - (v_{\max} - v_{\min}) e^{-\gamma(\|W_1(t_k)\| - s(t))}$. Решая это уравнение с начальным условием $s(t_k) = 0$, получаем:

$$\ln |v_{\max} e^{-\gamma s(t)} - (v_{\max} - v_{\min}) e^{-\gamma \|W_1(t_k)\|}| = -\gamma v_{\max} t + \ln C, \quad (3)$$

где произвольная постоянная взята для удобства в виде $\ln C$. Для отыскания постоянной C воспользуемся начальным условием $s(t_k) = 0$. В результате получим $\ln |v_{\max} - (v_{\max} - v_{\min}) e^{-\gamma \|W_1(t_k)\|}| e^{\gamma v_{\max} t_k} = \ln C$. Подставим найденное значение $\ln C$ в (3), после преобразований получим:

$$\frac{|v_{\max} e^{-\gamma s(t)} - (v_{\max} - v_{\min}) e^{-\gamma \|W_1(t_k)\|}|}{|v_{\max} - (v_{\max} - v_{\min}) e^{-\gamma \|W_1(t_k)\|}|} = e^{-\gamma v_{\max} (t - t_k)}.$$

Можно несколько упростить запись этого выражения, если ввести обозначение $\alpha = 1 - \frac{v_{\min}}{v_{\max}}$. Тогда $\frac{e^{-\gamma s(t)} - \alpha e^{-\gamma \|W_1(t_k)\|}}{1 - \alpha e^{-\gamma \|W_1(t_k)\|}} = e^{-\gamma v_{\max}(t-t_k)}$. Отсюда находим искомую зависимость на $[t_k, t_{k+1})$: $s(t) = -\frac{1}{\gamma} \ln \left[\alpha e^{-\gamma \|W_1(t_k)\|} + (1 - \alpha e^{-\gamma \|W_1(t_k)\|}) e^{-\gamma v_{\max}(t-t_k)} \right]$ или

$$s(t) = \|W_1(t_k)\| - \frac{1}{\gamma} \ln \left[\alpha + (e^{\gamma \|W_1(t_k)\|} - \alpha) e^{-\gamma v_{\max}(t-t_k)} \right]. \quad (4)$$

Нетрудно видеть, что функция $s(t)$ всегда неотрицательна. Таким образом, с учетом обозначения $W_1(t_k) = P^L(t_k) - P(t_k)$, получаем, что выражение (1) для закона движения принимает вид: $P(t) = P(t_k) + \frac{W_1(t_k)}{\|W_1(t_k)\|} s(t)$. Функция $s(t)$ на отрезке времени $[0, T]$ является псевдопериодической в том смысле, что $s(t_k) = 0$ для всех $k = 0, 1, \dots, n-1$, и на каждом участке $[t_k, t_{k+1})$ она повторяет одну и ту же функциональную зависимость, отличаясь лишь длиной $t_{k+1} - t_k$ этого участка времени и, возможно, «высотой», которая определяется параметром $W_1(t_k)$. Схематично вид этой функции показан на рис. 3. На этом рисунке функция $s(t)$ задана на отрезке $[0, 16]$, который распадается на участки $[0, 3)$, $[3, 5)$, $[5, 11)$, $[11, 16]$. Остальные параметры имеют следующие значения: $\gamma = 1.4$, $v_{\min} = 0.5$ м/с, $v_{\max} = 5$ м/с (и соответственно, $\alpha = 0.9$), $W_1(3) = W_1(5) = W_1(11) = W_1(16) = 4$.

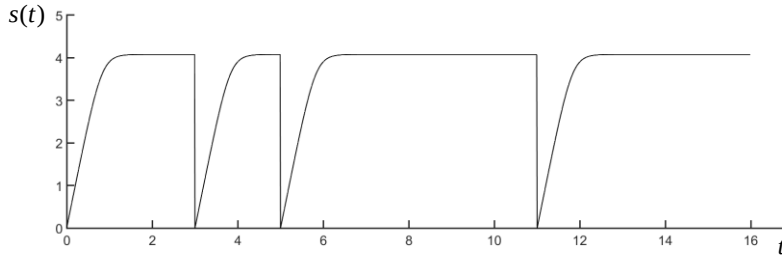


Рис. 3. Вид функции $s(t)$ на отрезке $[0, T]$

Линейная экстраполяция траектории лидера. Следующая задача состоит в том, чтобы найти аналитическое выражение для разности между положением ведомого АНПА и его позицией в строе. Эта позиция непрерывно движется в соответствии с законом движения лидера, который в свою очередь неизвестен. Поэтому необходимо получить приближенное выражение для траектории этой позиции. Это можно сделать на основании экстраполяции траектории лидера на основе информации о его движении в предыдущий период. Траектория лидера неизвестна, но в силу предположений его скорость v_L все время постоянна или меняется незначительно так, что ее можно считать постоянной. Так как по условию все АНПА имеют одинаковые характеристики, то $v_{\min} \leq v_L \leq v_{\max}$.

Для правдоподобной оценки уклонения аппарата от его позиции будем считать, что лидер на каждом отрезке времени между получением информации о его местоположении движется прямолинейно. Чем меньше задержки в передаче навигационной информации, тем меньше на этом участке будет расхождение гипотетической прямолинейной траектории лидера с реальной. Это утверждение становится еще более обоснованным в виду инерционности движения в водной среде. Время t_{k+1} и местоположение $P^L(t_{k+1})$ лидера неизвестны, поэтому сделаем дополнительные предположения и допущения.

Будем считать, что на промежутке $[t_k, t_{k+1})$ лидер движется прямолинейно в том же направлении, что и его движение на предыдущем (уже известном) участке, т.е. будем использовать линейную экстраполяцию траектории. Иными словами, считаем, что лидер движется по прямой линии из точки $P^L(t_k)$ в направлении, задаваемом единичным вектором $b_{k-1}^0 = \frac{P^L(t_k) - P^L(t_{k-1})}{\|P^L(t_k) - P^L(t_{k-1})\|}$, причем с постоянной скоростью $v = \frac{\|P^L(t_k) - P^L(t_{k-1})\|}{t_k - t_{k-1}}$, которая равна средней скорости лидера на предыдущем участке движения (рис. 4).

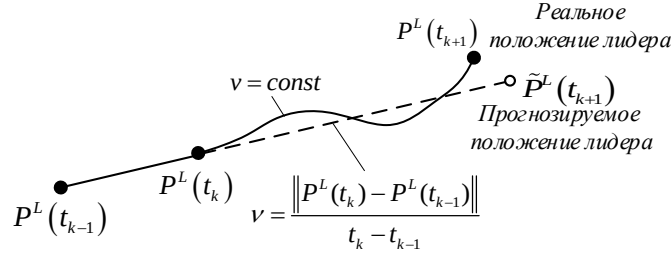


Рис. 4. Линейная аппроксимация движения лидера

Экстраполированное движение лидера (виртуальной позиции) на участке времени $[t_k, t_{k+1})$ описывается следующим уравнением:

$$P^L(t) = P^L(t_k) + b_{k-1}^0 v \cdot (t - t_k) = P^L(t_k) + (P^L(t_k) - P^L(t_{k-1})) \frac{t - t_k}{t_k - t_{k-1}}.$$

Поэтому прогнозируемым положением $\tilde{P}^L(t_{k+1})$ лидера в момент t_{k+1} будет $\tilde{P}^L(t_{k+1}) = P^L(t_k) + (P^L(t_k) - P^L(t_{k-1})) \frac{t_{k+1} - t_k}{t_k - t_{k-1}}$, и, если, например, $t_{k+1} - t_k \approx t_k - t_{k-1}$, то $\tilde{P}^L(t_{k+1}) \approx 2P^L(t_k) - P^L(t_{k-1})$.

Реальное положение $P^L(t_{k+1})$ позиции (лидера) в момент t_{k+1} будет, вообще говоря, отличаться от прогнозируемого $\tilde{P}^L(t_{k+1})$ (см. рис. 3). Но в момент t_{k+1} становится известным относительно точное реальное положение лидера. Именно эта информация (а не прогнозированная!) будет использоваться для дальнейших расчетов в процессе движения группы аппаратов.

Расстояние между аппаратом и его позицией в каждый момент времени. Расстояние между ведомым АНПА и его позицией в строю в момент времени $t \in [t_k, t_{k+1})$ равно:

$$\|P^L(t) - P(t)\| = \left\| P^L(t_k) + (P^L(t_k) - P^L(t_{k-1})) \frac{t - t_k}{t_k - t_{k-1}} - P(t_k) - \frac{P^L(t_k) - P(t_k)}{\|P^L(t_k) - P(t_k)\|} s(t) \right\|.$$

С учетом принятого обозначения $P^L(t_k) - P(t_k) = W_1(t_k)$, запишем:

$$\|P^L(t) - P(t)\| = \left\| W_1(t_k) + (P^L(t_k) - P^L(t_{k-1})) \frac{t - t_k}{t_k - t_{k-1}} - \frac{W_1(t_k)}{\|W_1(t_k)\|} s(t) \right\|.$$

Наконец, вводя дополнительные обозначения $\Delta t_{k-1} = t_k - t_{k-1}$ и $W_2(t_{k-1}) = P^L(t_k) - P^L(t_{k-1})$ (рис. 5), получаем:

$$\|P^L(t) - P(t)\| = \left\| W_2(t_{k-1}) \frac{t - t_k}{\Delta t_{k-1}} + W_1(t_k) \left(1 - \frac{s(t)}{\|W_1(t_k)\|} \right) \right\|. \quad (5)$$

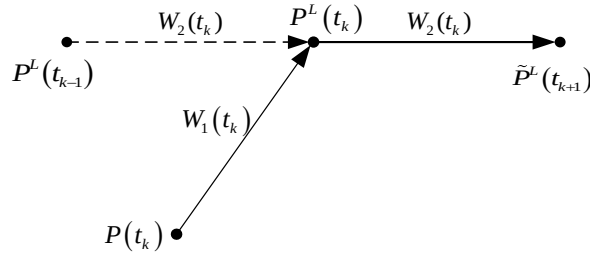


Рис. 5. Векторы движения и прогнозируемые положения лидера

Применим к последнему выражению неравенство треугольника:

$$\|P^L(t) - P(t)\| \leq \|W_2(t_{k-1})\| \frac{t - t_k}{\Delta t_{k-1}} + \|W_1(t_k)\| - s(t).$$

Эта формула тем точнее, чем меньше угол между направлениями движения АНПА и лидера. В пределе, когда направления движения совпадают, т.е. векторы W_1 и W_2 коллинеарны, получается равенство. Поскольку ведомый АНПА все время находится в режиме «догнать свою позицию в строю», то следует ожидать, что указанный угол действительно будет на всех временных участках небольшим. Следующее утверждение дает верхнюю оценку для расстояния между лидером и ведомым аппаратом.

Утверждение. $\|W_1(t_k)\| \leq (v_L + v_{\min})\Delta t_{k-1}$, где $\Delta t_{k-1} = t_k - t_{k-1}$.

Доказательство. Предположим, что в момент $t = t_{k-1}$ ведомый аппарат находится в малой окрестности точки $P^L(t_{k-1})$. Гипотетически предположим ситуацию, что в этот же момент $t = t_{k-1}$ лидер меняет свою траекторию и движется прямолинейно в направлении ведомого аппарата. Это значит, что за время $\Delta t_{k-1} = t_k - t_{k-1}$ он переместится на расстояние $v_L \Delta t_{k-1}$, перейдя в точку $P^L(t_k)$.

В свою очередь ведомый аппарат по сделанному выше предположению будет двигаться в противоположном по отношению к лидеру направлении. Поскольку после прохождения точки $P^L(t_{k-1})$ он может двигаться только с минимальной скоростью $v_{\min}$, то его максимально возможное удаление от точки $P^L(t_{k-1})$ будет гарантированно ограничено величиной $v_{\min} \Delta t_{k-1}$ (на самом деле меньшей). Точка, в которую он переместится, будет точкой $P(t_k)$.

Так как при сделанном предположении лидер и ведомый аппарат перемещаются вдоль одной и той же прямой линии, но в противоположных направлениях, то из построения следует, что максимально возможное расстояние, на которое лидер и ведомый подводный аппарат могут отдалиться друг от друга на отрезке времени $\Delta t_{k-1} = t_k - t_{k-1}$, равно $P^L(t_k) - P(t_k) = v_L \Delta t_{k-1} + v_{\min} \Delta t_{k-1} = (v_L + v_{\min})\Delta t_{k-1}$. Из этого следует справедливость доказываемого неравенства. ■

Усредненное отклонение аппарата от его позиции на отрезке времени и на всей траектории. Найдем усредненную оценку для отклонения $d([t_k, t_{k+1}))$ АНПА от целевой позиции на промежутке времени $[t_k, t_{k+1})$. Для этого необходимо вычислить интеграл на промежутке $[t_k, t_{k+1}]^{**}$ от функции, задающей расстояние между АНПА и его позицией в

**) Переход к интегралу позволяет рассматривать вместо промежутков (полуинтервалов) $[t_k, t_{k+1})$ замкнутые промежутки (отрезки) $[t_k, t_{k+1}]$.

каждый момент времени, и поделить найденное значение на $\Delta t_k = t_{k+1} - t_k$, т.е.

$d([t_k, t_{k+1}]) = \frac{1}{t_{k+1} - t_k} \int_{t_k}^{t_{k+1}} \|P^L(t) - P(t)\| dt$. Применяя найденное выше неравенство для $\|P^L(t) - P(t)\|$, запишем:

$$d([t_k, t_{k+1}]) \leq \|W_1(t_k)\| + \|W_2(t_{k-1})\| \frac{\Delta t_k}{2\Delta t_{k-1}} - \frac{1}{\Delta t_k} \int_{t_k}^{t_{k+1}} s(t) dt.$$

После подстановки в правую часть выражения (4) для функции $s(t)$ получаем:

$$d([t_k, t_{k+1}]) \leq \|W_2(t_{k-1})\| \frac{\Delta t_k}{2\Delta t_{k-1}} + \frac{1}{\gamma \Delta t_k} \int_{t_k}^{t_{k+1}} \ln \left[\alpha + \left(e^{\gamma \|W_1(t_k)\|} - \alpha \right) e^{-\gamma v_{\max}(t-t_k)} \right] dt.$$

Будем далее для простоты считать, что $\Delta t_k = \Delta t$ для всех k . В этом случае $t_k = k\Delta t$. Для случая равных промежутков доказанное выше Утверждение принимает вид: $\|W_1(t_k)\| \leq (v_L + v_{\min})\Delta t = \|W_1\|$ для всех $k = 1, 2, \dots$. Кроме того, поскольку по условию скорость лидера считается постоянной, то $\|W_2(t_k)\| = v_L \Delta t = \|W_2\|$ для всех $k = 1, 2, \dots$. Усредненная оценка расстояния не зависит от номера k и одинакова для всех промежутков равной длины, но зависит от длины этого промежутка, т.е. $d([t_k, t_{k+1}]) = d(\Delta t)$. Поэтому можно записать:

$$d([t_k, t_{k+1}]) = d(\Delta t) \leq \frac{1}{2} \|W_2\| + \frac{1}{\gamma \Delta t} \int_0^{\Delta t} \ln \left[\alpha + \left(e^{\gamma \|W_1\|} - \alpha \right) e^{-\gamma v_{\max} t} \right] dt.$$

Точность параметров, появляющихся в расчетах усредненных уклонений, не является критически важной, так как градиентный метод, лежащий в основе управления группой АНПА, будет их корректировать в процессе вычислений. В частности, можно считать, что временные отрезки лишь примерно равны, т.е. $\Delta t_k \approx \Delta t$ для всех k . Однако дискретно поступающие навигационные данные, напротив, существенно влияют на окончательные результаты, поскольку они содержат информацию о реальных траекториях подводных аппаратов.

Будем для простоты считать, что для некоторого n в точности выполняется равенство $T = n\Delta t$. Тогда для получения интегральной оценки d уклонения ведомого АНПА от его позиции на всей траектории движения за лидером необходимо просуммировать все уклонения и усреднить сумму по длине траектории. Учитывая, что уклонения по условию не зависят от номера промежутка k , интегральная оценка d уклонения на всей траектории будет совпадать с оценкой уклонения на любом временном отрезке длины Δt , Иными словами

$$d = \frac{1}{T} \sum_{k=0}^{n-1} \Delta t d(\Delta t) \leq \frac{1}{2} \|W_2\| + \frac{1}{\gamma \Delta t} \underbrace{\int_0^{\Delta t} \ln \left[\alpha + \left(e^{\gamma \|W_1\|} - \alpha \right) e^{-\gamma v_{\max} t} \right] dt}_I. \quad (6)$$

Эта формула дает хорошее приближение к точному значению. Однако она затруднительна для простой интерпретации. Поэтому проведем дополнительные оценки правой части, заведомо потеряв в точности ради более понятной интерпретируемости. Для этого рассмотрим слагаемые в правой части. Как было показано выше $\|W_2\| = v_L \Delta t$. Рассмотрим теперь интеграл I . В силу Утверждения для равных промежутков: $\|W_1\| = (v_L + v_{\min})\Delta t$. Поскольку на отрезке $[0, \Delta t]$ всегда выполняется неравенство $e^{-\gamma v_{\max} t} \leq 1$, то $\ln \left[\alpha + \left(e^{\gamma \|W_1\|} - \alpha \right) e^{-\gamma v_{\max} t} \right] \leq \gamma (v_L + v_{\min}) \Delta t$ и, значит, $I \leq \gamma (v_L + v_{\min}) (\Delta t)^2$. С учетом этого получаем

$$d \leq \frac{1}{2} v_L \Delta t + (v_L + v_{\min}) \Delta t = \left(\frac{3}{2} v_L + v_{\min} \right) \Delta t. \quad (7)$$

Данная оценка интеграла является грубой, и ее точность вначале быстро падает с ростом Δt , хотя в асимптотике и стабилизируется. Тем не менее, она позволяет сделать полезные качественные выводы. Во-первых, надо отметить, что слагаемое $v_{\min} \Delta t$ определяется параметром $v_{\min}$, задающим минимально возможную скорость движения подводного аппарата. Эта величина является технической характеристикой АНПА, на которую нет возможности влиять в процессе выполнения задания. Поскольку по физическому смыслу скорость $v_{\min}$ маленькая (можно считать, что она на порядок меньше, чем $v_{\max}$, и с большой вероятностью, чем v_L), второе слагаемое не вносит заметного вклада в верхнюю оценку уклонения на траектории. Основное влияние на количественную оценку сохранения строя в процессе движения АНПА помимо задержек вносит первое слагаемое, которое определяется скоростью лидера v_L , а именно, среднее уклонение на всей траектории уменьшается с уменьшением этой скорости, что согласуется с интуитивным ожиданием, поскольку на начальном этапе каждого участка ведомый аппарат в соответствии с моделью его движения перемещается со скоростью близкой к максимальной.

Хотя на самом деле зависимость точности сохранения строя носит нелинейный характер, полученная оценка показывает, что она условно линейно зависит от средней скорости движения аппарата-лидера и с ростом задержек уклонение растет пропорционально $\sim 1.5v_L \Delta t$. Например, если $v_L = 2$ м/с, что является правдоподобным значением при выполнении большинства задач (например, при сканировании донных поверхностей), и задержка в получении навигационной информации равна 2 с, то верхняя оценка уклонения в строе будет $\approx 1.5 \cdot 2 \cdot 2 = 6$ метров. При этом надо принимать во внимание, что скорость перемещения лидера технически ограничена максимальной скоростью $v_{\max}$ и, следовательно, уклонение ведомого аппарата не может превышать $1.5v_{\max} \Delta t$.

Формула (7) сильно завышает реальные значения, особенно при больших интервалах Δt , что вызвано грубыми оценками параметра $\|W_1\|$ и интеграла I . Следующая ниже таблица содержит конкретные значения оценок (6) и (7) уклонений подводного аппарата от его позиции в строе в зависимости от величины задержки Δt поступления навигационных данных. В последнем столбце приведены относительные погрешности $\frac{| \text{Оценка (7)} - \text{Оценка (6)} |}{\text{Оценка (6)}} \cdot 100\%$.

Оценка (6)

Таблица 1

Оценки уклонений при различных задержках

Задержка Δt , с	Оценка уклонения (6), м	Оценка уклонения (7), м	Относительная погрешность оценки (7), %
0.1	0.2418	0.3500	44.7455
0.2	0.4388	0.7000	59.5424
0.3	0.6098	1.0500	72.1919
0.4	0.7649	1.4000	83.0279
0.6	1.0488	2.1000	100.2296
0.8	1.3160	2.8000	112.7623
1.0	1.5774	3.5000	121.8778
2.0	2.8931	7.0000	141.9565
3.0	4.2457	10.5000	147.3062
4.0	5.6192	14.0000	149.1471
5.0	7.0028	17.5000	149.9006
6.0	8.3918	21.0000	150.2444
7.0	9.7839	24.5000	150.4104
8.0	11.1780	28.0000	150.4910
9.0	12.5735	31.5000	150.5277
10.0	13.9698	35.0000	150.5406

Заключение. В работе рассматривается относительно простая модель движения подводного аппарата в промежутках между получением навигационных данных. Его движение осуществляется в направлении последней известной позиции лидера. Более эффективным был бы алгоритм управления движением в предполагаемое местонахождение лидера через промежуток времени Δt . Перефразируя Уолтера Гретцки, надо перемещаться туда, где лидер будет, а не туда, где он был. Это может быть предметом следующего исследования.

Многочисленные предположения о «прямолинейности» движения лидера и ведомого аппарата упрощают анализ точности строя, но делают полученную оценку более грубой. Такие упрощения вполне приемлемы на небольших участках траектории, т.е. при небольших величинах задержек Δt , но могут исказить реальную картину при увеличении таких участков, т.е. с ростом «навигационной неопределенности». Для улучшения оценок в подобных случаях необходимо проводить более тонкий анализ. В частности, при больших задержках важно учитывать возможное искривление траектории подводного аппарата-лидера. С этой целью целесообразно осуществлять экстраполяцию его траектории не по двум значениям, а например, по трем или даже четырем, если известно, что траектории движения лидера не предполагают резких поворотов.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. *Инзарцев А.В., Киселев Л.В., Костенко В.В. и др.* Подводные робототехнические комплексы: системы, технологии, применение. – Владивосток: Изд-во ФГБУН Ин-т проблем морских технологий ДВО РАН, 2018. – 368 с.
2. *Инзарцев А.В., Киселев Л.В., Матвиенко Ю.В.* Навигация и управление автономных подводных роботов // Известия ЮФУ. Технические науки. – 2010. – № 3 (104). – С. 164-169.
3. *Yang Y., Xiao Y., Li T.* A Survey of Autonomous Underwater Vehicle Formation: Performance, Formation Control, and Communication Capability // IEEE Communications Surveys & Tutorials. – 2021. – Vol. 23, No. 2. – P. 815-841.
4. *Hadi B.* Adaptive Formation Motion Planning and Control of Under-Actuated Autonomous Underwater Vehicles Using Deep Reinforcement Learning. – 2023.
5. *Schaub H., Vadali S.R., Alfriend K.T.* Spacecraft formation flying control using mean orbit elements // Journal of the Astronautical Sciences. – 2000. – Vol. 48, No. 1. – P. 69-87.
6. *Das B., Subudhi B., Pati B.B.* Cooperative Formation Control of Autonomous Underwater Vehicles: An Overview // International Journal of Automation and Computing. – 2016. – Vol. 13, No. 3. – P. 199-225.
7. *Hadi B., Kharrati H., Tatar Y.* A Review of the Path Planning and Formation Control for Autonomous Underwater Vehicles // Journal of Intelligent & Robotic Systems. – 2021. – Vol. 101. – Article 44.
8. *Das A.K., Fierro R., Kumar V., Ostrowski J.P., Spletzer J., Taylor C.J.* A vision-based formation control framework // IEEE Transactions on Robotics and Automation. – 2002. – Vol. 18, No. 5. – P. 813-825.
9. *Lawton J.R., Beard R.W., Young B.J.* A decentralized approach to formation maneuvers // IEEE Transactions on Robotics and Automation. – 2003. – Vol. 19, No. 6. – P. 933-941.
10. *Lewis M.A., Tan K.-H.* High precision formation control of mobile robots using virtual structures // Autonomous Robots. – 1997. – Vol. 4, No. 4. – P. 387-403.
11. *Li L., Yan Z., Yu S. et al.* Formation Control of Multiple Autonomous Underwater Vehicles under Communication Delay, Packet Discreteness and Dropout // Journal of Marine Science and Engineering. – 2022. – Vol. 10, No. 7. – Article 920.
12. *Yan T. et al.* Formation Control of Multiple Autonomous Underwater Vehicles: A Comprehensive Overview of Recent Advances // Intelligence & Robotics. – 2023.
13. *Chen Y., Li C., Chen J. et al.* A Formation Control Method for AUV Group Under Communication Constraints // Frontiers in Bioengineering and Biotechnology. – 2022. – Vol. 10. – Article 848641.
14. *Bechlioulis C.P., Karras G.C., Heshmati-Alamdari S., Kyriakopoulos K.J.* Robust Formation Control for Multiple Underwater Vehicles // Frontiers in Robotics and AI. – 2019. – Vol. 6. – Article 90.
15. *Fan X., Guo Y., Liu H., Wei B., Lyu W.* Improved Artificial Potential Field Method Applied for AUV Path Planning // Mathematical Problems in Engineering. – 2020. – Article 6523158. – 21 p. – DOI: 10.1155/2020/6523158.
16. *Белоглазов Д.А., Гайдук А.Р., Косенко Е.Ю. и др.* Групповое управление подвижными объектами в неопределенных средах / под ред. В.Х. Пшихопова. – М.: Физматлит, 2015. – 304 с.
17. *Филимонов А.Б., Филимонов Н.Б., Бараишков А.А.* Вопросы построения потенциальных полей в задачах локальной навигации мобильных роботов // Автометрия. – 2019. – Т. 55, № 4. – С. 65-70.

18. Guldner J., Utkin V.I. Tracking the gradient of artificial potential fields: Sliding mode control for mobile robots // *International Journal of Control*. – 1996. – Vol. 63, No. 3. – P. 417-432.
19. Yan Z., Li L., Yu S. et al. Adaptive Formation Control of Unmanned Underwater Vehicles with Collision Avoidance and Environment Adaptation // *Journal of Marine Science and Engineering*. – 2022. – Vol. 10, No. 4. – Article 516.
20. Xia G., Zhang Y., Zhang W., Chen X., Yang H. Multi-timescale 3-D Coordinated Formation Control for Multi-Underactuated AUV with Uncertainties: Design and Stability Analysis Using Singular Perturbation Methods // *Ocean Engineering*. – 2021. – Vol. 230. – Article 109053. – DOI: 10.1016/j.oceaneng.2021.109053.
21. Каркищенко А.Н. Градиентный метод формирования строя подводных аппаратов на основе вероятностного функционала // *Мехатроника, автоматизация, управление*. – 2025. – Т. 26, № 4. – С. 209-219.
22. Zhen Q. et al. Formation Control of a Multi-AUVs System Based on Virtual Structure and Artificial Potential Field // *Ocean Engineering*. – 2022.

REFERENCES

1. Inzartsev A.V., Kiselev L.V., Kostenko V.V., Matvienko Yu.V., Pavin A.M., Shcherbatyuk A.F. Podvodnye robototekhnicheskie komplekсы: sistemy, tekhnologii, primeneniye [Underwater robotic complexes: systems, technologies, application], Izd-vo FGBUN In-t problem morskikh tekhnologiy DVO RAN, 2018, 368 p.
2. Inzartsev A.V., Kiselev L.V., Matvienko Yu.V. Navigatsiya i upravleniye avtonomnykh podvodnykh robotov [Navigation and control of autonomous underwater robots], *Izvestiya YuFU. Tekhnicheskie nauki* [Izvestiya SFedU. Engineering Sciences], 2010, No. 3 (104), pp. 164-169.
3. Yang Y., Xiao Y., Li T. A Survey of Autonomous Underwater Vehicle Formation: Performance, Formation Control, and Communication Capability, *IEEE Communications Surveys & Tutorials*, 2021, Vol. 23, No. 2, pp. 815-841.
4. Hadi B. Adaptive Formation Motion Planning and Control of Under-Actuated Autonomous Underwater Vehicles Using Deep Reinforcement Learning, 2023.
5. Schaub H., Vadali S.R., Alfriend K.T. Spacecraft formation flying control using mean orbit elements, *Journal of the Astronautical Sciences*, 2000, Vol. 48, No. 1, pp. 69-87.
6. Das B., Subudhi B., Pati B.B. Cooperative Formation Control of Autonomous Underwater Vehicles: An Overview, *International Journal of Automation and Computing*, 2016, Vol. 13, No. 3, pp. 199-225.
7. Hadi B., Kharrati H., Tatar Y. A Review of the Path Planning and Formation Control for Autonomous Underwater Vehicles, *Journal of Intelligent & Robotic Systems*, 2021, Vol. 101, Article 44.
8. Das A.K., Fierro R., Kumar V., Ostrowski J.P., Spletzer J., Taylor C.J. A vision-based formation control framework, *IEEE Transactions on Robotics and Automation*, 2002, Vol. 18, No. 5, pp. 813-825.
9. Lawton J.R., Beard R.W., Young B.J. A decentralized approach to formation maneuvers, *IEEE Transactions on Robotics and Automation*, 2003, Vol. 19, No. 6, pp. 933-941.
10. Lewis M.A., Tan K.-H. High precision formation control of mobile robots using virtual structures, *Autonomous Robots*, 1997, Vol. 4, No. 4, pp. 387-403.
11. Li L., Yan Z., Yu S. et al. Formation Control of Multiple Autonomous Underwater Vehicles under Communication Delay, Packet Discreteness and Dropout, *Journal of Marine Science and Engineering*, 2022, Vol. 10, No. 7, Article 920.
12. Yan T. et al. Formation Control of Multiple Autonomous Underwater Vehicles: A Comprehensive Overview of Recent Advances, *Intelligence & Robotics*, 2023.
13. Chen Y., Li C., Chen J. et al. A Formation Control Method for AUV Group Under Communication Constraints, *Frontiers in Bioengineering and Biotechnology*, 2022, Vol. 10, Article 848641.
14. Bechlioulis C.P., Karras G.C., Heshmati-Alamdari S., Kyriakopoulos K.J. Robust Formation Control for Multiple Underwater Vehicles, *Frontiers in Robotics and AI*, 2019, Vol. 6, Article 90.
15. Fan X., Guo Y., Liu H., Wei B., Lyu W. Improved Artificial Potential Field Method Applied for AUV Path Planning, *Mathematical Problems in Engineering*, 2020, Article 6523158, 21 p. DOI: 10.1155/2020/6523158.
16. Beloglazov D.A., Gayduk A.R., Kosenko E.Yu. i dr. Gruppovoe upravleniye podvizhnymi ob"ektami v neopredelennykh sredakh [Group control of moving objects in uncertain environments], ed. by V.Kh. Pshikhopova. Moscow: Fizmatlit, 2015, 304 p.
17. Filimonov A.B., Filimonov N.B., Barashkov A.A. Voprosy postroeniya potentsial'nykh poley v zadachakh lokal'noy navigatsii mobil'nykh robotov [Issues of constructing potential fields in problems of local navigation of mobile robots], *Avtometriya* [Autometria], 2019, Vol. 55, No. 4, pp. 65-70.
18. Guldner J., Utkin V.I. Tracking the gradient of artificial potential fields: Sliding mode control for mobile robots, *International Journal of Control*, 1996, Vol. 63, No. 3, pp. 417-432.

19. Yan Z., Li L., Yu S. *et al.* Adaptive Formation Control of Unmanned Underwater Vehicles with Collision Avoidance and Environment Adaptation, *Journal of Marine Science and Engineering*, 2022, Vol. 10, No. 4, Article 516.
20. Xia G., Zhang Y., Zhang W., Chen X., Yang H. Multi-timescale 3-D Coordinated Formation Control for Multi-Underactuated AUV with Uncertainties: Design and Stability Analysis Using Singular Perturbation Methods, *Ocean Engineering*, 2021, Vol. 230, Article 109053. DOI: 10.1016/j.oceaneng.2021.109053.
21. Karkishchenko A.N. Gradientnyy metod formirovaniya stroya podvodnykh apparatov na osnove veroyatnostnogo funktsionala [Gradient method for forming a formation of underwater vehicles based on a probabilistic functional], *Mekhatronika, avtomatizatsiya, upravlenie* [Mechatronics, Automation, Control], 2025, Vol. 26, No. 4, pp. 209-219.
22. Zhen Q. *et al.* Formation Control of a Multi-AUVs System Based on Virtual Structure and Artificial Potential Field, *Ocean Engineering*, 2022.

Каркищенко Александр Николаевич – НИИ Робототехники и процессов управления; e-mail: karkishalex@gmail.com; г. Таганрог, Россия; тел.: +79034610539; д.ф.-м.н.; профессор; г.н.с.

Karkishchenko Alexander Nikolaevich – Scientific Research Institute of the Robotics and Control Processes; e-mail: karkishalex@gmail.com; Taganrog, Russia; phone: +79034610539; dr. of phys. and math. sc.; professor; chief researcher.

УДК 681.51: 676.082.4

DOI 10.18522/2311-3103-2026-4-226-241

Д.А. Ковалёв, Д.Х. Имаев, С.Е. Душин

МОДЕЛИРОВАНИЕ И ИДЕНТИФИКАЦИЯ СОДОРЕГЕНЕРАЦИОННОГО КОТЛОАГРЕГАТА КАК ОБЪЕКТА УПРАВЛЕНИЯ

Ключевой агрегат производства сульфатной целлюлозы (крафт-процесса) – содорегенерационный котлоагрегат (СРК) – обеспечивает восстановление химикатов черного щелока. Актуальна задача повышения степени восстановления щелока без изменения технологии процесса и конструкции аппаратов, т.е. информационно-алгоритмическими средствами автоматического управления процессом. Более высокое гарантированное качество восстановления обеспечивают алгоритмы регулирования, синтезированные на основе адекватных математических моделей. Цель статьи – системный анализ СРК как объекта управления и построение усовершенствованной математической модели, пригодной для синтеза высококачественных систем автоматической стабилизации. В основе методологии лежит принцип последовательного раскрытия неопределённости моделей сложных систем управления (А.А. Вавилов). В отличие от известных подходов, динамика СРК в окрестности выбранного режима описана в виде многоуровневой модели класса LTI. Комбинирование аналитических методов с обработкой экспериментальных данных для конкретного режима позволило получить определенные до значений параметров математические модели реального СРК. Разработанная многоуровневая линейная стационарная модель СРК дает возможность анализировать динамику подсистем и агрегата в целом, устойчивость режимов и влияние возмущений, и обосновать необходимость автоматического поддержания требуемого качества технологических процессов в окрестности выбранного режима. Модели в символьном виде применимы к большинству подобных агрегатов технологического процесса производства сульфатной целлюлозы. Параметры полностью определенных моделей должны быть скорректированы по результатам обработки данных, относящихся к конкретному агрегату и режиму. Направление дальнейших исследований связано с развитием методов анализа динамики сложных систем управления по моделям иерархической структуры, а также синтеза систем стабилизации режима и ослабления влияния возмущений.

Содорегенерационный котлоагрегат; степень восстановления; черный щелок; моделирование; идентификация; система автоматического регулирования.

D.A. Kovalev, D.H. Imaev, S.E. Dushin

MODELING AND IDENTIFICATION OF A COGENERATION BOILER UNIT FOR KRAFT PULP PRODUCTION AS A CONTROL PLANT

The key unit of kraft pulp production – the recovery boiler – ensures the regeneration of black liquor chemicals. The task of increasing the liquor reduction degree without changing the process technology or equipment design, i.e., through information and algorithmic means – automatic control is relevant. A higher

guaranteed quality of recovery is provided by control algorithms synthesized on the basis of adequate mathematical models. The aim of this paper is a system analysis of the recovery boiler as a control plant and the development of an enhanced mathematical model suitable for synthesizing high-quality automatic stabilization systems. The methodology is based on the principle of sequential uncertainty removal for complex control system models (A.A. Vavilov). Unlike known approaches, the dynamics of the recovery boiler in the vicinity of a selected operating point are described as a multi-level LTI-class model. Combining analytical methods with experimental data processing for a specific operating mode made it possible to obtain mathematical models of an actual recovery boiler with parameters determined down to specific numerical values. The developed multilevel linear stationary model of the recovery boiler makes it possible to analyze the dynamics of subsystems and the unit as a whole, the stability of modes and the influence of disturbances, and justify the need to automatically maintain the required quality of technological processes in the vicinity of the selected mode. Symbolic models are applicable to most similar kraft process units. Parameters of fully defined models should be adjusted based on the results of processing data related to a specific unit and mode. The direction of further research is associated with the development of methods for analyzing the dynamics of complex control systems according to models of a hierarchical structure, as well as the synthesis of systems for stabilizing the regime and reducing the influence of perturbations, as well as the synthesis of systems for regime stabilization and mitigation of the influence of disturbances.

Recovery boiler; reduction degree; black liquor; modeling; identification; automatic control system.

Введение. Производство целлюлозы сульфатным способом сводится к варке древесной щепы в растворе сульфида натрия и гидроксида натрия. Так называемый крафт-процесс обладает рядом достоинств, к которым относятся высокая прочность волокна, возможность переработки древесины различных пород и экономичность [1–3].

Важнейшей частью комплекса производства сульфатной целлюлозы является содо-регенерационный котлоагрегат (СРК), который позволяет отделить химикаты от органических веществ и вернуть их в процесс варки. В результате получается целлюлоза и так называемый черный щелок, содержащий отработанные химикаты, растворенный лигнин и другие органические вещества.

Режимы функционирования СРК определяются Технологическими регламентами и Режимными картами [4–6]. Устойчивость режимов и их инвариантность к малым по уровню возмущениям обеспечивают системы автоматического регулирования, реализующие принцип обратной связи. Технологии автоматизации позволяют достичь целей управления и регулирования информационно-алгоритмическими средствами без необходимости изменения технологии восстановления щелоков и конструкции СРК. Цель регулирования СРК, прежде всего, заключается в поддержании минимально допустимой степени восстановления сульфата натрия. Более высокая степень восстановления щелока при тех же свойствах СРК достигается за счет добавляемого в сухом виде сульфата натрия, что удорожает процесс.

Традиционные системы автоматической стабилизации не позволяют в точности достичь цели регулирования, поставленной в виде ограничения. Приближенное (практическое) решение заключается в подмене цели регулирования – поддержания степени восстановления на уровне, несколько превышающем ограничение. Однако приближение заданного уровня восстановления к ограничению без его нарушения требует уменьшения ошибки системы стабилизации. Таким образом, технический показатель – качество процессов регулирования – напрямую определяет показатель экономический.

Системы автоматического управления и регулирования применяют специализированное программное обеспечение, разработанное на основе математического описания объекта управления, среды его функционирования и требований к процессам достижения цели. Более высокое гарантированное качество регулирования обеспечивают качественные алгоритмы. Их синтез основан на более адекватных математических моделях СРК, построенных в результате детального исследования технологических процессов, комбинирования аналитических методов и обработки экспериментальных данных.

В период с 1960 по 1990 гг. отечественные исследователи активно занимались разработкой технологии процессов в СРК, их математическим моделированием и проблемами автоматизации управления [7–13]. В настоящее время подобными исследованиями в основном занимаются зарубежные ученые [14–24], среди которых особо выделяются работы финских специалистов.



Рис. 2. СРК как ОУ

Обозначения технологических параметров – переменных математической модели – сведены в табл. 1.

Таблица 1

Параметры технологических процессов СРК

Обозн.	Наименование параметра	Размерность
$G_{гр}$	Массовый расход греющего пара, подаваемого в выпарную станцию	кг/с
G_s	Массовый расход сульфата натрия, добавленного в чёрный щёлок	кг/с
$G_{п}$	Массовый расход греющего пара, подаваемого в щелоковый подогреватель	кг/с
G_2	Объемный расход чёрного щёлока на выходе из щелокового подогревателя	м ³ /с
Q_{01}	Объемный расход первичного воздуха, подаваемого в топку	нм ³ /с
Q_{02}	Объемный расход вторичного воздуха, подаваемого в топку	нм ³ /с
Q_{03}	Объемный расход третичного воздуха, подаваемого в топку	нм ³ /с
G_{20}	Массовый расход греющего пара подаваемого в подогреватель воздуха	кг/с
G_2	Массовый расход питательной воды, подаваемой в экономайзер	кг/с
G_{31}	Массовый расход охлаждающей воды, подаваемой в поверхностный пароохладитель	кг/с
G_{22}	Массовый расход чёрного щёлока, подаваемого на сжигание в топку СРК	кг/с
$U_{эф}$	Напряжения коронирующих электродов электрофилтра	кВ
Q_{51}	Объемный расход дымовых газов на выходе из электрофилтра	нм ³ /ч
G_{43}	Массовый расход зеленого щёлока, из бака-растворителя плава СРК	кг/с
G_{42}	Массовый расход слабого белого щёлока, подаваемого в бак-растворитель плава СРК	кг/с
$\rho_{вх}$	Плотность чёрного щёлока на входе в выпарную станцию	кг/м ³
ϑ_{11}	Температура чёрного щёлока на входе в щелоковый подогреватель	°С
ϑ_0	Температура воздуха на входе в подогреватель воздуха	°С
G_{25}	Массовый расход перегретого пара на выходе второй части пароперегревателя	кг/с
$Q_{эф}$	Объемный расход дымовых газов на входе в электрофилтр	нм ³ /ч
Q_{oc}	Объемный расход уловленных пылевых частиц в электрофилтре	нм ³ /ч
G_{41}	Массовый расход плава, поступающего в бак-растворитель плава из топки СРК	кг/с
G_{44}	Массовый расход пылегазовых выбросов, из бака-растворителя плава СРК	кг/с

Окончание табл. 1

Обозн.	Наименование параметра	Размерность
ϑ_{25}	Температура перегретого пара на выходе второй части пароперегревателя	$^{\circ}\text{C}$
ρ_1	Плотность дымовых газов на выходе из электрофильтра	кг/м^3
ρ_{43}	Плотность зеленого щёлка на выходе из бака-растворителя плава СРК	кг/м^3
ξ	Степень восстановления сульфата натрия	%
$Q_{\text{вх}}$	Объёмный расход чёрного щёлка на входе в выпарную станцию	$\text{м}^3/\text{с}$
Q_0	Объёмный расход воздуха подаваемого в подогреватель воздуха	$\text{нм}^3/\text{с}$
ϑ_{01}	Температура первичного воздуха, подаваемого в топку	$^{\circ}\text{C}$
ϑ_{02}	Температура вторичного воздуха, подаваемого в топку	$^{\circ}\text{C}$
ϑ_{03}	Температура третичного воздуха, подаваемого в топку	$^{\circ}\text{C}$
ϑ_{01}	Температура первичного воздуха, подаваемого в топку	$^{\circ}\text{C}$
$\vartheta_{\text{т}}$	Температура уходящих дымовых газов из топки СРК	$^{\circ}\text{C}$
$Q_{\text{т}}$	Объёмный расход уходящих дымовых газов из топки СРК	$\text{нм}^3/\text{с}$
$\rho_{\text{эф}}$	Плотность дымовых газов на входе в электрофильтр	кг/нм^3
$\vartheta_{\text{эф}}$	Температура дымовых газов на входе в электрофильтр	$^{\circ}\text{C}$
ϑ_1	Температура дымовых газов на выходе из электрофильтра	$^{\circ}\text{C}$

На рис 1 возмущения среды функционирования: $\rho_{\text{вх}}$; ϑ_{11} ; ϑ_0 ; G_{25} ; $Q_{\text{эф}}$; $Q_{\text{ос}}$; G_{41} ; G_{44} . Потенциальными управляющими (регулирующими) воздействиями являются: $G_{\text{гр}}$; G_s ; $G_{\text{п}}$; Q_2 ; Q_{01} ; Q_{02} ; Q_{03} ; G_{20} ; G_2 ; G_{31} ; G_{22} ; $U_{\text{эф}}$; Q_1 ; G_{43} ; G_{42} .

В соответствии с принципом последовательного раскрытия неопределенности моделей ССУ [25] – [27] перечисление переменных дает модель нулевого ранга неопределенности ($R = 0$).

Попытки централизованного управления, основанные на моделировании СРК в целом, вызывают множество практических и теоретических проблем. Необходимость функционально-целевой декомпозиции требуют представления СРК как сложной системы, образованной взаимодействием подсистем: подготовки черного щелока к сжиганию PS1, подготовки воздуха PS2, сжигания черного щелока PS3, парообразования и питательной воды PS4, очистки дымовых газов PS5, зеленого и белого щелока PS6.

Диаграмма ориентированного гиперграфа – несимметричного отношения на множестве многомерных подсистем типа Multiple Input – Multiple Output (MIMO) – изображена на рис. 3.

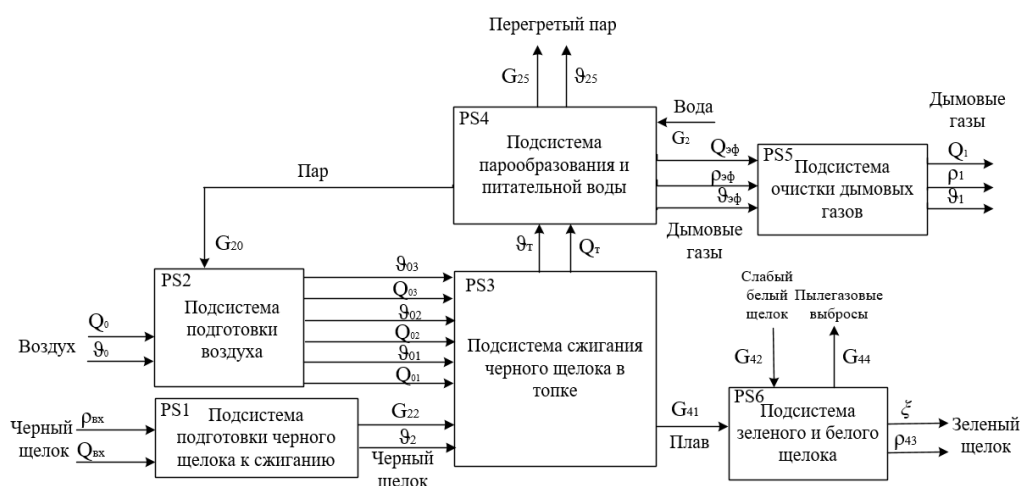


Рис. 3. СРК как объект управления сложной структуры

По терминологии системного подхода А. А. Вавилова [25] – [26] получена модель СРК первого ранга неопределенности ($R = 1$). Дальнейшее раскрытие неопределенности достигается раскрытием неопределенности моделей подсистем. Если ТП регенерации щелоков рассматривать как объект управления L -го уровня иерархии, то СРК – это подсистема ($L-1$)-го уровня; перечисленные ранее подсистемы – компоненты ($L-2$)-го уровня.

3. Моделирование подсистем СРК. Ввиду малости отклонений переменных от выбранного режима далее рассматриваются линейные стационарные модели (класса LTI – Linear Time-Invariant) в виде обыкновенных дифференциальных уравнений с постоянными коэффициентами [4, 7–13].

1. Подсистема подготовки черного щелока к сжиганию PS1 включает процессы выпаривания, смешения с сухим сульфатом натрия, подогрева и подачи в форсунки распыления (рис. 4). В табл. 2 сведены технологические параметры процесса выпаривания черного щелока.

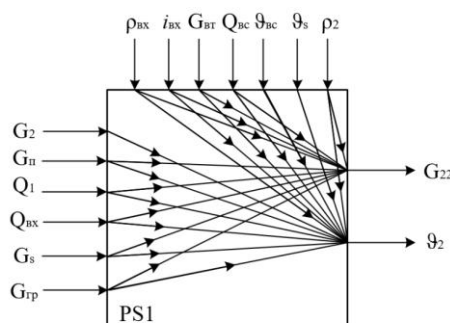


Рис. 4. Подсистема подготовки черного щелока к сжиганию

Таблица 2

Параметры процесса выпаривания черного щелока

Обозначение	Наименование параметра	Размерность
$\rho_{вс}$	Черный щёлок на выходе из выпарной станции	кг/м ³
$i_{вх}$	Удельная энтальпия чёрного щёлока на входе в выпарную станцию	кДж/кг
$i_{вс}$	Удельная энтальпия чёрного щёлока на выходе из выпарной станции	кДж/кг
$Q_{вс}$	Объёмный расход чёрного щёлока на выходе из выпарной станции	м ³ /с
$i_{гр}$	Удельная энтальпия пара, подаваемого в выпарную станцию	кДж/кг
$q_{гр}$	Удельная энтальпия отводимого конденсата из выпарной станции	кДж/кг
$G_{вт}$	Массовый расход вторичного пара на выходе из выпарной станции	кг/с
$i_{вт}^п$	Удельная энтальпия вторичного пара, отводимого из выпарной станции	кДж/кг
$V_{вс}$	Объём рабочего пространства выпарной станции	м ³

Процесс выпаривания черного щелока описывается дифференциальными уравнениями (см. [4, 6, 8]):

$$T_{вс} \frac{d\rho_{вс}}{dt} + \Delta\rho_{вс} = K_{Q_{вх}} \Delta Q_{вх} + K_{\rho_{вх}} \Delta\rho_{вх} + K_{i_{вх}} \Delta i_{вх} + K_{G_{гр}} \Delta G_{гр} - K_{G_{вт}} \Delta G_{вт} - K_{Q_{вс}} \Delta Q_{вс}, \quad (1)$$

где $T_{вс} = V_{вс}/Q_{вс}$, $A = \left[\rho_{вс} \left(\frac{\partial i_{вс}}{\partial \rho_{вс}} \right) + i_{вс} \right] Q_{вс}$, $K_{Q_{вх}} = \frac{\rho_{вх} i_{вх}}{A}$, $K_{\rho_{вх}} = \frac{Q_{вх} i_{вх}}{A}$, $K_{i_{вх}} = \frac{Q_{вх} \rho_{вх}}{A}$, $K_{G_{гр}} = \frac{i_{гр} q_{гр}}{A}$, $K_{G_{вт}} = (i_{вт}^п - i_{вс})/A$, $K_{Q_{вс}} = \rho_{вс} - i_{вс}/A$.

Процесс смешения черного щелока с сульфатом натрия в баке-смесителе описывается уравнениями [4, 6, 8]:

$$T_{\vartheta_{11}} \frac{d\Delta\vartheta_{11}}{dt} + \Delta\vartheta_{11} = a_1 \Delta Q_{\text{вс}} + a_2 \Delta \rho_{\text{вс}} + a_3 \Delta \vartheta_{\text{вс}} - a_4 \Delta G_s + a_5 \Delta \vartheta_s + a_6 \Delta Q_{21} + a_7 \Delta \rho_2 + a_8 \Delta \vartheta_2, \quad (2)$$

где $T_{\vartheta_{11}} = \frac{V}{Q_1}$, $a_1 = \frac{\rho_{\text{вс}}(i_{\text{вс}} - i_1)}{Q_1 \rho_{11} (\frac{\partial i_1}{\partial \vartheta_{11}})}$, $a_2 = \frac{Q_{\text{вс}}(i_{\text{вс}} - i_1)}{Q_1 \rho_{11} (\frac{\partial i_1}{\partial \vartheta_{11}})}$, $a_3 = \frac{Q_{\text{вс}} \rho_{\text{вс}} (\frac{\partial i_{\text{вс}}}{\partial \vartheta_{\text{вс}}})}{Q_1 \rho_{11} (\frac{\partial i_1}{\partial \vartheta_{11}})}$, $a_4 = \frac{i_1 - q_s}{Q_1 \rho_{11} (\frac{\partial i_1}{\partial \vartheta_{11}})}$,
 $a_5 = \frac{G_s (\frac{\partial q_s}{\partial \vartheta_s})}{Q_1 \rho_{11} (\frac{\partial i_1}{\partial \vartheta_{11}})}$, $a_6 = \frac{\rho_2 (i_2 - i_1)}{Q_1 \rho_{11} (\frac{\partial i_1}{\partial \vartheta_{11}})}$, $a_7 = \frac{Q_{21} (i_2 - i_1)}{Q_1 \rho_{11} (\frac{\partial i_1}{\partial \vartheta_{11}})}$, $a_8 = \frac{Q_{21} \cdot 2 (\frac{\partial i_2}{\partial \vartheta_2})}{Q_1 \rho_{11} (\frac{\partial i_1}{\partial \vartheta_{11}})}$.

Параметры процесса смешения сведены в табл. 3.

Таблица 3

Параметры процесса смешения

Обозначение	Наименование параметра	Размерность
Q_{21}	Объёмный расход чёрного щелока, подаваемого в бак-смеситель из рециркуляционной линии на выходе щелокового подогревателя	м ³ /с
Q_1	Объёмный расход чёрного щелока на выходе из бака-смесителя чёрного щелока с сульфатом натрия	м ³ /с
ρ_{11}	Плотность чёрного щелока на выходе из бака-смесителя	кг/м ³
$\vartheta_{\text{вс}}$	Температура чёрного щелока на выходе из выпарной станции	°С
ϑ_s	Температура сульфата натрия, добавляемого в бак-смеситель	°С
ρ_2	Плотность чёрного щелока на выходе из щелокового подогревателя	кг/м ³
ϑ_2	Температура чёрного щелока на выходе из щелокового подогревателя	°С
V	Объём рабочего пространства из бака-смесителя чёрного щелока с сульфатом натрия	м ³
i_1	Удельная энтальпия чёрного щелока на выходе из бака-смесителя чёрного щелока с сульфатом натрия	кДж/кг
q_s	Удельная энтальпия сульфата натрия, добавляемого в бак-смеситель	кДж/кг
i_2	Удельная энтальпия чёрного щелока на выходе из щелокового подогревателя	кДж/кг

Процесс подогрева черного щелока описывается уравнениями [4, 6, 8]:

$$T_{\vartheta_2} \frac{d\Delta\vartheta_2}{dt} + \Delta\vartheta_2 = K_1 \Delta Q_1 + K_2 \Delta \vartheta_{11} + K_3 \Delta G_{\text{п}} - K_4 \Delta Q_2, \quad (3)$$

где $T_{\vartheta_2} = V_{\text{п}}/Q_2$, $A_1 = Q_2 \left[\rho_2 \left(\frac{\partial i_2}{\partial \vartheta_2} \right) + i_2 \left(\frac{\partial \rho_2}{\partial \vartheta_2} \right) \right]$, $K_1 = \rho_1 i_1 / A_1$, $K_2 = Q_1 B_1 / A_1$, $K_3 = i_{\text{п}} / A_1$,
 $K_4 = \rho_2 i_2 / A_1$, $B_1 = \left[\rho_1 \left(\frac{\partial i_1}{\partial \vartheta_{11}} \right) + i_1 \left(\frac{\partial \rho_1}{\partial \vartheta_{11}} \right) \right]$.

Параметры процесса подогрева черного щелока сведены в табл. 4.

Таблица 4

Параметры процесса подогрева черного щелока

Обозначение	Наименование параметра	Размерность
Q_2	Объёмный расход чёрного щелока на выходе из щелокового подогревателя	м ³ /с
$V_{\text{п}}$	Объём рабочего пространства щелокового подогревателя	м ³ /с
$i_{\text{п}}$	Удельная энтальпия пара, подаваемого в щелоковый подогреватель	кДж/кг

Процесс подачи черного щелока в форсунки можно описать в виде:

$$T_{G_{22}} \frac{d\Delta G_{22}}{dt} + \Delta G_{22} = a_{21}Q_1 + a_{22}G_{\text{п}} - a_{23}\rho_2 + a_{24}\rho_1, \quad (4)$$

где $T_{G_{22}} = V_{\text{п}}/Q_2$, $a_{21} = \rho_2/\rho_1$, $a_{22} = 1/\rho_2$, $a_{23} = Q_2/\rho_2$, $a_{24} = Q_1/\rho_2$.

Для остальных подсистем СРК разработаны аналогичные системы дифференциальных уравнений (см. [4, 7–13]).

Раскрытие структур подсистем СРК ($L-2$)-го уровня дает модели второго ранга ($R = 2$).

2. Подсистема сжигания черного щелока в топке СРК (PS3) включает процессы: карбонизации едкого натра; карбонизации сульфида натрия; восстановления сульфата натрия в сульфид натрия; сгорания углерода и водорода; горения органической серы и сероводорода; регенерации серы и температурного режима в топочном пространстве СРК.

Процесс карбонизации едкого натра, можно описать в виде [4, 7, 9]:

$$T_{Na_2CO_3}^{11} \frac{d\Delta G_{Na_2CO_3}^{11}}{dt} + \Delta G_{Na_2CO_3}^{11} = k_{NaOH}^{\text{чщ}} \Delta G_{NaOH}^{\text{чщ}}, \quad (5)$$

где $T_{Na_2CO_3}^{11} = \frac{x_{NaOH}}{x_{Na_2CO_3}} \tau_{NaOH}^{\text{чщ}}$; $k_{NaOH}^{\text{чщ}} = \left(1 + \frac{x_{CO_2} - x_{H_2O}}{x_{NaOH}}\right)$.

Параметры процесса карбонизации едкого натра сведены в табл. 5.

Таблица 5

Параметры процесса карбонизации едкого натра

Обозначение	Наименование параметра	Размерность
$G_{Na_2CO_3}^{11}$	Массовый расход карбоната натрия в зоне реакции карбонизации	кг/с
$G_{NaOH}^{\text{чщ}}$	Массовый расход едкого натра в зоне реакции карбонизации	кг/с
x_{NaOH}	Стехиометрический коэффициент едкого натра	
$x_{Na_2CO_3}$	Стехиометрический коэффициент карбоната натрия	
x_{CO_2}	Стехиометрический коэффициент углекислого газа	
x_{H_2O}	Стехиометрический коэффициент воды	
$\tau_{NaOH}^{\text{чщ}}$	Время прохождения капель черного щелока зоны реакции карбонизации	с

Процесс карбонизации сульфида натрия представим в виде [4, 7, 9]:

$$T_{Na_2CO_3}^{12} \frac{d\Delta G_{Na_2CO_3}^{12}}{dt} + \Delta G_{Na_2CO_3}^{12} = k_{Na_2S}^{\text{чщ}} \Delta G_{Na_2S}^{\text{чщ}}, \quad (6)$$

где $T_{Na_2CO_3}^{12} = \frac{x_{Na_2S}}{x_{Na_2CO_3}} \tau_{Na_2S}^{\text{чщ}}$; $k_{Na_2S}^{\text{чщ}} = \left(1 + \frac{x_{H_2O} + x_{CO_2} + x_{H_2S}}{x_{Na_2S}}\right)$.

Параметры процесса карбонизации сульфида натрия сведены в табл. 6.

Таблица 6

Параметры процесса карбонизации сульфида натрия

Обозначение	Наименование параметра	Размерность
$G_{Na_2CO_3}^{12}$	Массовый расход карбоната натрия в зоне реакции карбонизации	кг/с
$G_{Na_2S}^{\text{чщ}}$	Массовый расход сульфида натрия в зоне реакции карбонизации	кг/с
x_{Na_2S}	Стехиометрический коэффициент сульфида натрия	
x_{H_2S}	Стехиометрический коэффициент сероводорода	
$\tau_{Na_2S}^{\text{чщ}}$	Время прохождения капель черного щелока зоны реакции карбонизации	с

Процесс восстановления сульфата натрия в сульфид натрия представим в виде [4, 7, 9]:

$$(T_{Na_2S}^{41}p + 1)\Delta G_{Na_2S}^{41} = k_{Na_2SO_4}^{32}\Delta G_{Na_2SO_4}^{32}, \quad (7)$$

где $k_{Na_2SO_4}^{32} = (1 + \frac{x_C - x_{CO_2}}{x_{Na_2SO_4}})$.

Параметры процесса восстановления сульфата натрия сведены в табл. 7.

Таблица 7

Параметры процесса восстановления сульфата натрия

Обозначение	Наименование параметра	Размерность
$G_{Na_2S}^{41}$	Массовый расход сульфата натрия в зоне реакции регенерации серы	кг/с
x_C	Стехиометрический коэффициент углерода	
$x_{Na_2SO_4}$	Стехиометрический коэффициент сульфата натрия	

Процессы сгорания углерода, водорода, а также органической серы и сероводорода опишем в виде [4, 7, 9]:

$$(T_{CO_2}^{21}p + 1)\Delta G_{CO_2}^{21} = k_C^{21}\Delta G_C^{21} + k_{O_2}^{211}\Delta G_{O_2}^{211}, \quad (8)$$

$$(T_{H_2O}^{212}p + 1)\Delta G_{H_2O}^{212} = k_{H_2}^{чщ}\Delta G_{H_2}^{чщ} + k_{O_2}^{212}\Delta G_{O_2}^{212}, \quad (9)$$

$$(T_{SO_2}^{211}p + 1)\Delta G_{SO_2}^{211} = k_{S_{орг}}^{чщ}\Delta G_{S_{орг}}^{чщ} + k_{O_2}^{215}\Delta G_{O_2}^{215}, \quad (10)$$

$$(T_{SO_2}^{212}p + 1)\Delta G_{SO_2}^{212} = k_{H_2S}^{окз}\Delta G_{H_2S}^{окз} + k_{O_2}^{214}\Delta G_{O_2}^{214} - k_{H_2O}^{211}\Delta G_{H_2O}^{211}. \quad (11)$$

Параметры процесса сгорания углерода, водорода, органической серы и сероводорода сведены в табл. 8.

Таблица 8

Параметры процесса сгорания углерода, водорода органической серы и сероводорода

Обозначение	Наименование параметра	Размерность
$G_{CO_2}^{21}$	Массовый расход углекислого газа в зоне реакции сгорания углерода	кг/с
G_C^{21}	Массовый расход углерода	кг/с
$G_{O_2}^{211}$	Массовый расход кислорода	кг/с
$G_{H_2O}^{212}$	Массовый расход воды в зоне реакции сгорания водорода	кг/с
$G_{H_2}^{чщ}$	Массовый расход водорода	кг/с
$G_{O_2}^{212}$	Массовый расход кислорода	кг/с
$G_{SO_2}^{211}$	Массовый расход оксида серы (сернистого газа)	кг/с
$G_{S_{орг}}^{чщ}$	Массовый расход органической серы	кг/с
$G_{O_2}^{215}$	Массовый расход кислорода	кг/с
$G_{SO_2}^{212}$	Массовый расход оксида серы (сернистого газа)	кг/с
$G_{H_2S}^{окз}$	Массовый расход сероводорода	кг/с
$G_{O_2}^{214}$	Массовый расход кислорода	кг/с
$G_{H_2O}^{211}$	Массовый расход воды	кг/с

Процесс регенерации серы описывается уравнениями [4, 7, 9]:

$$(T_{Na_2SO_4^T}p + 1)\Delta G_{Na_2SO_4^T}^{21} = k_{SO_2}\Delta G_{SO_2}^{221} + k_{O_2}\Delta G_{O_2}^{213} - k_{CO_2}\Delta G_{CO_2}^{22}, \quad (12)$$

где $T_{Na_2SO_4^T} = \frac{x_{Na_2CO_3}}{x_{Na_2SO_4}}\tau_{Na_2CO_3}^{21}$.

Параметры процесса регенерации серы сведены в табл. 9.

Таблица 9

Параметры процесса регенерации серы

Обозначение	Наименование параметра	Размерность
$G_{Na_2SO_4}^{21}$	Массовый расход сульфид натрия в зоне реакции восстановления	кг/с
$G_{SO_2}^{221}$	Массовый расход оксида серы (сернистого газа)	кг/с
$G_{O_2}^{213}$	Массовый расход кислорода	кг/с
$G_{CO_2}^{22}$	Массовый расход углекислого газа	кг/с
$\tau_{Na_2CO_3}^{21}$	Время прохождения капель черного щелока зоны реакции регенерации серы	с

3. Подсистема зеленого и белого щелока СРК (PS6) включает процессы смешения плава с белым щелоком в баке-растворителе плава для получения зеленого щелока требуемых параметров. Модель подсистемы PS6 имеет следующий вид [4, 12, 13]:

$$\frac{d\Delta h}{dt} = K_{11}\Delta Q_{42} - K_{12}\Delta Q_{43} - K_{13}\Delta Q_{44} + K_{17}\Delta G_{41}; \quad (13)$$

$$\frac{d\Delta\rho_{43}}{dt} = -K_{21}\Delta Q_{42} - K_{23}\Delta Q_{44} + K_{24}\Delta\rho_{42} - K_{25}\Delta\rho_{43} - K_{26}\Delta\rho_{44} + K_{27}\Delta G_{41} - -K_{28}\Delta h, \quad (14)$$

где $Q_{42} = \frac{G_{42}}{\rho_{42}}$, $Q_{43} = \frac{G_{43}}{\rho_{43}}$, $Q_{44} = \frac{G_{44}}{\rho_{44}}$, $K_{11} = K_{12} = \frac{1}{F}$, $K_{13} = \frac{\rho_{44}[(1-K_{44})\rho_{432} + K_{44}\rho_{431}]}{\rho_{431}\rho_{432} * F}$,
 $K_{17} = 1/(\rho_{432} * F)$, $K_{21} = (\rho_{43} - \rho_{42})/V$, $K_{23} = \frac{\rho_{44}[-K_{44}\rho_{43}(\rho_{432} - \rho_{431}) + \rho_{432}(\rho_{43} - \rho_{431})]}{\rho_{431}\rho_{432}V}$,
 $K_{24} = \frac{\rho_{43}Q_{42}}{\rho_{42} * V}$, $K_{25} = \frac{Q_{43}}{V}$, $K_{26} = \frac{\rho_{43}\rho_{44}(\rho_{432} - \rho_{431})Q_{44}}{\rho_{431}\rho_{432} * V} \left(\frac{\partial K_{44}}{\partial \rho_{44}} \right)$, $K_{27} = \frac{\rho_{432} - \rho_{43}}{\rho_{432} * V}$, $K_{28} = \frac{Q_{43}}{V} \left(\frac{\partial \rho_{43}}{\partial h} \right)$.

Параметры процесса смешения плава с белым щелоком в табл. 10.

Таблица 10

Параметры процесса смешения плава с белым щелоком

Обозначение	Наименование параметра	Размерность
G_{42}	Массовый расход слабого белого щёлока	кг/с
G_{43}	Массовый расход зелёного щёлока	кг/с
G_{44}	Массовый расход пылегазовых выбросов	кг/с
ρ_{431}	Плотность воды в зелёном щёлоке на выходе из бака-растворителя плава	кг/м ³
K_{42}	Концентрация абсолютно сухих веществ в слабом белом щёлоке	кг а.с.в./ кг раствора
K_{43}	Концентрация абсолютно сухих веществ в зелёном щёлоке	кг а.с.в./кг раствора
K_{44}	Концентрация абсолютно сухих веществ в пылегазовых выбросах	кг а.с.в./кг раствора
G_{41}	Расход абсолютно сухого вещества, поступающего в бак-растворитель плава с потоком плава	кг/с
ρ_{432}	Плотность абсолютно сухого вещества зелёного щёлока	кг/м ³
ρ_{42}	Плотность слабого белого щёлока	кг/м ³
ρ_{43}	Плотность зелёного щёлока	кг/м ³
ρ_{44}	Плотность пылегазовых выбросов	кг/м ³
h	Номинальное значение уровня зелёного щёлока в баке-растворителе плава	м
V	Рабочий объём зелёного щёлока в баке-растворителе плава при номинальном значении уровня	м ³
F	Поверхность зеркала испарения щёлока в баке-растворителе плава	м ²

В данном исследовании приведены модели подсистем СРК для щелокового тракта: подсистемы подготовки черного щелока к сжиганию PS1 → подсистемы сжигания черного щелока PS3 → подсистемы зеленого и белого щелока PS6. Для остальных подсистем СРК разработаны аналогичные модели в виде систем дифференциальных уравнений. Аналитический подход, основанный на законах природы, которым подчиняются процессы в СРК, дает модели подсистем PS1 ... PS6 в символьном виде (ранга $R = 2$). Коэффициенты в виде символов означают, что уравнения справедливы для любого режима.

Для полной определенности моделей (до ранга $R = 3$) необходимы значения коэффициентов дифференциальных уравнений (1)–(14), которые, в свою очередь, зависят от параметров режима. Формулы соответствующих зависимостей приведены ранее.

4. Экспериментальная оценка параметров СРК. Значения параметров ТП зависят от режима работы СРК конкретного ЦБК. Их можно определить только экспериментально – путем обработки доступных архивных данных.

Для оценивания значений параметров ТП выполнена статистическая обработка массива экспериментальных данных, полученных для реального СРК. Для этого из массива выбран один из трех режимов работы (700 тонн а. с. в/сутки), в котором СРК отработал 73% времени календарного года. Результаты обработки массивов данных – оценки параметров ТП подсистем СРК PS1, PS3 и PS6 сведены в табл. 11.

Таблица 11

Значения параметров ТП подсистем СРК

Обозначение	Значение параметра	Размерность	Обозначение	Значение параметра	Размерность
Подсистема подготовки черного щелока к сжиганию (PS1)					
$Q_{вх}$	0,04	м ³ /с	q_s	70	кДж/кг
$\rho_{вх}$	1110	кг/м ³	Q_{21}	0,003	м ³ /с
$i_{вх}$	228	кДж/кг	ρ_2	1400	кг/м ³
$G_{гр}$	6,16	кг/с	ϑ_2	120	°С
$i_{гр}$	2630	кДж/кг	Q_1	0,008	м ³ /с
$q_{гр}$	630	кДж/кг	ρ_1	1350	кг/м ³
$G_{вт}$	5,6	кг/с	ϑ_{11}	105	°С
$i_{вт}^n$	2734	кДж/кг	i_1	250	кДж/кг
$Q_{вс}$	0,013	м ³ /с	V	8	м ³
$\rho_{вс}$	1273	кг/м ³	$G_{п}$	5,6	кг/с
$i_{вс}$	271	кДж/кг	$i_{п}$	2734	кДж/кг
$V_{вс}$	50	м ³	Q_2	0,008	м ³ /с
$\vartheta_{вс}$	115	°С	i_2	330	кДж/кг
G_s	0,93	кг/с	$V_{п}$	4	м ³ /с
ϑ_s	40	°С			
Подсистема сжигания черного щелока в топке (PS3)					
x_{NaOH}	80		x_{Na_2S}	78	
x_{CO_2}	44		x_{H_2S}	34	
$x_{Na_2CO_3}$	106		x_C	24	
x_{H_2O}	18		$x_{Na_2SO_4^*}$	142	
Подсистема зеленого и белого щелока (PS6)					
F	133	м ²	ρ_{44}	0,16	кг/м ³
V	69,3	м ³	K_{44}	0,21	кг а.с.в./кг раствора
ρ_{42}	1040	кг/м ³	G_{42}	11,2	кг/с
ρ_{43}	1160	кг/м ³	G_{43}	13,2	кг/с
ρ_{431}	1000	кг/м ³	G_{44}	0,7	кг/с
ρ_{432}	4500	кг/м ³			

Используя значения параметров ТП подсистем СРК выполнен расчет коэффициентов дифференциальных уравнений по приведенным выше зависимостям (табл. 12).

Таблица 12

Коэффициенты дифференциальных уравнений подсистем СРК

Подсистема подготовки черного щелока к сжиганию (PS1)									
Коэффициент	$T_{\text{вс}}$	A	$K_{Q_{\text{вх}}}$	$K_{\rho_{\text{вх}}}$	$K_{i_{\text{вх}}}$	$K_{G_{\text{гп}}}$	$\Delta G_{\text{вт}}$		
Расчетное значение	3846,2	7,1	45370,4	1,6	6,3	283,8	349,5		
Коэффициент	$K_{Q_{\text{вс}}}$	$T_{\theta_{11}}$	a_1	a_2	a_3	a_4	a_5		
Расчетное значение	89,7	1000	1039,6	0,1	1,5	7	0,1		
Коэффициент	a_6	a_7	a_8	$T_{\theta_2} = T_{G_{22}}$	A_1	B_1	K_1		
Расчетное значение	616,4	0,1	0,3	512,5	56,9	3214,3	5935,5		
Коэффициент	K_2	K_3	K_4	a_{21}	a_{22}	a_{23}	a_{24}		
Расчетное значение	0,5	48,1	8125	0,96	0,01	0,01	0,01		
Подсистема сжигания черного щелока в топке (PS3)									
Коэффициент	$T_{Na_2CO_3}^{11}$	$k_{NaOH}^{\text{чщ}}$	$\tau_{NaOH}^{\text{чщ}}$	$T_{Na_2CO_3}^{12}$	$k_{Na_2S}^{\text{чщ}}$	$\tau_{Na_2S}^{\text{чщ}}$	$T_{Na_2S}^{41}$	$k_{Na_2SO_4}^{32}$	
Расчетное значение	1,1	1,3	1,4	1,0	1,4	1,4	2,5	0,6	
Коэффициент	$T_{CO_2}^{21}$	k_C^{21}	$k_{O_2}^{211}$	$T_{H_2O}^{212}$	$k_{H_2}^{\text{чщ}}$	$k_{O_2}^{212}$	$T_{SO_2}^{211}$	$k_{S_{\text{орг}}}^{\text{чщ}}$	$k_{O_2}^{215}$
Расчетное значение	0,27	2,0	1,38	0,89	1,89	2,13	0,5	2	1,5
Коэффициент	$T_{SO_2}^{212}$	$k_{H_2S}^{\text{окз}}$	$k_{O_2}^{214}$	$k_{H_2O}^{211}$	$T_{Na_2SO_4}^{\text{т}}$	k_{SO_2}	k_{O_2}	k_{CO_2}	$\tau_{Na_2CO_3}^{21}$
Расчетное значение	0,53	1,88	3,67	3,67	1,02	1	2	3	1,4
Подсистема зеленого и белого щелока (PS6)									
Коэффициент	$K_{11} = K_{12}$		K_{13}	K_{17}		K_{21}		K_{23}	K_{24}
Расчетное значение	0,01		0,01	0,01		1,73		0,01	0,01
Коэффициент	K_{25}		K_{26}	K_{27}		K_{28}			
Расчетное значение	0,01		0,01	0,01		0,09			

Заключение. Разработана многоуровневая линейная стационарная модель содоре-генерационного котлоагрегата, которая позволяет анализировать динамику подсистем и агрегата в целом – устойчивость режимов и влияние возмущений, обосновать необходимость автоматического поддержания требуемого качества процессов в окрестности выбранного режима.

Модели в символьном виде применимы к большинству подобных агрегатов технологического процесса производства сульфатной целлюлозы. Параметры полностью определенных моделей должны быть скорректированы по результатам обработки данных, относящихся к конкретному агрегату и режиму.

Направление дальнейших исследований связано с развитием методов анализа динамики сложных систем управления по моделям иерархической структуры, а также синтеза систем стабилизации режима и ослабления влияния возмущений.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. *Липовков И.З.* Сжигание сульфатного щёлока. – М.: Лесная промышленность, 1979. – 128 с.
2. *Непенин Ю.Н.* Технология целлюлозы: В 3 т. Т. 2. Производство сульфатной целлюлозы. – М.: Лесн. пром-сть, 1990. – 600 с.
3. Технология целлюлозно-бумажного производства. В 3 т. Т. 3. Автоматизация, стандартизация, экономика и охрана окружающей среды в ЦБП. Ч. 1. Автоматизация, стандартизация и экономика в ЦБП. – СПб.: Политехника, 2008. – 621 с.
4. *Доронин В.А.* Автоматизация содорегенерационных агрегатов. – М.: Лесная промышленность, 1988. – 320 с.
5. *Сиваков В.П., Вураско А.В., Музыкантова В.И.* Регенерация химикатов и тепла в содорегенерационных котлоагрегатах. Устройство и диагностирование. – Екатеринбург: УГЛТУ, 2015. – 140 с.
6. *Новикова А.И.* Технологические расчеты материального и теплового балансов содорегенерационного отдела завода сульфатной целлюлозы: учеб. пособие. – Ленинград: ЛТА, 1983. – 78 с.
7. *Житков В.В.* Исследование процесса сжигания черного щелока в топках содорегенерационных агрегатов впрыскивающего типа: дисс. ... канд. техн. наук. ВНИИБ. – Ленинград, 1973. – 230 с.
8. *Вьюков Б.Е.* Исследование процесса сжигания черных сульфатных щелоков для целей управления: дисс. ... канд. техн. наук: 05.21.03. Ленинградская лесотехническая академия. – Ленинград., 1974. – 142 с.
9. *Сморodin С.Н.* Исследование топочных процессов при сжигании сульфатного щелока в вихревой топке содорегенерационного котлоагрегата с целью повышения ее эффективности: специальность 05.14.04 "Промышленная теплоэнергетика": дисс. ... канд. техн. наук. – СПб., 2000. – 112 с. – EDN QDIMMN.
10. *Пожитков В.В.* Математическое моделирование процессов загрязнения поверхностей нагрева содорегенерационного агрегата сульфат-целлюлозного производства с целью оптимизации их работы: дисс. ... канд. техн. наук: 05.21.03. Ленинградский технологический институт целлюлозно-бумажной промышленности. – Ленинград, 1989. – 217 с.
11. *Абдалла Исхаг Абдалла.* Разработка модели и системы управления производительностью процесса выпаривания щелоков сульфат-целлюлозного производства: специальность 05.21.03 "Технология и оборудование химической переработки биомассы дерева; химия древесины": дисс. ... канд. техн. наук. – СПб., 2003. – 145 с. – EDN NMNLEL.
12. *Братищева А.В.* Очистка выбросов растворителя плава содорегенерационного котлоагрегата сульфатцеллюлозного производства в конденсаторе наклонного типа: специальность 05.21.03 "Технология и оборудование химической переработки биомассы дерева; химия древесины": дисс. ... канд. техн. наук. – СПб., 2007. – 158 с. – EDN NOPQSD.
13. *Горьковой Е.В.* Диагностика состояния содорегенерационного котлоагрегата с целью повышения эффективности его работы: специальность 05.21.03 "Технология и оборудование химической переработки биомассы дерева; химия древесины", 05.13.06 "Автоматизация и управление технологическими процессами и производствами (по отраслям)": дисс. ... канд. техн. наук. – СПб., 2012. – 159 с. – EDN QFUFDF.
14. Black Liquor Recovery Boiler Advisory Committee (BLRBAC). Recommended guidelines for personnel safety for black liquor recovery boilers. – Atlanta: BLRBAC, 2018. – 46 p. – URL: <https://blrbac.net/wp-content/uploads/2020/10/personnel-safety-april-2018.pdf> (дата обращения: 30.04.2025).
15. *Rämö J., Lappalainen J., Hiltunen Y., Järvinen M., Alopaeus V.* Solids content of black liquor measured by online time-domain NMR // Applied Sciences. – 2019. – Vol. 9, Issue 10. – Article 2169. – DOI: 10.3390/app9102169.
16. *Vakkilainen E.K., Hamaguchi M.* Evaluation of the energy efficiency improvement potential through back-end heat recovery in the Kraft recovery boiler // Energies. – 2021. – Vol. 14, Issue 6. – Article 1550. – DOI: 10.3390/en14061550.
17. *Paoliello F.* Corrosion damage and in-service inspection of retractable sootblower lances in recovery boilers // TAPPI Journal. – 2021. – Vol. 20, Issue 10. – P. 655-662. – DOI: 10.32964/tj20.10.655.
18. *Raju V., Engblom M., Rantala E., Enestam S., Mansikkasalo J.* Kraft recovery boiler operation with splash plate and/or beer can nozzles - a case study // TAPPI Journal. – 2021. – Vol. 20, Issue 10. – P. 625-636. – DOI: 10.32964/TJ20.10.625.
19. *Saari J., Kaikko J., Sermyagina E., Hamaguchi M., Cardoso M., Vakkilainen E., Haider M.* Recovery boiler back-end heat recovery // TAPPI Journal. – 2023. – Vol. 22, Issue 3. – P. 174-183. – DOI: 10.32964/tj22.3.174.
20. *Belisario A.B., Edberg A., Björk M., Almeida G.M., Vakkilainen E.K.* On the diagnosis of a fouling condition in a kraft recovery boiler: combining process knowledge and data-based insights // TAPPI Journal. – 2023. – Vol. 22, Issue 3. – P. 162-171. – DOI: 10.32964/tj22.3.162

21. *Rajan K., Berton P., Rogers R.D., Shamshina J.L.* Is Kraft pulping the future of biorefineries? A perspective on the sustainability of lignocellulosic product development // *Polymers*. – 2024. – Vol. 16, Issue 23. – Article 3438. – DOI: 10.3390/polym16233438
22. *Niemi J., Balint R., Engblom M., Lindberg D.* Modeling of temperature gradient-induced melt movement within kraft recovery boiler ash deposits // *Energy & Fuels*. – 2024. – Vol. 39, Issue 1. – P. 454-464. – DOI: 10.1021/acs.energyfuels.4c04516.
23. *Niemi J., Balint R., Engblom M., Lehmusto J., Lindberg D., Hupa L.* Equilibrium model approach to predict local chemical changes in recovery boiler deposits // *Energy*. – 2024. – Vol. 306. – Article 132507. – DOI: 10.1016/j.energy.2024.132507.
24. *Kayal B., Nasr Y., El Zakhem H., El Bachawati M.* Optimizing brown stock washing in the pulp and paper industry: a system dynamics approach // *Processes*. – 2025. – Vol. 13, Issue 2. – Article 368. – DOI: 10.3390/pr13020368.
25. *Вавилов А.А., Имаев Д.Х., Плескунин В.И., Фомин Б.Ф., Зальтцер Г., Зауер В., Станек В., Франк М., Фридрих В., Шаллер Я., Штремпель Х.* Имитационное моделирование производственных систем / под общ. ред. А.А. Вавилова. – М.: Машиностроение, 1983. – 416 с.
26. *Вавилов А.А., Имаев Д.Х.* Эволюционный синтез систем управления: учеб. пособие. – Л.: ЛЭТИ, 1983. – 80 с.
27. *Имаев Д.Х., Станиславски В.* Исследование динамики иерархических систем на примере парогенераторов. – СПб.: Санкт-Петербургский государственный электротехнический университет «ЛЭТИ» им. В.И. Ульянова (Ленина), 2011. – 219 с. – ISBN 978-5-7629-1136-8. – EDN SCTSTX.

REFERENCES

1. *Lipovkov I.Z.* Szhiganie sul'fatnogo shcheloka [Combustion of sulfate liquor]. Moscow: Lesnaya promyshlennost', 1979, 128 p.
2. *Nepenin Yu.N.* Tekhnologiya tsellyulozy: V 3 t. T. 2. Proizvodstvo sul'fatnoy tsellyulozy [Cellulose Technology: In 3 vol. Vol. 2. Production of sulfate cellulose]. Moscow: Lesn. prom-st', 1990, 600 p.
3. Tekhnologiya tsellyulozno-bumazhnogo proizvodstva. V 3 t. T. 3. Avtomatizatsiya, standartizatsiya, ekonomika i okhrana okruzhayushchey sredy v TSBP. Ch. 1. Avtomatizatsiya, standartizatsiya i ekonomika v TSBP [Pulp and paper production technology. In 3 volumes. Volume 3. Automation, standardization, economics and environmental protection in the pulp and paper industry. Part 1. Automation, standardization and economics in the pulp and paper industry]. Saint Petersburg: Politehnika, 2008, 621 p.
4. *Doronin V.A.* Avtomatizatsiya sodoregeneratsionnykh agregatov [Automation of sodium recovery units]. Moscow: Lesnaya promyshlennost', 1988, 320 p.
5. *Sivakov V.P., Vurasko A.V., Muzykantova V.I.* Regeneratsiya khimikatov i tepla v sodoregeneratsionnykh kotloagregatakh. Ustroystvo i diagnostirovanie [Chemical and heat recovery in sodium recovery boiler units. Design and diagnostics]. Ekaterinburg: UGLTU, 2015, 140 p.
6. *Novikova A.I.* Tekhnologicheskie raschety material'nogo i teplovogo balansov sodoregeneratsionnogo otdela zavoda sul'fatnoy tsellyulozy: ucheb. posobie [Technological calculations of material and heat balances for the sodium recovery department of a sulphate pulp plant: a textbook]. Leningrad: LTA, 1983, 78 p.
7. *Zhitkov V.V.* Issledovanie protsessa szhiganiya chernogo shcheloka v topkakh sodoregeneratsionnykh agregatov vpryskivayushchego tipa: diss. ... kand. tekhn. nauk [Study of the black liquor combustion process in injection-type sodium recovery unit furnaces: cand. of eng. sc. diss.] VNIIB. Leningrad, 1973, 230 p.
8. *V'yukov B.E.* Issledovanie protsessa szhiganiya chernykh sul'fatnykh shchelokov dlya tseley upravleniya: diss. ... kand. tekhn. nauk [Study of the process of combustion of black sulfate liquors for control purposes: cand. of eng. sc. diss.]: 05.21.03. Leningrad Forest Engineering Academy. Leningrad, 1974, 142 p.
9. *Smorodin S.N.* Issledovanie topochnykh protsessov pri szhiganii sul'fatnogo shcheloka v vikhrevooy topke sodoregeneratsionnogo kotloagregata s tsel'yu povysheniya ee effektivnosti: spetsial'nost' 05.14.04 "Promyshlennaya teploenergetika": diss. ... kand. tekhn. nauk [Study of combustion processes during combustion of sulfate liquor in a vortex furnace of a soda recovery boiler unit in order to improve its efficiency: specialty 05.14.04 "Industrial heat power engineering": cand. of eng. sc. diss.]. Saint Petersburg, 2000, 112 p. EDN QDIMN.
10. *Pozhitkov V.V.* Matematicheskoe modelirovanie protsessov zagryazneniya poverkhnostey nagreva sodoregeneratsionnogo agregata sul'fat-tsellyuloznogo proizvodstva s tsel'yu optimizatsii ikh raboty: diss. ... kand. tekhn. nauk [Mathematical modeling of the processes of contamination of heating surfaces of a soda recovery unit in sulfate pulp production in order to optimize their operation: cand. of eng. sc. diss.]: 05.21.03. Leningrad Technological Institute of the Pulp and Paper Industry. Leningrad, 1989, 217 p.

11. Abdalla Iskhag Abdalla. Razrabotka modeli i sistemy upravleniya proizvoditel'nost'yu protsessa vyparivaniya shchelokov sul'fat-sellyuloznogo proizvodstva: spetsial'nost' 05.21.03 "Tekhnologiya i oborudovanie khimicheskoy pererabotki biomassy dereva; khimiya drevesiny": diss. ... kand. tekhn. nauk [Development of a model and performance control system for the evaporation process of sulphate pulp production liquors: specialty 05.21.03 "Technology and equipment for chemical processing of wood biomass; wood chemistry": cand. of eng. sc. diss.]. Saint Petersburg, 2003, 145 p. EDN NMNLEL.
12. Brattseva A.V. Ochistka vybrosov rastvoritelya plava sodoregeneratsionnogo kotloagregata sul'fatsellyuloznogo proizvodstva v kondensatore naklonnogo tipa: spetsial'nost' 05.21.03 "Tekhnologiya i oborudovanie khimicheskoy pererabotki biomassy dereva; khimiya drevesiny": diss. ... kand. tekhn. nauk [Cleaning of solvent emissions from the melt of a soda recovery boiler unit of a sulphate cellulose plant in an inclined-type condenser: specialty 05.21.03 "Technology and equipment for chemical processing of wood biomass; wood chemistry": cand. of eng. sc. diss.]. Saint Petersburg, 2007, 158 p. EDN NOPQSD.
13. Gor'kovoy E.V. Diagnostika sostoyaniya sodoregeneratsionnogo kotloagregata s tsel'yu povysheniya effektivnosti ego raboty: spetsial'nost' 05.21.03 "Tekhnologiya i oborudovanie khimicheskoy pererabotki biomassy dereva; khimiya drevesiny", 05.13.06 "Avtomatizatsiya i upravlenie tekhnologicheskimi protsessami i proizvodstvami (po otraslyam)": diss. ... kand. tekhn. nauk [Diagnostics of the state of a soda recovery boiler unit in order to improve its operating efficiency: specialty 05.21.03 "Technology and equipment for chemical processing of wood biomass; wood chemistry", 05.13.06 "Automation and control of technological processes and production (by industry)": cand. of eng. sc. diss.]. Saint Petersburg, 2012, 159 p. EDN QFUFDF.
14. Black Liquor Recovery Boiler Advisory Committee (BLRBAC). Recommended guidelines for personnel safety for black liquor recovery boilers. Atlanta: BLRBAC, 2018, 46 p. Available at: <https://blrbac.net/wp-content/uploads/2020/10/personnel-safety-april-2018.pdf> (accessed 30 April 2025).
15. Rämö J., Lappalainen J., Hiltunen Y., Järvinen M., Alopaeus V. Solids content of black liquor measured by online time-domain NMR, *Applied Sciences*, 2019, Vol. 9, Issue 10, Article 2169. DOI: 10.3390/app9102169.
16. Vakkilainen E.K., Hamaguchi M. Evaluation of the energy efficiency improvement potential through back-end heat recovery in the Kraft recovery boiler, *Energies*, 2021, Vol. 14, Issue 6, Article 1550. DOI: 10.3390/en14061550.
17. Paoliello F. Corrosion damage and in-service inspection of retractable sootblower lances in recovery boilers, *TAPPI Journal*, 2021, Vol. 20, Issue 10, pp. 655-662. DOI: 10.32964/tj20.10.655.
18. Raju V., Engblom M., Rantala E., Enestam S., Mansikkasalo J. Kraft recovery boiler operation with splash plate and/or beer can nozzles - a case study, *TAPPI Journal*, 2021, Vol. 20, Issue 10, pp. 625-636. – DOI: 10.32964/TJ20.10.625.
19. Saari J., Kaikko J., Sermyagina E., Hamaguchi M., Cardoso M., Vakkilainen E., Haider M. Recovery boiler back-end heat recovery, *TAPPI Journal*, 2023, Vol. 22, Issue 3, pp. 174-183. DOI: 10.32964/tj22.3.174.
20. Belisario A.B., Edberg A., Björk M., Almeida G.M., Vakkilainen E.K. On the diagnosis of a fouling condition in a kraft recovery boiler: combining process knowledge and data-based insights, *TAPPI Journal*, 2023, Vol. 22, Issue 3, pp. 162-171. DOI: 10.32964/tj22.3.162
21. Rajan K., Berton P., Rogers R.D., Shamshina J.L. Is Kraft pulping the future of biorefineries? A perspective on the sustainability of lignocellulosic product development, *Polymers*, 2024, Vol. 16, Issue 23, Article 3438. DOI: 10.3390/polym16233438
22. Niemi J., Balint R., Engblom M., Lindberg D. Modeling of temperature gradient-induced melt movement within kraft recovery boiler ash deposits, *Energy & Fuels*, 2024, Vol. 39, Issue 1, pp. 454-464. DOI: 10.1021/acs.energyfuels.4c04516.
23. Niemi J., Balint R., Engblom M., Lehmusto J., Lindberg D., Hupa L. Equilibrium model approach to predict local chemical changes in recovery boiler deposits, *Energy*, 2024, Vol. 306, Article 132507. DOI: 10.1016/j.energy.2024.132507.
24. Kayal B., Nasr Y., El Zakhem H., El Bachawati M. Optimizing brown stock washing in the pulp and paper industry: a system dynamics approach, *Processes*, 2025, Vol. 13, Issue 2, Article 368. DOI: 10.3390/pr13020368.
25. Vavilov A.A., Imaev D.X., Pleskunin V.I., Fomin B.F., Zal'tser G., Zauer V., Stanek V., Frank M., Fridrikh V., Shaller YA., Shtrempel' X. Imitatsionnoe modelirovanie proizvodstvennykh sistem [Simulation modeling of production systems], ed. by A.A. Vavilova. Moscow: Mashinostroyeniye, 1983, 416 p.
26. Vavilov A.A., Imaev D.Kh. Evolyutsionnyy sintez sistem upravleniya: ucheb. posobie [Evolutionary synthesis of control systems: a tutorial]. Leningrad: LETI, 1983, 80 p.
27. Imaev D.Kh., Stanislavski V. Issledovanie dinamiki ierarkhicheskikh sistem na primere parogeneratorov [Study of the Dynamics of Hierarchical Systems Using Steam Generators as an Example]. Saint Petersburg: Sankt-Peterburgskiy gosudarstvennyy elektrotekhnicheskii universitet «LETI» im. V.I. Ul'yanova (Lenina), 2011, 219 p. ISBN 978-5-7629-1136-8. EDN SCTSTX.

Ковалёв Дмитрий Александрович – Санкт-Петербургский государственный университет промышленных технологий и дизайна; e-mail: d.a.kovalyov@yandex.ru; г. Санкт-Петербург, Россия; тел.: +78127865366; к.т.н.; доцент; зав. кафедрой автоматизации технологических процессов и производств.

Имаев Дамир Хабибович – Санкт-Петербургский государственный электротехнический университет «ЛЭТИ» им. В.И. Ульянова (Ленина); e-mail: damir.imaev@mail.ru; г. Санкт-Петербург, Россия; тел.: +78122343798; д.т.н.; профессор кафедры автоматики и процессов управления.

Душин Сергей Евгеньевич – Санкт-Петербургский государственный электротехнический университет «ЛЭТИ» им. В.И. Ульянова (Ленина); e-mail: dushins@yandex.ru; г. Санкт-Петербург, Россия; тел.: +78122343798; д.т.н.; профессор кафедры автоматики и процессов управления.

Kovalev Dmitrii Aleksandrovich – Saint Petersburg State University of Industrial Technologies and Design; e-mail: d.a.kovalyov@yandex.ru; Saint Petersburg Russia; phone: +78127865366; cand. of eng. sc.; associate professor; head of the Department of Automation of Technological Processes and Production.

Imaev Damir Habibovich – Saint Petersburg Electrotechnical University "LETI" named after V.I. Ulyanov (Lenin); e-mail: damir.imaev@mail.ru; Saint Petersburg Russia; phone: +78122343798; dr. of eng. sc.; professor of the Department of Automation and Control Processes.

Dushin Sergey Evgenievich – Saint Petersburg Electrotechnical University "LETI" named after V.I. Ulyanov (Lenin); e-mail: dushins@yandex.ru; Saint Petersburg, Russia; phone: +78122343798; dr. of eng. sc.; professor of the Department of Automation and Control Processes.

УДК 681.5

DOI 10.18522/2311-3103-2026-4-241-249

А.В. Мартиросян, Д.В. Ромашин

КОНЦЕПТУАЛЬНАЯ МОДЕЛЬ ФАКТОРОВ, ВЛИЯЮЩИХ НА ЭФФЕКТИВНОСТЬ ПРОЦЕССА ПОДГОТОВКИ И СЕПАРАЦИИ ГАЗА

Представлена концепция адаптивного управления при разделении природного газа. Применяемые на практике системы управления, как правило, основаны на фиксированных алгоритмах регулирования и не учитывают динамическое изменение физико-технических и эксплуатационных параметров, что приводит к снижению качества сепарации и росту энергозатрат. В статье описан подход сочетания системного анализа и метод Парето. Целью работы является разработка концепции адаптивного управления процессом подготовки природного газа на основе системного анализа. Недавние исследования показывают, что методы, основанные на данных, позволяют более точно настраивать параметры, лучше реагировать на колебания исходного газа и повышать эффективность удаления примесей. Для достижения поставленной цели в работе решены задачи выявления и классификации факторов, влияющих на качество и экономичность сепарации газа, а также формирования их взаимосвязей в рамках единой концептуальной модели управления. Особое внимание уделяется влиянию точного контроля ключевых параметров, таких как давление, температура и расход, на эффективность этих процессов. Анализ последних исследований демонстрирует растущее использование нейронных сетей и моделей машинного обучения в очистке газа для прогнозирующего управления, обнаружения аномалий и оптимизации рабочих параметров. Сравнительная оценка классических ПИД-регуляторов, нечетких, адаптивных и нейронных методов управления подтверждает преимущества интеллектуального регулирования с точки зрения стабильности, адаптивности и энергоэффективности. Основным результатом работы является обоснование ключевых факторов, определяющих эффективность сепарации, среди которых наибольшее влияние оказывают давление, температура и расход газа. Полученная модель формирует методологическую основу для разработки интеллектуальных и адаптивных систем управления процессом подготовки газа.

Сепарация газа; очистка газов; адаптивное управление; нечеткая логика; нейронные сети; машинное обучение; оптимизация процессов; интеллектуальные системы; метод Парето.

A.V. Martirosyan, D.V. Romashin

CONCEPTUAL MODEL OF FACTORS INFLUENCING THE EFFICIENCY OF GAS PREPARATION AND SEPARATION PROCESS

The paper presents the concept of adaptive control in natural gas separation. The control systems used in practice are usually based on fixed control algorithms and do not consider dynamic changes in physical, technical and operational parameters, which leads to a decrease in separation quality and an increase in energy consumption. The article describes the approach of combining system analysis, modeling and the Pareto method. This paper presents a concept for adaptive control in natural gas separation. Control systems used in practice are typically based on fixed control algorithms and do not account for dynamic changes in physical, technical, and operational parameters, which leads to reduced separation quality and increased energy costs. This paper describes an approach combining systems analysis and the Pareto method. The aim of this paper is to develop a concept for adaptive control of the natural gas treatment process based on systems analysis. Recent studies demonstrate that data-driven methods enable more precise parameter adjustment, better responsiveness to raw-gas fluctuations and improved impurity removal efficiency. To achieve this goal, this paper addresses the challenges of identifying and classifying factors affecting the quality and efficiency of gas separation, as well as integrating their relationships within a unified conceptual control model. Particular attention is paid to the impact of precise control of key parameters, such as pressure, temperature, and flow rate, on the efficiency of these processes. An analysis of recent research demonstrates the growing use of neural networks and machine learning models in gas purification for predictive control, anomaly detection, and optimization of operating parameters. A comparative evaluation of classic PID controllers, fuzzy, adaptive, and neural control methods confirms the advantages of intelligent control in terms of stability, adaptability, and energy efficiency. The main result of this study is the substantiation of key factors determining separation efficiency, among which pressure, temperature, and gas flow rate have the greatest impact. The resulting model forms a methodological basis for the development of intelligent and adaptive control systems for gas purification processes.

Gas separation; gas treatment; adaptive control; fuzzy logic; neural networks; machine learning; process optimization; intelligent systems; Pareto method.

Введение. Нефтегазовая отрасль играет ключевую роль в мировой экономике, обеспечивая более половины потребляемой энергии. Природный газ, как самое чистое ископаемое топливо, является важным ресурсом для перехода к новым энергетическим технологиям. В условиях усложняющейся добычи и растущих требований к качеству газа необходимо совершенствовать методы его очистки и разделения [1, 2].

На газоперерабатывающих предприятиях процесс сепарации природного газа остаётся одним из наиболее энергоёмких и критичных к качеству управления этапов. Существующие системы автоматики, как правило, используют жёсткие алгоритмы регулирования, не адаптирующиеся к изменению свойств сырья и внешних условий. Это приводит к росту содержания примесей в товарном газе, увеличению эксплуатационных затрат и сокращению ресурса оборудования. В то же время современные методы интеллектуального управления, способные решить указанные проблемы, требуют предварительного системного анализа факторов, влияющих на процесс. Настоящая работа посвящена созданию такой аналитической основы.

В качестве примера предлагается рассмотреть процесс сепарации газа. Сепарация – процесс разделения (отделения, разъединения) твёрдой, жидкой и газовой (паровой) фаз потока природного газа с последующим извлечением из него твёрдой и жидкой фаз. Сепарация газа предназначена для предохранения от попадания влаги и твёрдых частиц в промысловые газосборные сети и технологическое оборудование газовых и газоконденсатных месторождений.

Добываемый газ содержит механические (частицы горных пород, строительный шлак, продукты коррозии и эрозии, жидкий конденсат и воду) и химические (сероводород, сероуглерод, меркаптан, азот, гелий, углекислый газ, сульфиды и дисульфиды) примеси. Эти примеси могут оказывать негативное воздействие на оборудование, приводя к увеличению случаев отказов, аварий и засоров твердыми частицами. Таким образом, повышенное количество примесей в потоке способно ускорить износ и частоту обслуживания оборудования, увеличивая эксплуатационные расходы.

Несмотря на непрерывную модернизацию технологического комплекса, логика управления многих установок остается неизменной. Сейчас на рассматриваемом объекте алгоритм управления не может адаптироваться к изменениям состава газа и колебаниям температуры окружающей среды. В последствии это приводит к снижению эффективности сепарации, увеличению содержания примесей на выходе и повышению энергопотребления для повторной очистки газа [3, 4]. Решить данную задачу на газоперерабатывающих заводах позволяет автоматизированные системы управления, которые непрерывно измеряют, анализируют и корректируют параметры в режиме непрерывного мониторинга. Меры управления направлены на стабилизацию технологического потока и улучшение очистки газа от жидких и механических примесей.

Для эффективного регулирования давления используют датчики давления передают сигналы контроллеру, который сравнивает текущее значение с заданным. При повышении давления контроллер посылает управляющий сигнал на частичное открытие клапана, тем самым увеличивая расход и снижая внутреннее давление установки. И наоборот, при снижении давления клапан закрывается. Этот принцип управления основан на пропорционально-интегрально-дифференциальном (ПИД) законе управления, обеспечивающем плавное и стабильное управление параметрами и предотвращающем колебания. Регулирование температуры направлено на обеспечение условий, необходимых для эффективного разделения газожидкостной смеси. Мониторинг осуществляется с помощью терморезисторных датчиков или термопар, установленных в критических зонах оборудования. Эффект регулирования достигается путем изменения расхода хладагента или охлаждающей жидкости через регулирующие клапаны. При повышении температуры подача охлаждающей жидкости увеличивается; при понижении температуры подача охлаждающей жидкости уменьшается.

Для эффективного регулирования уровня жидкости в сепараторе используют уровнемеры поплавкового, которые позволяют предотвратить переполнения и прорыва жидкости в газовую линию. При достижении предельно допустимого уровня контроллер подаёт сигнал на открытие дренажного клапана или включение насоса для отвода конденсата в сборную ёмкость. После снижения уровня клапан автоматически закрывается. Для контроля расхода газа используют вихревые или диафрагменные расходомеры, которые позволяют равномерно осуществлять поступление газового потока в сепаратор.

Все управляющие воздействия выполняются в замкнутом цикле автоматического регулирования. Данный процесс протекает измерительные устройства фиксируют текущее состояние процесса, контроллер анализирует отклонение параметров от заданных уставок и формирует управляющий сигнал для исполнительных механизмов. Исполнительные устройства (клапаны, насосы, дросселирующие узлы) изменяют состояние технологического потока, после чего обновлённые значения снова поступают в систему обратной связи. Для газоперерабатывающих предприятий важно адаптировать управление параметрами, такими как давление, температура и расход газа. способы регулирования обеспечивают стабильность, но не позволяют менять параметры при изменении сырья или внешних факторов (состав сырья, температура окружающей среды) [5, 6]. Чтобы решить эту задачу, нужен системный подход, объединяющий анализ связей между параметрами, прогнозирование и оценку факторов [7]. Современные исследования демонстрируют, что применение интеллектуальных технологий и системного анализа позволяет рассматривать процесс подготовки газа как задачу с множеством факторов. Эффективность этого процесса зависит от различных технических, термодинамических и эксплуатационных аспектов [8, 9]. Для своего исследования мы использовали метод Парето. Этот метод помогает определить приоритеты среди параметров по степени их влияния на конечный результат и выделить ключевые факторы, определяющие качество подготовки газа [7, 8]. Применение такого подхода позволяет перейти от управления отдельными параметрами к комплексной системе регулирования, основанной на приоритетах и обратной связи [9, 10].

Существующие системы автоматического регулирования на газоперерабатывающих предприятиях построены по принципу поддержания заданных уставок с помощью ПИД-законов. Однако, как показал анализ эксплуатационных данных, такие системы де-

монстрируют снижение эффективности при изменении состава сырья, колебаниях температуры окружающей среды и нестабильности подачи газа. Это приводит к росту содержания примесей на выходе и увеличивает энергопотребление на повторную очистку.

Альтернативой является применение адаптивных и интеллектуальных методов управления, которые способны перестраивать стратегию регулирования в зависимости от текущих условий. В литературе активно обсуждаются подходы на основе нечёткой логики, нейронных сетей и гибридных архитектур, однако их практическое внедрение сдерживается отсутствием системных моделей, связывающих технологические параметры с показателями эффективности. Данное исследование направлено на формирование такой методологической основы.

Цель работы – создать концептуальную модель адаптивного управления процессом подготовки газа.

Для этого нужно решить следующие задачи:

1. Изучить современные подходы к управлению процессом подготовки газа и выявить их ограничивающие факторы, влияющие на качество сепарации [11].
2. Определить, какие факторы влияют на качество и эффективность сепарации газа [12].
3. Разработать концептуальную модель управления, показывающую, как связаны параметры технологического процесса.
4. Оценить приоритетность факторов методом Парето, используя статистические данные и мнения экспертов.

Объект исследования – процесс сепарации природного газа. Предмет исследования – методы системного анализа и оценки факторов, влияющих на управление параметрами подготовки газа [13].

Концептуальная модель позволяет определить взаимосвязь между параметрами (давление, температура, расход газа) и внешними воздействиями. Это является основой для создания адаптивных систем регулирования [14, 15].

Практическая польза заключается в том, что разработанная модель позволит упростить параметры процесса, выявить наиболее важные факторы, влияющие на качество сепарации [16]. Это позволит определить дальнейшие задачи для совершенствования существующих систем управления. В перспективе данное решение можно интегрировать в состав комплексных интеллектуальных систем [17].

Материалы и методы. Методологическая часть исследования строится на применении системного анализа в процессе подготовки природного газа. Этот метод рассматривает объекты как совокупность взаимосвязанных элементов: сырье, оборудование, параметры управления, внешние факторы и показатели эффективности.

На данном этапе работы основное внимание направлено на создание концептуальной модели, которая будет служить теоретическим основанием для разработки математической модели и алгоритмов адаптивного управления. Цель модели – интегрировать физико-технические, организационные и информационные аспекты, влияющие на эффективность подготовки газа к сепарации [18].

Исследование основано на сочетании методов системного анализа, структурного моделирования и оценки приоритетов факторов по принципу Парето. Такой подход позволяет не только выявить ключевые параметры, определяющие качество технологического процесса, но и обосновать их взаимосвязь в контуре управления [19].

Согласно системной методологии, процесс подготовки газа рассматривается как многоуровневая динамическая система, в которой выделяются:

- ♦ входные параметры – давление, температура, состав сырья, расход газа и жидкости;
- ♦ управляемые переменные – давление в сепараторе, расход газа, температура теплообменников [20];
- ♦ возмущающие воздействия – изменения состава сырья, внешняя температура, нестабильность подачи;
- ♦ результирующие показатели (KPI) – степень очистки, потери конденсата, энергоёмкость и стабильность режима.

Взаимосвязи параметров анализировались с использованием элементов системного структурирования: выделения ключевых факторов, построения их иерархии и определения направлений влияния [21].

Далее, будут описаны этапы разработки концептуальной модели выполнялась в несколько последовательных этапов:

1. Сбор и анализ исходной информации. Были изучены технические характеристики сепараторов, данные технологических измерений (давление, температура, расход газа) и результаты промысловых наблюдений [22]. Дополнительно рассматривались исследования отечественных и зарубежных авторов по управлению процессами подготовки газа [23, 24].

2. Систематизация факторов. На основе анализа выделена совокупность факторов, влияющих на эффективность сепарации, сгруппированных по блокам: физико-технические параметры (температура, давление, расход газа); конструктивные и эксплуатационные особенности оборудования; характеристики сырья (состав, влажность, содержание примесей); организационно-технологические условия (режим работы, частота регулирования, квалификация персонала); внешние воздействия (температура окружающей среды, изменения дебита).

В дальнейшем была составлена структурированная система факторов, влияющих на качественную сепарацию газа – то есть на минимизацию количества примесей (жидких, механических, химических). Данные факторы будут использоваться для экспертного опроса и оценки значимости [25].

Наиболее значимые параметры определены по диаграмме Парето (рис. 1). Согласно принципу Парето, наиболее значимые параметры имеют вес $\lambda_i \geq 0,01$. Выбор метода Парето для ранжирования факторов обусловлен тем, что технологический процесс сепарации характеризуется большим числом взаимосвязанных параметров, многие из которых дублируют влияние друг друга. Применение метода Парето позволяет выделить ограниченное множество наиболее значимых факторов, определяющих до 80% суммарного эффекта, что критически важно для последующего построения управляющих моделей – сокращение размерности задачи снижает вычислительную сложность и повышает робастность системы. Кроме того, метод Парето хорошо сочетается с экспертными оценками и статистическими данными, что делает его применимым на начальных этапах исследования при отсутствии больших объемов промышленных данных.



Рис. 1. Диаграмма Парето

Согласно данному соотношению построена концептуальная модель подготовки газа к сепарации (рис. 2).

На основе разработанной концептуальной модели на следующем этапе предполагается построение адаптивной системы управления, включающей следующие функциональные блоки: Блок идентификации состояния – предназначен для оценки текущих значений давления, температуры, расхода и состава сырья на основе данных датчиков и корреляционных моделей. Блок прогнозирования – планируется использование рекуррентных нейронных сетей (LSTM) или уравнений состояния для краткосрочного прогноза изменения выходных показателей при текущих управляющих воздействиях. Блок оптимизации уставок – предполагается решение многокритериальной задачи (по Парето) с целевыми функциями:

минимизация содержания примесей, минимизация энергозатрат, максимизация стабильности режима. Блок нечёткого управления – на основе базы правил вида «ЕСЛИ ... ТО ...» предполагается коррекция уставок ПИД-регуляторов или прямое управление исполнительными механизмами. Детальная реализация перечисленных блоков, включая выбор конкретных алгоритмов и их настройку, является предметом дальнейших исследований.

Математический аппарат для оценки качества прогнозирования. При проведении исследования для количественной оценки точности прогнозирования параметров сепарации использовались следующие метрики, а именно MAE, RMSE, MAPE.

MAE (Mean Absolute Error) – средняя абсолютная ошибка:

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i|. \quad (1)$$

RMSE (Root Mean Square Error) – среднеквадратичная ошибка:

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2}. \quad (2)$$

R^2 – коэффициент детерминации, показывающий долю объяснённой дисперсии:

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2}. \quad (3)$$

MAPE (Mean Absolute Percentage Error) – средняя абсолютная процентная ошибка:

$$\text{MAPE} = \frac{100}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right|. \quad (4)$$

Кроме того, учитывая специфику технологического процесса, предлагается использовать специализированную функцию потерь `EquipmentAwareLoss`, которая штрафует прогнозы, попадающие в критические зоны (например, давление выше 8 МПа или температура выше 150°C). Разработанная функция потерь имеет вид:

$$L = \frac{1}{n} \sum_{i=1}^n \left(\alpha_i \cdot |y_i - \hat{y}_i| + \beta \cdot \text{penalty}_{\text{critical}}(y_i, \hat{y}_i) \right). \quad (5)$$

Планируемый сценарий апробации. В качестве сравнительного базиса будет использоваться классический ПИД-регулятор, настроенный на номинальный режим. Ожидается, что адаптивная система позволит сократить время переходных процессов и снизить амплитуду колебаний давления по сравнению с традиционным подходом.

В рамках подготовки к реализации указанного сценария авторами уже проведён сравнительный анализ методов прогнозирования технологических параметров сепарации на основе промышленных данных. Из 37868 исходных записей после многоэтапной фильтрации сформирована репрезентативная выборка из 928 записей, соответствующих штатному режиму работы. Предварительное тестирование показало, что: `Random Forest` обеспечивает наивысшую точность прогнозирования давления ($R^2 = 0,889$) и температуры ($R^2 = 0,864$, $\text{MAPE} = 1,41\%$); `XGBoost` показывает наилучшие результаты для расхода газа ($R^2 = 0,422$, $\text{MAE} = 199,2 \text{ м}^3/\text{ч}$); классическая модель `ARIMA` дала отрицательные значения R^2 для всех параметров, что подтверждает необходимость использования нелинейных методов машинного обучения для данного технологического процесса. Полученные результаты создают основу для интеграции прогнозных моделей в контур адаптивного управления и будут использованы при имитационном моделировании.

Концептуальная модель, полученная на данном этапе, не претендует на завершённую прикладную реализацию. Её задача – сформировать методологическую основу для последующего построения математических моделей, имитационных экспериментов и алгоритмов адаптивного управления.

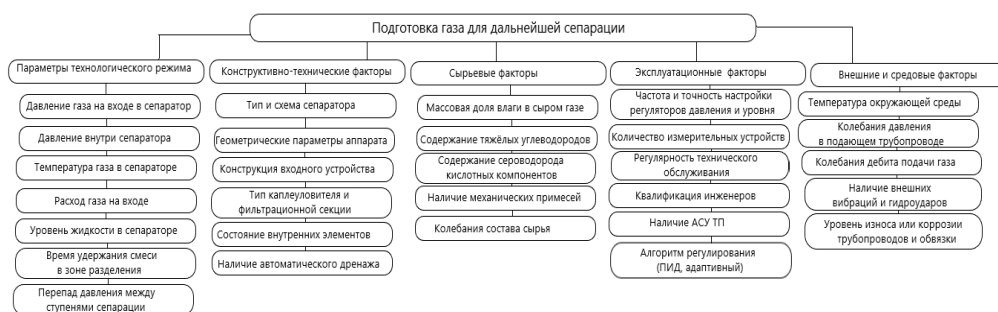


Рис. 2. Концептуальная модель управляемых процессов подготовки газа к сепарации

Заключение. Проведённое исследование позволило сформировать системное представление о совокупности факторов, влияющих на качество процессов сепарации и подготовки газа. Разработана концептуальная модель, отражающая структуру взаимосвязей между физико-техническими, конструктивными, сырьевыми, эксплуатационными и внешними параметрами. Модель базируется на принципах системного анализа и предназначена для определения ключевых направлений повышения эффективности подготовки газа.

Установлено, что наибольшее влияние на минимизацию примесей и стабильность технологического процесса оказывают параметры технологического режима (давление, температура, расход газа). Взаимосвязь этих факторов формирует основу для построения адаптивной системы управления процессом подготовки газа, способной учитывать изменчивость состава сырья и внешние условия эксплуатации. Для верификации модели применялись методы экспертных оценок. Значимость факторов подтверждена статистическими критериями – коэффициентами конкордации и корреляции Пирсона, что позволило уточнить устойчивость выявленных связей.

В дальнейшем предполагается разработка архитектуры адаптивного регулятора, включающей блоки идентификации, прогнозирования, оптимизации и нечёткого управления, а также проведение имитационного моделирования для оценки её эффективности по сравнению с классическим ПИД-регулятором в условиях переменного состава сырья. Результаты моделирования позволят уточнить количественные показатели улучшения стабильности и энергоэффективности.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Zemenkova M.Y., Chizhevskaya E.L., & Zemenkov Y.D. Intelligent monitoring of the condition of hydrocarbon pipeline transport facilities using neural network technologies, *Journal of Mining Institute*, 2022, 258, pp. 933-944. Available at: <https://doi.org/10.31897/PMI.2022.105>.
2. Ereemeeva A.M., Ilyushin Y.V. Automation of the control system for drying grain crops of the technological process for obtaining biodiesel fuels, *Sci Rep.*, 2023, 13, 14956. Available at: <https://doi.org/10.1038/s41598-023-41962-0>.
3. Qiu M., Ji Z., & Ma L. An energy-efficiency evaluation method for high-sulfur natural gas purification system using artificial neural networks and particle swarm optimization, *International Journal of Energy Research*, 2022, 46 (3). Available at: <https://doi.org/10.1002/er.7376>.
4. Li Y., Qin G., Xiong Z., & Fan L. Gas-liquid separation performance of a micro axial flow cyclone separator, *Chemical Engineering Science*, 2022, 249. Available at: <https://doi.org/10.1016/j.ces.2021.117234>.
5. Martirosyan A., Ilyushin Y., Afanaseva O., Kukharova T., Asadulagi M., & Khloponina V. 'Development of an Oil Field's Conceptual Model', *International Journal of Engineering*, 2025, 38 (2), pp. 381-388. Available at: <https://doi.org/10.5829/ije.2025.38.02b.12>.
6. Urzakov K.R., Belozero V.V., & Latypov B.M. Study of the dynamics for gas accumulation in the annulus of production wells, *Journal of Mining Institute*, 2021, 250, pp. 606-614. Available at: <https://doi.org/10.31897/PMI.2021.4.14>.
7. Nazarychev A., Iliev I., Manukian D., Beloev H., Suslov K., & Beloev I. Review of Operating Conditions, Diagnostic Methods, and Technical Condition Assessment to Improve Reliability and Develop a Maintenance Strategy for Electrical Equipment, *Energies*, 2025, 18 (21), 5832. Available at: <https://doi.org/10.3390/en18215832>.

8. Safiullin R.N., Zalyubovskiy A.F., Safiullin R.R., Sorokin K.V., & Avksentiev S.Y. An Algorithm to Process Big Data of Remote Technical Condition Assessment System for Mining Machines, *International Journal of Engineering, Transactions B: Applications*, 2026, 39(3), pp. 613-622. Available at: <https://doi.org/10.5829/ije.2026.39.03c.05>.
9. Tian J., Hu Y., Lv J., Feng Z., & Song W. Modelling and performance analysis of power system in gas engine-driven heat pump, *Applied Thermal Engineering*, 2023, 223. Available at: <https://doi.org/10.1016/j.applthermaleng.2023.120015>.
10. Yang Z., Li X., Xiang Q., He Q., Faber M.H., Zio E., Su H., & Zhang J. A resilience evaluation method of natural gas pipeline system based on uncertainty analysis, *Process Safety and Environmental Protection*, 2023, 177. Available at: <https://doi.org/10.1016/j.psep.2023.07.058>.
11. Afanaseva O.V., & Tulyakov T.F. Analysis of the effectiveness of traditional and modern technologies for monitoring power transmission lines, *Izvestiya SFedU. Engineering Sciences*, 2025, pp. 182-188. Available at: <https://doi.org/10.18522/2311-3103-2025-4-182-188>.
12. Li X.P., Shang D.Y., Li F.J., & Cao W.L. Dynamic Modeling and DME Evaluation of Power Transmission Line Inspection Robots, *Journal of Northeastern University*, 2020, 41, pp. 1280-1284. Available at: <https://doi.org/10.12068/j.issn.1005-3026.2020.09.011>.
13. Chen M., Tian Yu., & Xing Sh. (2021). Environment Perception Technologies for Power Transmission Line Inspection Robots, *Journal of Sensors*. 1. Available at: <https://doi.org/10.1155/2021/5559231>. EDN XLWPAO.
14. Mammadov Ja.F., Valiyeva B., Huseynova A.S.G., & Hasanova Ye.M.G. Development of diagnostic subsystem for manufacturing active elements in instrument-making industry, *Vestnik of Astrakhan State Technical University. Series: Management, Computer Science and Informatics*, 2022, 1, pp. 16-21. Available at: <https://doi.org/10.24143/2073-5529-2022-1-16-21>. EDN JPEWZU.
15. Mal'Tsev P.A., Shatilova N.A., Abramkin S.E., & Podkina M.E. Designing a Digital Twin in Oil and Gas Technologies: Technology Integration, 2024 XXVII International Conference on Soft Computing and Measurements (SCM), 2024, pp. 352-355. Available at: <https://doi.org/10.1109/SCM62608.2024.10554148>.
16. Asadulagi M.-A.M., Pershin I.M., & Tsapleva V.V. Research on Hydrolithospheric Processes Using the Results of Groundwater Inflow Testing, *Water*, 2024, 16, 487. Available at: <https://doi.org/10.3390/w16030487>.
17. Katysheva E. Analysis of the Interconnected Development Potential of the Oil, Gas and Transport Industries in the Russian Arctic, *Energies*, 2023, 16, 3124. Available at: <https://doi.org/10.3390/en16073124>.
18. Novozhilov I.M., Abramkin S.E., Maltsev P.A., Kuharova T.V., & Plotnikov A.V. Development of a Distributed Pressure Field Control System for Gas Production in a Structurally Complex Field, *LETI Transactions on Electrical Engineering & Computer Science*, 2025, 18 (4), pp. 55-66 Available at: <https://doi.org/10.32603/2071-8985-2025-18-4-55-66>.
19. Nikolaev A.V., & Kychkin A.V. Development and integration of an underground mining enterprise ventilation process simulations into the demand response, *Journal of Mining Institute*, 2025, 275, pp. 18-29. Retrieved from. Available at: <https://pmi.spmi.ru/pmi/article/view/16568>.
20. Kazantseva V.V., Ozhigin D.S., Kosarev N.S., Satbergenova A.K., & Ozhigina S.B. Development of complex system of geotechnical monitoring of technogenic objects based on geospatial data, *Journal of Mining Institute*, 2025. Retrieved from <https://pmi.spmi.ru/pmi/article/view/16590>.
21. He Y.W., He Z.Y., Tang Y., Xu Y.J., Long J.C., & Sepehrnoori K. Shale gas production evaluation framework based on data-driven models, *Petroleum Science*, 2023, 20 (3). Available at: <https://doi.org/10.1016/j.petsci.2022.12.003>.
22. Sidorenko A., Kriukov A., Meshkov A., & Sidorenko S.A. Comprehensive Technical and Economic Analysis of Rubber-Tyred Transport Implementation in Longwall Mining: A Case Study on the V.D. Yalovsky Coal Mine, *Mining*, 2025, 5, 65. Available at: <https://doi.org/10.3390/mining5040065>.
23. Marinina O., Malikov A., Lyubek Y., Pasternak S., Reshneva E., & Stolbovskaya N. Selection of Enhanced Oil Recovery Method on the Basis of Clustering Wells, *Processes*, 2024, 12, 2082. Available at: <https://doi.org/10.3390/pr12102082>.
24. Tokarev I.S. Development of parameters for an industry-specific methodology for calculating the electric energy storage system for gas industry facilities, *Journal of Mining Institute*, 2025, 272, pp. 171-180. Retrieved from: <https://pmi.spmi.ru/pmi/article/view/16516>.
25. Montenegro J., Veloz C., & Ricaurte M. Estimating water content of natural gas mixtures considering heavy hydrocarbons contribution using artificial neural networks, *Petroleum Science and Technology*, 2023, 41 (8). Available at: <https://doi.org/10.1080/10916466.2022.2075012>.

Мартirosян Александр Витальевич – Санкт-Петербургский горный университет императрицы Екатерины II; e-mail: Martirosyan_AV@pers.spmi.ru; г. Санкт-Петербург, Россия; тел.: +78123288566; к.т.н.; доцент кафедры системного анализа и управления.

Ромашин Даниил Васильевич – Санкт-Петербургский горный университет императрицы Екатерины II; e-mail: Daniil.V-Romashin@yandex.ru; г. Санкт-Петербург, Россия; тел.: +79620194779; аспирант кафедры системного анализа управления.

Martirosyan Alexander Vitalevich – Saint Petersburg Mining University of Empress Catherine II; e-mail: Martirosyan_AV@pers.spmi.ru; Saint Petersburg; Russia; phone: +78123288566; cand. of eng. sc.; associate professor of the Department of System Analysis and Management.

Romashin Daniil Vasilyevich – Saint Petersburg Mining University of Empress Catherine II; e-mail: Romashin@yandex.ru; e-mail: Romashin@yandex.ru; Saint Petersburg; Russia; phone: +79620194779; postgraduate student of the Department of System Analysis and Management.

УДК 519.21

DOI 10.18522/2311-3103-2026-4-249-259

В.Х. Пшихопов, А.Л. Веревкин, А.Р. Гайдук, Л.С. Веревкина

МАТРИЧНЫЙ МЕТОД ОЦЕНКИ НАДЁЖНОСТИ СЛОЖНЫХ МУЛЬТИАГЕНТНЫХ РЕКОФИГУРИРУЕМЫХ СИСТЕМ

Представлены разработанные матричный метод одновременной оценки надёжности нескольких вариантов сложных мультиагентных реконфигурируемых систем, а также стадийный метод резервирования агентов, отличающиеся от известных простотой и широкой областью применения. Матричный метод оценки надёжности заключается в матричном отображении параметров надёжности всех агентов конкретного вариантов группы агентов системы и векторном представлении условий выполнения миссии. Предполагается, что надёжность выполнения миссии целевой группой агентов определяется только надёжностью самих агентов и характером их резервирования. Метод стадийного резервирования отличается тем, что традиционные типовые схемы резервирования используются как стадии процесса резервирования группы агентов, что позволяет существенно повысить надёжность вариантов системы. Агенты системы делятся на активных, непосредственно выполняющих текущую миссию, и пассивных, находящихся в режиме ожидания активации. Активные агенты, в свою очередь, подразделяются на основные и резервные, причём основные агенты начинают выполнение миссии, а резервные агенты, ожидая замены отказавших активных агентов, могут быть как в состоянии горячего резерва, так и в состоянии холодного резерва. Специфические условия выполнения миссии учитываются вектором поправочных коэффициентов, что существенно повышает точность оценки надёжности и прогноза времени безотказной работы группы агентов. Матричный метод оценки надёжности позволяет определить расчётную надёжность всех возможных вариантов системы и найти целесообразный вариант, который имеет требуемые параметры надёжности для выполнения данной миссии в конкретных условиях. Метод стадийного резервирования позволяет существенно повысить надёжность вариантов системы, по сравнению с традиционными методами резервирования, что значительно расширяет область применения сложных мультиагентных реконфигурируемых систем. Важной особенностью предложенных методов оценки и прогнозирования надёжности является их аналитический характер, что позволяет использовать при решении задач обеспечения требуемой надёжности выполнения миссий сложными мультиагентными системами современные быстродействующие вычислительные средства.

Сложная мультиагентная система; реконфигурация; надёжность, матричный метод; поправочный коэффициент; стадийное резервирование.

V.Kh. Pshikhopov, A.L. Verevkin, A.R. Gaiduk, L.S. Verevkina

MATRIX METHOD FOR SIMULTANEOUS RELIABILITY ASSESSMENT OF COMPLEX MULTI-AGENT RECONFIGURABLE SYSTEMS

The article presents the developed matrix method of simultaneous assessment of the reliability of several variants of complex multi-agent reconfigurable systems, as well as the staged method of reservation, which differ from the known ones in their simplicity and wide scope of application The matrix reli-

bility assessment method consists in a matrix display of the reliability parameters of all agents of a specific variant of a group of system agents and a vector representation of the mission conditions. It is assumed that the mission reliability of a target group of agents is determined solely by the reliability of the agents themselves and the nature of their redundancy. The staged redundancy method is distinguished by its use of traditional standard redundancy schemes as stages in the process of redundant agent groups, significantly increasing the reliability of system variants. System agents are divided into active agents, directly performing the current mission, and passive agents, awaiting activation. Active agents, in turn, are divided into primary and backup agents. Primary agents initiate mission execution, while backup agents, awaiting replacement of failed active agents, can be in either an active state (hot standby) or a cold standby state. Specific mission conditions are taken into account by a vector of correction factors, significantly improving the accuracy of reliability assessment and agent group uptime prediction. An important feature of the proposed methods for assessing and predicting reliability is their analytical nature, which allows the use of modern high-speed computing tools in solving problems of ensuring the required reliability of missions performed by complex multi-agent systems.

Complex multiagent system; reconfiguration; reliability; matrix method; correction factor; staged redundancy.

Введение. В настоящее время всё более широкое применение находят сложные мультиагентные реконфигурируемые системы (СМАРС) с агентами различных типов: беспилотными авиационными системами (БАС), роботизированными колесными платформами (РКП), автономными необитаемыми подводными аппаратами (АНПА). Для успешного выполнения своих миссий СМАРС, используют GPS, Глонас, современные системы связи и искусственный интеллект [1–5]. Так в 2025 году в Китае была продемонстрирована скоординированная работа 15947 дронов, стартовавших с наземной площадки (рис. 1, слева) с целью создания светового шоу (рис. 1, справа). Управление дронами осуществлялось с единого компьютера в городе Люяне [5]. Эта СМАРС успешно выполнила требования регистрации в книге рекордов Гиннеса, согласно которым дроны должны были одновременно подняться в воздух и оставаться в полете не менее одной минуты.



Рис. 1. Световое шоу дронов в Китае

По данным мировой печати к 2033 году мировой рынок агродронов, управляемых искусственным интеллектом, покажет четырехкратный рост и достигнет 14 миллиардов долларов [6]. Эти дроны будут способны выполнять широкий спектр задач, от посева и мониторинга состояния посевов до обрезки деревьев и аккуратного сбора отдельных фруктов, таких как яблоки, персики, сливы и айва, без нанесения им повреждений [7].

Использование таких сложных мультиагентных систем ставит необходимость разработки новых методов оценки времени их бесперебойной работы и обеспечения требуемой продолжительности миссии. Трудность заключается в том, что СМАРС, могут формировать большое число возможных вариантов групп своих агентов. Оценка показателей надёжности каждой из них с использованием разных традиционных методов [8–12] требует существенных временных и вычислительных затрат.

Постановка задачи оценки надёжности. Как и в других работах, в статье принято, что закон распределения является экспоненциальным, надёжность каждого μ -го агента A_μ определяется его интенсивностью отказов λ_μ , наработкой на отказ T_μ и вероятностью безотказной работы $p_\mu(t)$. Эти показатели оцениваются с помощью статистических методов на основе данных, полученных в ходе испытаний [13, 14].

В случае экспоненциального закона распределения интенсивность отказов каждого агента в нормальном режиме обычно принимается постоянной, т.е. $\lambda_\mu = \text{const}$, поэтому вероятность его безотказной работы определяется выражением

$$p_\mu(t) = e^{-\lambda_\mu t}, \quad (1)$$

а средняя наработка на отказ:

$$T_\mu = 1/\lambda_\mu. \quad (2)$$

Экспоненциальный закон распределения типичен для ситуации, когда множество компонентов или устройств функционируют в штатном режиме. Это характерно для радиоэлектронных и мехатронных компонентов и устройств, в частности, робототехнических систем [15].

Ввиду того, что рабочие условия агентов в рамках миссий, как правило, отличаются от условий тестовых статистических испытаний на надёжность, каждый агент характеризуется индивидуальным поправочным коэффициентом интенсивности отказов (ПКИО) k_μ , который также может учитывать специальные условия функционирования и рассчитываться по различным формулам, как показано в работах [13–15]. Этим коэффициентом определяются: условная интенсивность отказов $\lambda_{\mu y}$, условная наработка на отказ $T_{\mu y}$ и соответствующая вероятность безотказной работы $p_\mu(t)$ μ -го агента по формулам:

$$\lambda_{\mu y} = \lambda_\mu k_\mu, \quad T_{\mu y} = \frac{1}{\lambda_{\mu y}} = \frac{T_\mu}{k_\mu}, \quad p_\mu(t) = e^{-\lambda_{\mu y} t}. \quad (3)$$

Отметим, что условная наработка на отказ $T_{\mu y}$ – представляет собой ожидаемый период функционирования агента, по истечении которого условная вероятность его безотказной работы снижается до значения $e^{-1} \approx 0,368$. Однако, в реальных условиях эксплуатации, как правило, требуются более высокие значения этого параметра, варьирующиеся от 0,7 до 0,95 и выше.

В результате реконфигурации мультиагентной системы образуется некоторое подмножество, группа агентов, называемых вариантами этой системы, ориентированных на выполнение миссии, поставленной перед этой системой. Во многих случаях рассматриваемые СМАРС могут выполнить некоторую миссию, используя не все агенты, входящие в их состав. При расчёте надёжности вариантов СМАРС, как и в случаях технических систем [14–19], в зависимости от характера взаимосвязей между их компонентами, применяются различные схемы расчёта надёжности вариантов. На рис. 2 изображена последовательная схема оценки надёжности системы без резервирования, состоящей из n агентов A_μ , $\mu = \overline{1, n}$.

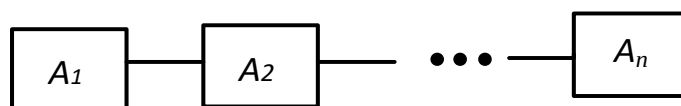


Рис. 2. Последовательная схема оценки надёжности

Здесь предполагается, что отказ любого агента влечёт невозможность выполнения системой своей функции. В этом случае, для вычисления характеристик надёжности системы применяются следующие выражения [14, 15]:

$$\Lambda_C = \sum_{\mu=1}^n \lambda_{\mu}, \quad \Lambda_{CY} = \sum_{\mu=1}^n \lambda_{\mu} k_{\mu}, \quad (4)$$

$$T_{CY} = 1 / \Lambda_{CY}, \quad p_C(t) = e^{-\Lambda_{CY} t} = \exp(-t / T_{CY}), \quad (5)$$

где Λ_C – интенсивность отказов, Λ_{CY} , T_{CY} – условные интенсивность отказов и наработка на отказ; $p_C(t)$ – вероятность безотказной работы системы – группы агентов (рис. 2); время t задаётся в интервале $0 \leq t \leq t_{\max}$, $t_{\max}$ – предполагаемое время работы группы в часах.

Выражения для определения Λ_{CY} , T_{CY} при других схемах расчета надёжности и учёта резервирования агентов приводятся в работах [14, 20, 21].

Матричный метод оперативной оценки надёжности вариантов СМАРС. Одновременная оценка условных интенсивностей отказов потенциальных вариантов СМАРС без резервирования, вычисляется по следующей формуле:

$$\Lambda_{CY} = A_{CI} K_{CI}, \quad (6)$$

где $A_{CI} = [a_{ij}]$ – $m \times n$ -матрица интенсивностей; $K_{CI} = [k_j]$ – n -вектор-столбец ПКНО активных агентов i -го варианта. Элементы A_{CI} и K_{CI} определяются выражениями:

$$a_{ij} = \begin{cases} \lambda_{\mu}, i = \overline{1, m}; j = \overline{1, n}; \\ 0, i = \overline{1, m-1}; j = \overline{i+1, n}, \end{cases} \quad (7)$$

$$k_j = k_{\mu}, \mu = j, j = \overline{1, n}, \quad (8)$$

$$m = N_C - N_1 + 1, \quad n = N_C. \quad (9)$$

В (7)–(9) $\Lambda_{CY}^T = [\Lambda_{CYi}]$, $i = \overline{1, m}$ – вектор интенсивностей отказов возможных вариантов рассматриваемой СМАРС, т.е. $\Lambda_{CY}^T = [\Lambda_{CY1} \Lambda_{CY2} \dots \Lambda_{CYm}]$; i – номер варианта СМАРС; j – номера активных агентов i -го варианта; m – число потенциальных вариантов; N_C – общее число агентов СМАРС; N_1 – число агентов её 1-го варианта;

Интенсивность Λ_{CYi} отказов i -го варианта СМАРС и число его агентов N_i , определяются выражениями:

$$\Lambda_{CYi} = A_{CIi} K_{CI}, \quad N_i = N_1 + i - 1, \quad i = \overline{1, m}, \quad (10)$$

где A_{CIi} – i -я строка матрицы A_{CI} .

Вектор условных времён наработки на отказ вариантов СМАРС:

$$T_{CY} = [1 / \Lambda_{CY1} \quad 1 / \Lambda_{CY2} \quad \dots \quad 1 / \Lambda_{CYm}]^T. \quad (11)$$

Так как СМАРС, в общем случае, может выполнить поставленную миссию с некоторой вероятностью при использовании не всех её агентов, то возникает вопрос о выборе наиболее подходящей конфигурации для успешного выполнения конкретной миссии. Согласно техническим требованиям, для того чтобы СМАРС могла гарантированно достичь цели, необходимо выполнение следующих условий:

$$p_C(t) \geq p_{TP}, \quad 0 \leq t \leq T_{TP}, \quad (12)$$

где p_{TP} – требуемое значение вероятности безотказной работы варианта СМАРС в течение требуемого времени T_{TP} . Значения p_{TP} и T_{TP} – требуемые показатели надежности выполнения конкретной миссии.

Вариант системы, показатели надёжности которого удовлетворяют условиям (12), принимается целесообразным для выполнения миссии с заданными значениями p_{TP} и T_{TP} .

Из выражений (11) и (12) следует, что если имеется m вариантов некоторой СМАРС, каждый из которых характеризуется значением групповой наработки на отказ:

$$t_{pTi} = T_{Cyi} \ln(1/p_{TP}), \quad i = \overline{1, m}, \quad (13)$$

и выполнено условие:

$$t_{pTmax} \geq T_{TP}, \quad (14)$$

где $t_{pTmax} = (\max_{i \in [1, m]} T_{Cyi}) \ln(1/p_{TP})$, то целесообразный вариант этой СМАРС для выполнения миссии, с требуемыми значениями p_{TP} и T_{TP} удобно находить на основе выражения:

$$t_{pT, i^*-1} < T_{TP} \leq t_{pT, i^*}, \quad i^* \in [1, m] \quad (15)$$

где i^* – номер целесообразного варианта, t_{pT, i^*} – групповая наработка на отказ этого варианта.

Таким образом, ключевым моментом является определение такого набора агентов (целесообразного варианта) данной СМАРС, который обладает вероятностью выполнения поставленной конкретной миссии не хуже требуемого значения p_{TP} по истечении интервала времени T_{TP} с момента начала выполнения этой миссии. Выражения (6)–(15) позволяют решить эту задачу аналитически.

Алгоритм выбора целесообразного варианта СМАРС. На основе этих выражений разработан алгоритм выбора целесообразного варианта некоторой СМАРС для выполнения миссии с заданными значениями p_{TP} и T_{TP} . Блок-схема этого алгоритма приведена на рис. 3.

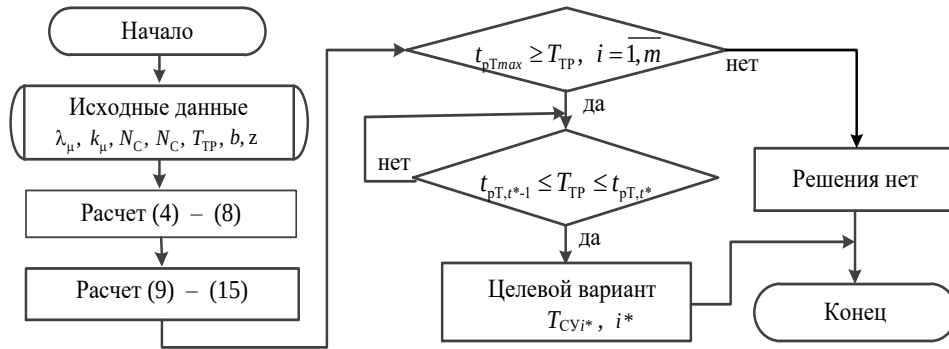


Рис. 3. Алгоритм выбора целесообразного варианта СМАРС

Матричный метод оценки надежности. Рассмотрим применение матричного метода для определения возможности применения без резервирования СМАРС, агентами которой являются один пульт и пять роботизированных колёсных платформ (РКП) (рис. 4) для выполнения миссии с требуемыми значениями вероятности безотказной работы не менее $p_{TP} = 0,85$ в течении $T_{TP} = 600$ часов.

Исходные данные: $\lambda_{\Pi} = 80 \cdot 10^{-6}$ 1/час – интенсивность отказов пульта управления; интенсивность отказов каждой РКП: $\lambda_{\text{ркп1}} = 400 \cdot 10^{-6}$ 1/час; $\lambda_{\text{ркп2}} = 500 \cdot 10^{-6}$ 1/час; $\lambda_{\text{ркп3}} = 450 \cdot 10^{-6}$ 1/час; $\lambda_{\text{ркп4}} = 550 \cdot 10^{-6}$ 1/час; $\lambda_{\text{ркп5}} = 300 \cdot 10^{-6}$ 1/час; $k_{\Pi} = 1,00$ – ПКИО пульта; и ПКИО РКП: $k_{\text{ркп},1} = 1,0$; $k_{\text{ркп},2} = 1,1$; $k_{\text{ркп},3} = 1,2$; $k_{\text{ркп},4} = 1,1$; $k_{\text{ркп},5} = 1,3$.

На основе соотношений (6)–(12) разработанного матричного метода и алгоритма выполнено имитационное моделирование оценки надёжности вомплекса с возможностью изменения состава агентов РКП от одной до пяти.



Рис. 4. СМАРС с пультом и пятью РКП

Принимая во внимание обязательное включение пульта управления во всех рассматриваемых вариантах, получаем значения параметров: $N_1 = 2$; $m = 5$; $n = 6$; тогда $A_{\text{СИ}}$ – матрица 5×6 ; $K_{\text{СИ}}$ – вектор с 6 элементами. Подставив численные значения в (7), (8) и (6) получим:

$$A_{\text{СИ}} = 10^{-6} \begin{bmatrix} 80 & 400 & 0 & 0 & 0 & 0 \\ 80 & 400 & 500 & 0 & 0 & 0 \\ 80 & 400 & 500 & 450 & 0 & 0 \\ 80 & 400 & 500 & 450 & 550 & 0 \\ 80 & 400 & 500 & 450 & 550 & 300 \end{bmatrix} \begin{bmatrix} 1,0 \\ 1,0 \\ 1,1 \\ 1,2 \\ 1,1 \\ 1,3 \end{bmatrix} = 10^{-6} \begin{bmatrix} 480 \\ 1030 \\ 1570 \\ 2175 \\ 2565 \end{bmatrix}. \quad (16)$$

Условная наработка на отказ каждого варианта данной СМАРС на основе результатов (16) определяется по формуле (11):

$$T_{\text{СУ}} = [2083 \quad 971 \quad 637 \quad 460 \quad 390]^T, \text{ час.} \quad (17)$$

При $p_{\text{ТР}} = 0,85$ по формуле (13) с учётом (17) находим: $t_{pT_{\text{max}}} = 338$, т.е. в данном случае условие (14) не выполняется ни при каком значении $i \in [1, 5]$, так как $T_{\text{ТР}} = 600$ часов.

Отсюда следует, что данная СМАРС без резервирования не может применяться для выполнения поставленной миссии из-за недостаточной надёжности. Что подтверждается графиками вероятности безотказной работы рассмотренных вариантов СМАРС приведенными на рис. 5. Следовательно, ни один из возможных вариантов рассматриваемой СМАРС не удовлетворяет условиям (12)–(15), и применить без резервирования рассматриваемую СМАРС, для выполнения поставленной миссии невозможно из-за низкой надёжности.

Для повышения надёжности чаще всего используется резервирование [9, 10].

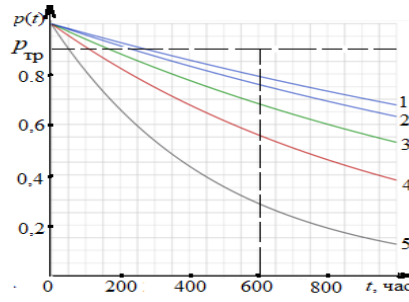


Рис. 5. Графики вероятности безотказной работы вариантов СМАРС

Метод стадийного резервирования агентов. Рассмотрим новый метод стадийного резервирования агентов, вытекающий из матричного метода расчёта надёжности. Метод стадийного резервирования заключается использовании традиционных типовых схем резервирования, как стадий процесса резервирования с целью существенного повышения надёжности технических систем.

Как отмечалось выше, в схемах расчёта надёжности СМАРС с резервированием выделяются основные и резервные агенты. Число основных агентов обозначается b , а число резервных – z . Таким образом, общее число агентов в СМАРС $N_C = b + z$ или больше.

В случае стадийного резервирования расчёт надёжности выполняется также матричным методом, но с некоторыми изменениями, так как матрица параметров агентов и вектор ПКМО определяются в этом случае по наработке на отказ T_μ или $T_{\mu\mu}$ агентов. При этом элементы вектора поправочных коэффициентов K_{CT} равны обратным значениям коэффициентов k_μ или единице. Нарботка на отказ определяется по аналогичной выражению (6) формуле:

$$T_{CY} = A_{CT} K_{CT}. \quad (18)$$

Элементы $m \times n$ -матрицы A_{CT} и n -вектора K_{CT} определяются по формулам, зависящим от схемы резервирования.

В качестве примера эффективности стадийного резервирования реконфигурируемой СМАРС, по сравнению с вариантом непосредственного использования всех агентов, рассчитаем надёжность рассмотренной выше СМАРС (рис. 4), когда пульт и первая РКП являются основными, а остальные четыре РКП – резервными. Это соответствует случаю избыточного холодного стадийного резервирования при $b = 2$ и $z = 4$. В этом случае: $N_C = 6$, $N_1 = 2$, $m = z = 4$, $n = z + 1 = 5$ [10, 11]. Элементы матрицы A_{CT} и вектора K_{CT} определяются выражениями:

$$a_{ij} = \begin{cases} T_{n1}, & i = \overline{1, m}; j = 1; \\ T_j, & i = \overline{1, m}; j = \overline{2, 5}; \\ 0, & i = \overline{1, m-1}; j = i + 2, n; \end{cases} \quad (19)$$

$$K_{CT}^T = [1 \ 1 \ 1 \ 1 \ 1], \quad (20)$$

$$T_{n1} = \frac{1}{\lambda_\pi k_\pi + \lambda_{ркп1} k_{ркп1}}, \quad T_j = \frac{1}{\lambda_\pi k_\pi + \lambda_{ркпj} k_{ркпj}}, \quad j = \overline{2, 5}, \quad (21)$$

где T_{n1} – наработка на отказ пульта и первой РКП (определена в выражении (17)); T_j – наработка на отказ пульта и j -ой резервной РКП.

Схема реализации рассматриваемой реконфигурируемой системы показана на рис. 5. На рис. 6 символами П и РКП обозначены те же, что и на рис. 4 пульт управления и роботизированная колёсная платформа. В горизонтальных строчках этого рисунка показаны только три варианта: 1в, 2в и 3в (для краткости), а вертикальные столбцы соответствуют стадиям резервирования: 1с, 2с, 3с и 4с.

	1с	2с	3с	4с
1в	— П — РКП1	— П — РКП2	—	—
2в	— П — РКП1	— П — РКП2	— П — РКП3	—
3в	— П — РКП1	— П — РКП2	— П — РКП3	— П — РКП4

Рис. 6. Варианты и стадии холодного резервирования

Элементы первого варианта на рис. 6 являются традиционной схемой холодного резервирования. Как видно, в последующих вариантах эта схема повторяется на каждой стадии. В этом и заключается стадийный метод резервирования, который может применяться не только при избыточном холодном резервировании, но и при других традиционных способах резервирования.

В данном случае, подставляя в формулу (21) приведённые выше значения параметров надёжности пульта и РКП, найдём значения $T_{п1}$ и T_j , $j = 2, 5$, а затем по (19) определяем значения a_{ij} , подставляя которые с учётом (20) в формулу (18), получим:

$$T_{cy} = \begin{bmatrix} 2083 & 1587 & 0 & 0 & 0 \\ 2083 & 1587 & 1613 & 0 & 0 \\ 2083 & 1587 & 1613 & 1460 & 0 \\ 2083 & 1587 & 1613 & 1460 & 2128 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 3670 \\ 5283 \\ 6743 \\ 8871 \end{bmatrix}. \quad (22)$$

На основе результатов (22) вектор групповой наработки на отказ вариантов рассматриваемой СМАРС при холодном стадийном резервировании и требуемой вероятности безотказной работы $p_{Tr} = 0,875$ имеет вид:

$$t_{pT} = [490 \quad 705 \quad 900 \quad 1184], \text{ час.} \quad (23)$$

Так как $T_{Tr} = 600$ час, а $t_{pT, \max} = 1184$ часов, то условие (14) выполняется и из неравенства (15) следует, что целесообразным вариантом при стадийном холодном резервировании для выполнения поставленной миссии является второй вариант, то есть $i^* = 2$. Как видно из схемы на рис. 5, этот вариант включает 4 агента, т.е. пульт и три РКП: РКП1, РКП2 и РКП3.

Корректность полученного результата подтверждается графиками, приведёнными на рис. 7.

На основе полученных результатов можно сделать вывод, что резервирование агентов является более эффективным методом повышения надёжности по сравнению с простым увеличением агентов в формируемых группах агентов.

Таким образом, используя рассмотренные выше методы, можно осуществить расчёт параметров надёжности других СМАРС, и осуществить определение целесообразного варианта для выполнения поставленных перед ними миссий. Для этого должны быть заданы указанные выше параметры надёжности агентов этих систем. Подчеркнём, что

предложенные в работе методы позволяют обоснованно найти не только целесообразный вариант группы агентов для выполнения поставленной миссии, но и схему резервирования агентов этой группы.

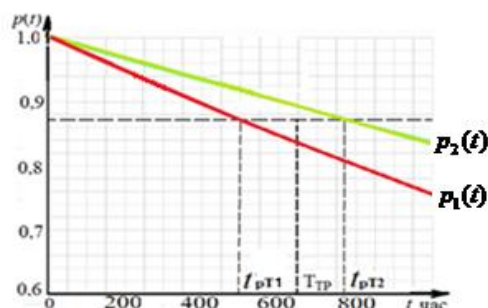


Рис. 7. Вероятности безотказной работы 1-го и 2-го вариантов

Примеры расчёта СМАРС с горячим резервированием агентов приведены в работе [15]. В случае необходимости оптимизации количества агентов некоторого варианта может использоваться генетический алгоритм, предложенный в работе [16].

Заключение. Рассмотренные методы ориентированы на применение при решении актуальных в настоящее время задач обеспечения надёжности выполнения миссий сложными мультиагентными реконфигурируемыми системами. Особую актуальность этих задач придаёт стремительное развитие беспилотных систем воздушного, подводного и наземного базирования. Эти системы могут включать значительное число обособленных автоматически действующих единиц – агентов, из которых эти системы могут формировать, путём реконфигурации группы агентов, ориентированные на решение той или иной конкретной задачи (миссии). Однако, как и любая техническая система, эти группы могут достичь поставленных целей, если только они обладают достаточно высокой надёжностью.

Рассмотренные в работе матричный методы одновременной оценки надёжности нескольких групп агентов, а также стадийный метод резервирования могут существенно повысить надёжность достижения поставленных целей. Они ориентированы на применение при использовании мультиагентных систем для решения задач промышленной, медицинской, социальной логистики, а также для решения специальных задач.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Аналитический обзор мирового рынка робототехники. – <https://robotrends.ru/robopedia/roi-bespilotnikov> (дата доступа: 19.12.2025).
2. Геоскан для управления роем дронов. – <https://robotrends.ru/pub/2406/geoskan-predstavil-reshenie-dlya-upravleniya-roem-dronov> (дата доступа: 19.12.2025).
3. Эскадрилья для тушения природных пожаров. – <https://www.rbc.ru/rbcfreenews/6167e9b99a7947e8e3e8da25> (дата доступа: 19.12.2025).
4. Веревкин А.Л. Алгоритм компьютерного моделирования надёжности групп различной численности АНПА комплекса сейсморазведки океана // Проблемы современной системотехники: Сб. научных статей. Выпуск XVI. – Таганрог: Изд-ль ИП Ступин С.А., 2022. – 98 с. – С. 81-89.
5. <https://ok.ru/videoembed/9279008606777> (дата доступа: 09.12.2025).
6. Рынок агродронов к 2033 году вырастет в четыре раза. – https://uza.uz/ru/posts/rynok-agrodronov-k-2033-godu-vyrastet-v-chetyre-raza_505760 (дата доступа: 19.12.2025).
7. Дроны для сбора фруктов. Forbes.ru. – <https://www.forbes.ru/tekhnologii/485668-zacem-fermery-v-ssa-vnedraut-robotizirovannye-ul-i-i-drony-dla-sbora-fruktoy> (дата доступа: 19.12.2025).
8. Штефан Ю.В., Федоров В.К. Основы теории надёжности: учебно-методическое пособие к семинарским занятиям и лабораторным работам. – М.: МАДИ, 2023. – 168 с.
9. ГОСТ Р МЭК 61165 Надёжность в технике. Применение Марковских методов.
10. Шорохова И.С., Кисляк Н.В., Мариев О.С. Статистические методы анализа. – Екатеринбург: Изд-во Урал. ун-та, 2015. – 300 с.

11. *Крянев А.В., Семенов С.С.* К вопросу о качестве и надёжности экспертных оценок при определении технического уровня сложных систем // *Функциональная надёжность. Теория и практика.* – <https://www.dependability.ru/jour/article/viewFile/34/98> (дата доступа: 27.10.2025).
12. ГОСТ 27.301 Надёжность в технике. Расчёт надёжности. Основные положения.
13. *Пилюхов В.Х., Вережкин А.Л., Вережкин А.Л.* Проектирование электротехнических устройств: учеб. пособие. Ч. 1. Стадии проектирования. Техническое задание. – Таганрог: Изд-во ЮФУ, 2013. – 50 с.
14. *Капитанов В.А., Медведев А.И.* Теория надёжности сложных систем. – 2-е изд. перераб. – М.: Физматлит, 2010. – 608 с.
15. *Павлов А.П.* Методические указания к выполнению практической части курсовой работы по дисциплине «Резервирование при производстве, эксплуатации и ремонте машин». – М.: МАДИ, 2022. – 106 с.
16. *Острейковский В.А.* Теория надёжности: учебник для вузов. – 2-е изд., испр. – М.: Высшая школа, 2008. – 464 с.
17. *Дулов О.А., Абрамов А.В.* Основы теории надёжности: Практикум по дисциплине «Основы теории надёжности». – Ульяновск: УлГТУ, 2015. – 62 с.
18. *Плютов Ю.А., Хомич Л.В.* Новые подходы к анализу надёжности // *Фундаментальные исследования.* – 2004. – № 6. – С. 70-72. – <https://fundamental-research.ru/ru/article/view?id=6298> (дата доступа: 19.12.2025).
19. *Болдырев А.С., Вережкин А.Л., Пилюхова К.В., Вережкин А.Л.* Методика и моделирование надёжности системы группового управления роботизированными платформами // *Известия ЮФУ. Технические науки.* – 2020. – № 3 (123). – С. 16-29.
20. *Болдырев А.С., Вережкин А.Л., Вережкин А.Л.* Алгоритмы моделирования надёжности систем управления группой роботизированных платформ с горячим резервированием // *Известия СПб ГТЭУ «ЛЭТИ».* – 2021. – № 4. – С. 49-57.
21. *Вережкин А.Л., Пилюхова К.В., Вережкин А.Л.* Генетический алгоритм оценки надёжности многоканальной системы контроля, с учетом вариантов резервирования // *Компьютерные и информационные технологии в науке, инженерии и управлении, «КомТех-2019»:* Матер. Всероссийской научно-технической конференции с международным участием имени профессора О.Н. Пьявченко / отв. ред. С.И. Клевцов. – Ростов-на-Дону; Таганрог: Изд-во ЮФУ, 2019. – С. 406-414.

REFERENCES

1. Analiticheskiy obzor mirovogo rynka robototekhniki [Analytical review of the global robotics market]. Available at: <https://robotrends.ru/robopedia/roi-bespilotnikov> (accessed 19 December 2025).
2. Geoskan dlya upravleniya roem dronov [Geoscan for drone swarm management]. Available at: <https://robotrends.ru/pub/2406/geoskan-predstavil-reshenie-dlya-upravleniya-roem-dronov> (accessed 19 December 2025).
3. Eskadriya dlya tusheniya prirodnikh pozharov [Wildfire Squadron]. Available at: <https://www.rbc.ru/rbcfreenews/6167e9b99a7947e8e3e8da25> (accessed 19 December 2025).
4. *Verezhkin A.L.* Algoritmy kompyuternogo modelirovaniya nadezhnosti grupp razlichnoy chislennosti ANPA kompleksa seysmorazvedki okeana [Algorithm for computer modeling of the reliability of groups of various sizes of AUVs for ocean seismic exploration], *Problemy sovremennoy sistemotekhniki: Sb. nauchnykh statey. Vypusk XVI* [Problems of modern systems engineering, collection of scientific articles. Issue XVI]. Taganrog: Izd-l' IP Stupin S.A., 2022, 98 p., pp. 81-89.
5. Available at: <https://ok.ru/videoembed/9279008606777> (accessed 19 December 2025).
6. Rynok agrodronov k 2033 godu vyrastet v chetyre raza [The agrodrones market will quadruple by 2033]. Available at: https://uza.uz/ru/posts/rynok-agrodronov-k-2033-godu-vyrastet-v-chetyre-raza_505760 (accessed 19 December 2025).
7. Drony dlya sbora fruktov. Forbes.ru [Fruit picking drones. Forbes.ru]. Available at: <https://www.forbes.ru/tekhnologii/485668-zacem-fermery-v-ssa-vnedraut-robotizirovannye-ul-i-i-drony-dla-sbora-fruktov> (accessed 19 December 2025).
8. *Shtefan Yu.V., Fedorov V.K.* Osnovy teorii nadezhnosti: uchebno-metodicheskoe posobie k seminariskim zanyatiyam i laboratornym rabotam [Fundamentals of the theory of reliability: a teaching manual for seminars and laboratory work]. Moscow: MADI, 2023, 168 p.
9. GOST R MEK 61165 Nadezhnost' v tekhnike. Primenenie Markovskikh metodov [GOST R IEC 61165 Reliability in technology. Application of Markov methods].
10. *Shorokhova I.S., Kislyak N.V., Mariev O.S.* Statisticheskie metody analiza [Statistical methods of analysis]. Ekaterinburg: Izd-vo Ural. un-ta, 2015, 300 p.

11. Kryanev A.V., Semenov S.S. K voprosu o kachestve i nadezhnosti ekspertnykh otsenok pri opredelenii tekhnicheskogo urovnya slozhnykh sistem [On the quality and reliability of expert assessments in determining the technical level of complex systems], *Funktsional'naya nadezhnost'. Teoriya i praktika* [Functional Reliability: Theory and Practice]. Available at: <https://www.dependability.ru/jour/article/viewFile/34/98> (accessed 27 October 2025).
12. GOST 27.301 Nadezhnost' v tekhnike. Raschet nadezhnosti. Osnovnye polozheniya [GOST 27.301 Reliability in engineering. Reliability calculation. Main provisions].
13. Pshikhopov V.Kh., Verevkin A.L., Verevkin A.L. Proektirovanie elektrotekhnicheskikh ustroystv: ucheb. posobie. Ch. 1. Stadii proektirovaniya. Tekhnicheskoe zadanie [Design of electrical devices: tutorial. Part 1. Design stages. Terms of reference]. Taganrog: Izd-vo YuFU, 2013, 50 p.
14. Kashtanov V.A., Medvedev A.I. Teoriya nadezhnosti slozhnykh sistem [Reliability theory of complex systems]. 2nd revised ed. Moscow: Fizmatlit, 2010, 608 p.
15. Pavlov A.P. Metodicheskie ukazaniya k vypolneniyu prakticheskoy chasti kursovoy raboty po distsipline «Rezervirovanie pri proizvodstve, ekspluatatsii i remonte mashin» [Guidelines for Completing the Practical Part of the Coursework on "Redundancy in the Production, Operation, and Repair of Machines]. Moscow: MADI, 2022, 106 p.
16. Ostreykovskiy V.A. Teoriya nadezhnosti: uchebnyk dlya vuzov [Reliability theory: textbook for universities]. 2nd revised ed. Moscow: Vysshaya shkola, 2008, 464 p.
17. Dulov O.A., Abramov A.V. Osnovy teorii nadezhnosti: Praktikum po distsipline «Osnovy teorii nadezhnosti» [Fundamentals of Reliability Theory: Workshop on the Course "Fundamentals of Reliability Theory"]. Ul'yanovsk: UGTU, 2015, 62 p.
18. Plyutov Yu.A., Khomich L.V. Novye podkhody k analizu nadezhnosti [New approaches to reliability analysis], *Fundamental'nye issledovaniya* [Fundamental Research], 2004, No. 6, pp. 70-72. Available at: <https://fundamental-research.ru/ru/article/view?id=6298> (accessed 19 December 2025).
19. Boldyrev A.S., Verevkin A.L., Pshikhopova K.V., Verevkin A.L. Metodika i modelirovanie nadezhnosti sistemy gruppovogo upravleniya robotizirovannymi platformami [Methodology and modeling of reliability of a group control system for robotic platforms], *Izvestiya YuFU. Tekhnicheskie nauki* [Izvestiya SFedU. Engineering Sciences], 2020, No. 3 (123), pp. 16-29.
20. Boldyrev A.S., Verevkin A.L., Verevkin A.L. Algoritmy modelirovaniya nadezhnosti sistem upravleniya gruppoy robotizirovannykh platform s goryachim rezervirovaniem [Algorithms for modeling the reliability of control systems for a group of robotic platforms with hot redundancy], *Izvestiya SPb GTEU «LETI»* [Bulletin of ETU "LETI"], 2021, No. 4, pp. 49-57.
21. Verevkin A.L., Pshikhopova K.V., Verevkin A.L. Geneticheskii algoritm otsenki nadezhnosti mnogokanal'noy sistemy kontrolya, s uchetom variantov rezervirovaniya [Genetic algorithm for assessing the reliability of a multi-channel control system, taking into account redundancy options] *Kompyuternye i informatsionnye tekhnologii v nauke, inzhenerii i upravlenii, «KomTekh-2019»: Mater. Vserossiyskoy nauchno-tekhnicheskoy konferentsii s mezhdunarodnym uchastiem imeni professora O.N. Pyavchenko* [Computer and Information Technologies in Science, Engineering and Management, "KomTech-2019": Proceedings of the All-Russian Scientific and Technical Conference with International Participation named after professor O.N. Pyavchenko], ed. S.I. Klevtsov. Rostov-on-Don; Taganrog: Izd-vo YuFU, 2019, pp. 406-414.

Пшихопов Вячеслав Хасанович – Южный федеральный университет; e-mail: vpshihopov@sfedu.ru; г. Таганрог, Россия; тел.: +78634371634; д.т.н.; профессор.

Веревкин Александр Леонидович – Южный федеральный университет; e-mail: verevkin.a@mail.ru; г. Таганрог, Россия; тел.: +78634371634; соискатель.

Гайдук Анатолий Романович – Южный федеральный университет; e-mail: argayduk@sfedu.ru; г. Таганрог, Россия; тел.: +78634371689; д.т.н.; профессор.

Веревкина Лина Станиславовна – Южный федеральный университет; e-mail: lverevkina@sfedu.ru; г. Таганрог, Россия; тел.: +78634371634; к.т.н.; доцент.

Pshikhopov Vyacheslav Khasanovich – Southern Federal University; e-mail: vpshihopov@sfedu.ru; Taganrog, Russia; phone: +78634371634; dr. of eng. sc.; professor.

Verevkin Alexander Leonidovich – Southern Federal University; e-mail: verevkin.a@mail.ru; Taganrog, Russia; phone: +78634371634; applicant.

Gaiduk Anatoly Romanovich – Southern Federal University; e-mail: argayduk@sfedu.ru; Taganrog, Russia; phone: +78634371689; dr. of eng. sc.; professor.

Verevkin Lina Stanislavovna – Southern Federal University; e-mail: lverevkina@sfedu.ru; Taganrog, Russia; phone: +78634371634; cand. of eng. sc.; associate professor.

Э.М. Эндреев, В.А. Шаповалов, А.А. Аджиева, Г.В. Куповых, В.В. Гривцов

**СУХИЕ ВТОРЖЕНИЯ В СРЕДНЕЙ ТРОПОСФЕРЕ КАК ФАКТОР
ФОРМИРОВАНИЯ ОПАСНЫХ НИСХОДЯЩИХ КОНВЕКТИВНЫХ ПОТОКОВ:
АНАЛИЗ И ДИАГНОСТИКА НА ЮГЕ ЕВРОПЕЙСКОЙ ЧАСТИ РОССИИ**

На основе 34-летнего архива (1990–2023 гг.) данных аэрологического зондирования 11 станций юга России, материалов реанализа ERA5 проведён систематический климатологический анализ роли сухих вторжений в средней тропосфере (700–500 гПа) как предиктора опасных нисходящих конвективных потоков и микропорывов. Сформированы отдельные выборки из 268 верифицированных случаев сильных приземных шквалов конвективного происхождения (скорость ветра ≥ 25 м/с) и 1042 контрольных дней без опасного ветра при наличии грозовой активности. Установлено, что в дни со шквалами дефицит точки росы на 700 гПа в среднем превышает контрольные значения в 2,4 раза (18,3 против 7,7°C), тогда как CAPE в обеих группах статистически значимо не различается. Введён и апробирован такой диагностический показатель как индекс нисходящей конвективной энергии с учетом охлаждения за счёт испарения осадков в сухом среднем тропосферном слое ($DCAPE_{dry}$), представляющий собой произведение стандартного DCAPE на нормированный дефицит точки росы в слое 700–500 гПа. Показано, что $DCAPE_{dry}$ при пороге ≥ 620 Дж/кг обеспечивает вероятность обнаружения шкала $POD = 84\%$ при доле ложных тревог $FAR = 19\%$, что принципиально улучшает существующие диагностические подходы ($DCAPE$ в изоляции: $POD = 61\%$, $FAR = 43\%$). Обсуждается роль испарительного охлаждения, механизмов генерации вихрей у поверхности и влияния орографии Предкавказья и северных склонов Большого Кавказа на интенсификацию нисходящих потоков. Орографические особенности (долинный фокусирующий эффект, фёновые зоны, конвергентные зоны на границе гор и равнин) систематически усиливают интенсивность нисходящих потоков, способствуя концентрации наиболее опасных шквалов в полосе от Нальчика до Ставрополя.

Сухие вторжения; нисходящий конвективный поток; шквал; микропорыв; $DCAPE$; дефицит точки росы; ERA5; испарительное охлаждение; юг России; Предкавказье; Большой Кавказ.

E.M. Andreyev, V.A. Shapovalov, A.A. Adzhieva, G.V. Kupovkh, V.V. Grivtsov

**DRY INTRUSIONS IN THE MIDDLE TROPOSPHERE AS A FACTOR IN THE
FORMATION OF DANGEROUS DOWNWARD CONVECTIVE FLOWS: ANALYSIS
AND DIAGNOSTICS IN THE SOUTH OF THE EUROPEAN PART OF RUSSIA**

A 34-year archive (1990–2023) of radiosonde data from 11 stations in southern Russia, ERA5 reanalysis is used to analyse mid-tropospheric dry intrusions (700–500 hPa) as a predictor of hazardous convective downdrafts, downbursts and microbursts. Separate samples of 268 verified cases of severe convective squalls (wind ≥ 25 m s⁻¹) and 1042 control thunderstorm days without damaging wind are compiled. Mean dewpoint depression at 700 hPa is 2.4 times higher in squall days than in controls (18.3 vs. 7.7°C), whereas CAPE shows no statistically significant difference between the two groups. A novel diagnostic parameter – Dry-Enhanced Downdraft CAPE ($DCAPE_{dry}$) – is derived as the product of standard DCAPE and the normalised dewpoint deficit in the 700–500 hPa layer. $DCAPE_{dry} \geq 620$ J kg⁻¹ achieves $POD = 84\%$ and $FAR = 19\%$ in squall detection, markedly outperforming DCAPE alone ($POD = 61\%$, $FAR = 43\%$). The roles of evaporative cooling, surface vortex generation, and Ciscaucasian orography in downdraft intensification are discussed. Orographic features (valley focusing effect, foehn zones, and convergent zones at the border between mountains and plains) systematically increase the intensity of descending flows, contributing to the concentration of the most dangerous squalls in the area from Nalchik to Stavropol.

Dry intrusions; convective downdraft; squall; downburst; microbursts; DCAPE; dewpoint depression; ERA5; evaporative cooling; southern Russia; Ciscaucasia; Caucasus Major.

Введение. Опасные нисходящие конвективные потоки – микропорывы (microbursts, $d < 4$ км) и шквалы (downbursts, $d > 4$ км) – являются одними из наиболее трудно прогнозируемых и опасных явлений конвективного происхождения. По данным мировой статистики, около 25% авиационных катастроф в зоне конвективной погоды связаны именно с нисходящими потоками, а приземная скорость ветра в них нередко превышает 40–50 м/с, причиняя серьёзный ущерб инфраструктуре, сельскому хозяйству и лесному фонду [1, 2].

Физика формирования мощных нисходящих потоков принципиально отличается от физики восходящих конвективных течений. Если восходящий поток «питается» положительной плавучестью от конденсации водяного пара, то нисходящий обязан своей мощностью двум основным источникам отрицательной плавучести: силе тяжести, действующей на осадки, и испарительному охлаждению воздуха при прохождении гидрометеоров через сухие слои атмосферы [3, 4]. Именно второй механизм определяет экстремальные случаи: чем суше воздух под облаком или в промежуточных слоях, тем интенсивнее испарение и тем сильнее охлаждение, ведущее к развитию мощного нисходящего потока [5, 6].

В этом контексте вертикальное распределение влажности в тропосфере приобретает двойственное значение. При анализе восходящих потоков (что составляет основу большинства конвективных прогностических схем) сухой воздух в средней тропосфере традиционно рассматривается как ингибитор конвекции. Однако с позиции нисходящих потоков тот же сухой слой выступает как катализатор: осадки, попадающие в него, интенсивно испаряются, охлаждают воздух и формируют холодный, плотный и тяжёлый «пузырь», стремительно опускающийся к поверхности. Это противоречие – когда одна и та же атмосферная конфигурация одновременно подавляет восходящий поток и усиливает нисходящий – является фундаментальным и не находит отражения в стандартных оперативных прогностических индексах.

Юг европейской части России (регион Предкавказья и Северного Кавказа) представляет особый интерес для изучения данного явления. Регион характеризуется высокой повторяемостью сухих вторжений из Средиземноморья и Каспия, которые формируют характерный «загруженный» тип зондирования – влажный пограничный слой, перекрытый мощной сухой воздушной массой на уровнях 850–500 гПа [7, 8]. При прорыве ингибирующего слоя возникающие конвективные системы нередко сопровождаются катастрофическими шквалами, которые оказываются совершенно неожиданными для стандартных прогностических систем, ориентированных на доступную конвективную потенциальную энергию – Convective Available Potential Energy (CAPE) как основной параметр.

Настоящая работа ставит следующие задачи: провести анализ аэрологических характеристик, предшествующих опасным конвективным шквалам на юге России; количественно сопоставить влажностные профили в дни с шквалами и контрольных случаях; апробировать новый диагностический параметр индекс нисходящей конвективной энергии с учетом охлаждения за счёт испарения осадков в сухом среднем тропосферном слое ($DCAPE_{dry}$), явно учитывающий дефицит влажности в среднем тропосферном слое; исследовать физические механизмы усиления нисходящих потоков в условиях сухих вторжений применительно к орографически сложному региону Предкавказья и Большого Кавказа.

1. Физика нисходящих конвективных потоков и роль влажности. 1.1. Механизмы формирования шквалов. Нисходящий конвективный поток формируется в результате суммарного действия трёх сил: отрицательной плавучести охлаждённого воздуха, силы тяжести, действующей на гидрометеоры, и возмущения давления. Основным термодинамическим параметром, характеризующим потенциал нисходящего потока, является Downdraft CAPE (DCAPE) – интеграл отрицательной плавучести воздушного объема при опускании от уровня насыщения нисходящего потока (DSLFF) до поверхности:

$$DCAPE = \int_{z_s}^0 g \frac{T_{ve}(z) - T_{vp}(z)}{T_{ve}(z)} dz,$$

где T_{ve} – виртуальная температура окружающей среды, T_{vp} – виртуальная температура опускающегося объема воздуха, z_s – высота уровня DSLFF. Опускающийся объем изначально насыщен, он охлаждается по влажноадиабатическому градиенту до уровня конденсации нисходящего потока (DCL), затем – по сухоадиабатическому. Чем суше исходная воздушная масса и соответственно ниже точка росы, тем выше расположен DCL, и воздух суше – охлаждается по более крутому градиенту при последующем опускании, и тем больше DCAPE [5, 9].

Ключевым дополнительным механизмом является испарительное охлаждение. Когда осадки (дождь или круп) попадают в слой сухого воздуха, значительная часть их массы переходит в пар. Согласно закону сохранения энергии, испарение 1 г воды погло-

щает примерно 2500 Дж теплоты, что соответствует охлаждению 1 кг воздуха приблизительно на 2,5°C. При интенсивных осадках (5–10 мм/ч) в очень сухом слое (дефицит точки росы $TD = 20\text{--}30^\circ\text{C}$) скорость испарительного охлаждения может достигать 5–10°C/км, создавая мощный источник отрицательной плавучести [6, 10]. Именно этот процесс превращает умеренный ливень в источник катастрофического шквала.

1.2. Сухое вторжение: концепция и региональная специфика. Сухое вторжение – это нисходящий поток сухого воздуха из верхней тропосферы или стратосферы, проникающий в среднюю тропосферу по потенциально-вихревым поверхностям вблизи вершин трогов и струйных течений [8, 11]. На синоптических диаграммах водяного пара (спутниковые каналы 6,2 и 7,3 мкм) сухие вторжения отчетливо видны как темные (холодные) языки, вторгающиеся в зону префронтальной влажной конвекции [11].

Для Предкавказья характерны два основных вида сухих вторжений:

- ♦ Средиземноморский тип – сухой тёплый воздух, связанный с фёновыми процессами над горными системами и субтропическими антициклонами над Средиземноморьем, формирующий мощные инверсионные слои на уровнях 700–600 гПа с $TD > 15\text{--}20^\circ\text{C}$.

- ♦ Каспийский (континентальный) тип – сухой и горячий воздух из аравийско-иранских пустынь, поступающий в регион с востока-юго-востока, характеризующийся экстремально низкой влажностью ($RH < 15\%$ на 700 гПа) и $TD > 25^\circ\text{C}$. Оба типа создают термодинамически «заряженную» атмосферу с высоким потенциалом катастрофических нисходящих потоков при прорыве конвекции.

1.3. Индекс DCAPE с учётом охлаждения за счёт испарения осадков в сухом среднем тропосферном слое ($DCAPE_{dry}$). Стандартный DCAPE в условиях сухого воздуха – это ключевой метеорологический показатель для прогнозирования сильных, разрушительных шквалов и сухих микропорывов, но он не в полной мере отражает роль сухих слоёв в усилении нисходящих потоков: он вычисляется для репрезентативного объема воздуха из заданного слоя, при этом не учитывая интегральный дефицит влажности по всей толщине слоя, через который проходят осадки. С целью устранения данного ограничения предложен модифицированный показатель $DCAPE_{dry}$:

$$DCAPE_{dry} = \int_{z_s}^0 g \frac{T_{ve}(z) - T_{vp}^{eff}(z)}{T_{ve}(z)} dz,$$

$$T_{vp}^{eff}(z) = T_{vp}(z) - \gamma \cdot \overline{\Delta TD}_{mid} \cdot \frac{T_{ve}(z)}{T_{ve,cp}},$$

его можно представить иначе

$$DCAPE_{dry} = \int_{z_s}^0 g \frac{T_{ve}(z) - T_{vp}(z)}{T_{ve}(z)} (1 + k \cdot \overline{\Delta TD}_{mid}) dz,$$

или

$$DCAPE_{dry} = DCAPE \cdot (1 + k \cdot \overline{\Delta TD}_{mid}).$$

где DCAPE – стандартное значение нисходящей конвективной потенциальной энергии (Дж/кг), $T_{vp}^{eff}(z)$ – эффективная виртуальная температура опускающейся частицы, γ – безразмерный коэффициент, $T_{ve,cp}$ – средняя виртуальная температура в слое, $\overline{\Delta TD}_{mid}$ – среднее значение дефицита точки росы в слое 500–700 гПа ($^\circ\text{C}$), $k = 0,042\ ^\circ\text{C}^{-1}$ – эмпирически подобранный коэффициент (его значение оптимизировалось методом логистической регрессии на обучающей выборке из 180 случаев за 1990–2010 гг. с последующей проверкой на независимой выборке 2011–2023 гг.). При $\overline{\Delta TD}_{mid} = 0$ формула вырождается в стандартный DCAPE. При $\overline{\Delta TD}_{mid} = 20^\circ\text{C}$ поправочный множитель составляет 1,84 – т.е. фактическая энергия нисходящего потока почти вдвое превышает оценку по стандартному DCAPE.

Физический смысл поправки: чем больше дефицит влажности в слое, через который проходят осадки, тем интенсивнее испарительное охлаждение, тем выше отрицательная плавучесть, не учитываемая стандартным DCAPE, который строится для изначально насыщенного объема воздуха.

2. Данные и методология. 2.1. Исходные данные. Аэрологическое зондирование. Использованы данные стандартного радиозондирования (сроки 00 и 12 UTC) с 11 станций юга России и Северного Кавказа: Ростов-на-Дону (WMO 34731), Краснодар (34941), Ставрополь (37228), Нальчик (37011), Минеральные Воды (37054), Грозный (37089), Элиста (34812), Волгоград (34731), Астрахань (35121), Анапа (33976), Сочи (37597) за период 1990–2023 гг. (всего 247 400 зондирований). Архив получен из глобального хранилища IGRA v2 (NOAA NCEI) с процедурой контроля качества (удаление зондирований с более чем 20% пропущенных данных в ключевых слоях).

Реанализ ERA5 использовался для восстановления полей, недоступных из дискретных зондирований: горизонтальный ветер, геопотенциал, поля CAPE, DCAPE, PWAT с разрешением $0,25^\circ \times 0,25^\circ$ (почасово). Точность ERA5 в воспроизведении DCAPE оценивалась сравнением с рассчитанными по зондированиям значениями; RMSE составил 89 Дж/кг (относительная ошибка ~18%), что сопоставимо с другими оценками для умеренных широт [12–14].

Архив опасных явлений. Случаи приземных шквалов конвективного происхождения (скорость ветра ≥ 25 м/с) идентифицированы на основе: базы данных опасных явлений Росгидромета за 1990–2023 гг.; архива ВГИ (Нальчик) [15]; данных сети метеостанций (автоматических и ручных наблюдений). Отбирались только случаи с подтверждённой конвективной природой шквала наличие грозы и/или крупы и с доступным радиозондом не позднее чем за 3 ч до события [16, 17]. Итоговая выборка: 268 шквалов. Контрольная выборка: 1042 дня с зарегистрированной грозой, но без приземного ветра ≥ 25 м/с.

2.2. Методология анализа. Климатологический анализ проводился по следующей схеме:

- ♦ Для каждого зондирования обеих выборок рассчитывался полный набор термодинамических индексов: CAPE (Surface-Based, SB-CAPE), CIN, LCL, LFC, DCAPE, К-индекс, дефицит точки росы на 700 гПа (TD700), средний дефицит точки росы в слоях 850–700 гПа и 700–500 гПа, высота нулевой изотермы (H_{FZ}), интегральное влагосодержание (PWAT).

- ♦ Вычислялась статистика распределений (медиана, квантили 25-й и 75-й, критерий Манна–Уитни для проверки статистической значимости различий). (3) Проводился корреляционный анализ (рангового Спирмена) между параметрами влажности и интенсивностью зафиксированных нисходящих потоков (максимальная приземная скорость ветра).

- ♦ Методом ROC-анализа оценивалось прогностическое качество различных индексов и пороговых критериев для обнаружения опасного шквала.

- ♦ Дополнительно построены составные (композиционные) карты полей ERA5 для сухих вторжений обоих типов.

3. Результаты. 3.1. Анализ термодинамических профилей. Сравнение медианных значений термодинамических параметров в выборках шквалов и контрольных дней выявило ряд принципиальных различий (табл. 1). Наиболее выраженное статистически значимое ($p < 0,001$ по критерию Манна–Уитни) различие зафиксировано именно в показателях влажности средней тропосферы. В частности, медианный дефицит точки росы на уровне 700 гПа составил 18,3°C для случаев со шквалом против 7,7°C в контроле – разница в 2,4 раза. Аналогичная картина наблюдается для среднего дефицита в слое 700–500 гПа (20,1 против 9,4°C).

Таблица 1

Медианные значения (квартили 25%–75%) термодинамических параметров в выборках шквалов и контрольных дней (11 станций, 1990–2023 гг.)

Параметр	Шквалы (N=268)	Контроль (N=1042)	p-знач. МУ
SB-CAPE, Дж/кг	1740 (1220–2410)	1890 (1340–2570)	0,087
CIN, Дж/кг	–58 (–95–28)	–31 (–52–14)	< 0,001
LCL, м	960 (780–1180)	880 (720–1080)	0,012
TD на 700 гПа, °C	18,3 (12,6–24,7)	7,7 (4,2–12,1)	< 0,001

Окончание табл. 1

Параметр	Шквалы (N=268)	Контроль (N=1042)	р-знач. МУ
TD ср. 700–500 гПа, °C	20,1 (15,4–26,2)	9,4 (5,8–14,3)	< 0,001
RH на 700 гПа, %	28 (18–38)	62 (48–73)	< 0,001
PWAT, мм	34,2 (28–41)	38,6 (33–45)	0,003
DCAPE, Дж/кг	680 (490–920)	290 (170–440)	< 0,001
DCAPE_{drv}, Дж/кг	1120 (740–1680)	340 (190–520)	< 0,001
Высота нулевой изот., м	2340 (1980–2710)	2180 (1840–2540)	0,021
К-индекс	24 (18–30)	31 (24–38)	< 0,001

МУ – тест Манна–Уитни для независимых выборок; TD – дефицит точки росы ($T - T_d$).

Показательно и то, что параметры энергии неустойчивости (SB-CAPE) и пограничного слоя (LCL, LFC) в двух выборках не демонстрируют статистически значимых различий: медианный CAPE для дней со шквалом (1740 Дж/кг) оказывается несколько ниже, чем в контроле (1890 Дж/кг), что на первый взгляд выглядит парадоксально. Однако этот результат полностью согласуется с концепцией «заряженной» атмосферы: высокий CAPE в условиях сухих вторжений реализуется во взрывную конвекцию, которая затем генерирует мощные нисходящие потоки за счёт испарительного охлаждения в том же сухом слое. Напротив, в дни с более влажной средней тропосферой CAPE реализуется более постепенно, а испарительное охлаждение под облаком незначительно.

Здесь следует отметить, что К-индекс, широко используемый синоптиками, для дней со шквалом ниже, чем в контроле (24 против 31), поскольку высокий дефицит точки росы на 700 гПа входит в формулу К-индекса со знаком минус, именно низкий К-индекс в данном контексте указывает на высокий потенциал опасных нисходящих потоков.

3.2. Стратификация случаев по типу сухого вторжения. Из 268 случаев шквалов 147 (54,9%) были классифицированы как связанные со Средиземноморским типом сухих вторжений (на основе ERA5-траекторного анализа и спутниковых карт водяного пара), 96 (35,8%) – с Каспийским (континентальным) типом, 25 (9,3%) – с вторжениями иного происхождения (арктические, местные термически генерируемые). Наиболее высокая максимальная скорость приземного ветра зафиксирована при Каспийском типе (медиана 34,2 м/с), что объясняется экстремально низкой относительной влажностью воздуха (медиана RH на 700 гПа = 12%) и соответственно максимальным испарительным охлаждением.

Таблица 2

Основные характеристики шквалов по типу сухого вторжения

Характеристика	Средиземноморский (N=147)	Каспийский (N=96)	Иные (N=25)
Медиана скорости ветра, м/с	28,4	34,2	26,1
TD на 700 гПа, °C	16,8	23,2	11,4
RH на 700 гПа, %	31	12	48
DCAPE_{drv}, Дж/кг	1050	1480	620
CAPE, Дж/кг	1680	1560	1920
Повторяемость в тёплый сезон, %	61	83	44
Время суток: доля ночных случаев, %	24	38	29

3.3. Корреляционный анализ. Матрица ранговых корреляций (Спирмена) между основными термодинамическими параметрами и максимальной зафиксированной скоростью приземного ветра (на основе объединённой выборки 268 шквалов) представлена в табл. 3. Наиболее высокий коэффициент корреляции с интенсивностью шквала демонстрирует $DCAPE_{dry}$ ($\rho = 0,74$), тогда как стандартный $DCAPE$ ($\rho = 0,61$) и TD на 700 гПа ($\rho = 0,58$) также статистически значимы, но менее информативны. Стандартный $CAPE$ практически не коррелирует с интенсивностью приземного ветра ($\rho = 0,09$, $p = 0,14$).

Таблица 3

Ранговые корреляции Спирмена (ρ) между термодинамическими параметрами и максимальной скоростью приземного ветра при шквале ($N=268$)

Параметр	ρ	p -значение
$DCAPE_{dry}$	0,74	< 0,001
$DCAPE$	0,61	< 0,001
TD на 700 гПа (дефицит т. р.)	0,58	< 0,001
TD ср. 700–500 гПа	0,55	< 0,001
RH на 700 гПа (обратная)	–0,52	< 0,001
Высота нулевой изотермы	0,38	< 0,001
$PWAT$ (обратная)	–0,31	< 0,001
$SB-CAPE$	0,09	0,14
K -индекс (обратная)	–0,44	< 0,001

Курсивом выделены корреляции с $p \geq 0,05$ (статистически незначимые); все остальные значимы при уровне 0,001.

3.4. Прогностическое качество индексов: ROC-анализ. Оценка качества прогноза различных индексов при разграничении случаев со шквалом и контрольных дней выполнена методом ROC-анализа. Для каждого индекса варьировался пороговый критерий, и строилась зависимость вероятности обнаружения ($POD = TP/(TP+FN)$) от доли ложных тревог ($FAR = FP/(FP+TN)$). Площадь под кривой ROC (AUC) использовалась как интегральная мера прогностической способности.

Таблица 4

Качество прогноза диагностических параметров при разграничении шквалов (≥ 25 м/с) и контрольных дней на независимой выборке 2011–2023 гг.

Параметр / критерий	Порог	POD , %	FAR , %	CSI , %	AUC
$SB-CAPE > 1500$ Дж/кг	>1500	52	51	34	0,56
K -индекс < 20	<20	58	47	38	0,63
$DCAPE > 400$ Дж/кг	>400	61	43	42	0,68
$TD_{700} > 12^\circ C$	>12	69	36	49	0,74
$DCAPE + TD_{700}$ (линейн.)	—	76	31	57	0,81
$DCAPE_{dry} \geq 620$ Дж/кг	≥ 620	84	19	70	0,89

POD = True Positive Rate; FAR = False Alarm Rate; CSI = Critical Success Index = $TP/(TP+FP+FN)$.

ДСАРЕ_{dry} с порогом ≥ 620 Дж/кг обеспечивает $AUC = 0,89$ – наилучший результат среди протестированных параметров. Для сравнения: стандартный ДСАРЕ с оптимальным порогом 400 Дж/кг даёт $AUC = 0,68$ (разность статистически значима при $p < 0,01$ по тесту ДеЛонга). Комбинированный линейный показатель (ДСАРЕ + TD₇₀₀) занимает промежуточное положение ($AUC = 0,81$), что подтверждает нелинейный, мультипликативный характер взаимодействия между потенциалом нисходящего потока и влажностью среды, адекватно воспроизводимый формулой ДСАРЕ_{dry}.

3.5. Пространственная изменчивость: роль орографии Предкавказья. Составные карты полей ERA5 (из 147 Средиземноморских случаев) на уровне 700 гПа демонстрируют устойчивый пространственный паттерн: ось максимального дефицита точки росы ($TD > 18^\circ\text{C}$) ориентирована с запада-северо-запада на восток-юго-восток, с максимальными значениями над Ставропольской возвышенностью и Кабардино-Балкарией. Это объясняется орографическим сжатием сухой воздушной массы над понижениями рельефа при сохранении её в верхних уровнях, а также нагревом воздуха при опускании – фёны на подветренных склонах Кавказа.

Данная структура приводит к пространственной концентрации наиболее опасных нисходящих потоков в «зоне аномалии» – вдоль полосы от Нальчика до Ставрополя и далее к Волгоградской области. Анализ архива шквалов подтверждает: 62% случаев Каспийского типа с наибольшей интенсивностью наблюдались именно в этой полосе. Это пространственное совпадение зон максимального дефицита влажности на 700 гПа и зон наибольшей повторяемости сильных шквалов является важным для регионального прогнозирования опасных явлений.

4. Физические механизмы усиления нисходящих потоков в условиях сухих вторжений. 4.1. Испарительное охлаждение: расчётные оценки. Для количественной оценки вклада испарительного охлаждения в развитие нисходящего потока проведены простые термодинамические расчёты при параметрах, характерных для Каспийского типа вторжений: интенсивность осадков 10 мм/ч, толщина сухого слоя 3 км, начальная $TD = 25^\circ\text{C}$ на уровне 700 гПа.

При данных условиях скорость испарительного охлаждения составляет около $6\text{--}8^\circ\text{C}/\text{км}$. Вертикальный интеграл дополнительного охлаждения по слою 3 км даёт около $20^\circ\text{C}\cdot\text{км}$, что при плотности воздуха $0,8 \text{ кг}/\text{м}^3$ и $C_p \approx 1005 \text{ Дж}/(\text{кг}\cdot^\circ\text{C})$ соответствует дополнительной отрицательной плавучести порядка 200–250 Дж/кг. Именно эта величина составляет разность между стандартным ДСАРЕ ($\sim 500\text{--}700$ Дж/кг при умеренной влажности) и ДСАРЕ_{dry} ($\sim 900\text{--}1200$ Дж/кг при экстремально сухом воздухе), что хорошо согласуется со статистическими данными из табл. 1 и объясняет физический смысл поправочного коэффициента k .

Результирующая скорость нисходящего потока у поверхности оценивается из соотношения $w_d \approx \sqrt{2 \cdot \text{ДСАРЕ}}$. Для стандартного ДСАРЕ = 600 Дж/кг это даёт $w_d \approx 35 \text{ м}/\text{с}$; для ДСАРЕ_{dry} = 1100 Дж/кг — $w_d \approx 47 \text{ м}/\text{с}$. Разница в 12 м/с вполне объясняет наблюдаемый в архиве разброс скоростей приземного ветра между «обычными» и наиболее опасными шквалами при схожих значениях ДСАРЕ.

4.2. Механизм генерации приземного вихря. Особую опасность нисходящие потоки представляют не только по скорости прямого удара о поверхность, но и через генерацию горизонтального вихря вблизи земли. При растекании холодного нисходящего потока по подстилающей поверхности формируется граница с сильным горизонтальным градиентом ветра и температуры. Вертикальное перемещение горизонтального вихря, ассоциированного с этой границей, может приводить к образованию смерча или, в менее выраженном виде, к генерации локальных вращательных движений, усиливающих разрушительное воздействие шквала.

Интенсивность генерации горизонтального вихря у поверхности пропорциональна горизонтальному градиенту плавучести. При более мощном испарительном охлаждении градиент плавучести между нисходящим потоком и окружающей тёплой воздушной массой выше, что усиливает и горизонтальный вихрь [18]. Данный механизм объясняет нередко наблюдаемую при шквалах Каспийского типа вращательную структуру приземного ветра, фиксируемую доплеровскими радарными.

4.3. Взаимодействие с орографией. В условиях Предкавказья нисходящие потоки дополнительно усиливаются орографически. Существуют три механизма. Первый: долинный фокусирующий эффект. При растекании холодного воздуха от центра конвективного процесса он нередко оказывается в долинах рек (Кубань, Кума, Терек, Сунжа), которые действуют как каналы, ускоряя поток и концентрируя его в узкое русло. Скорость ветра при выходе из ущелья может значительно превышать скорость на открытой равнине. Второй: горные фёны. В зонах с подветренными склонами (например, к северу от Эльбруса) нисходящий поток «суммируется» с местной фёновой циркуляцией, создавая исключительно сильные порывы. Третий: конвергентные зоны. На границах между горными и равнинными участками возникают зоны конвергенции потоков, усиливающие горизонтальные градиенты скорости ветра и его разрушительный потенциал [7, 15].

5. Сезонная и суточная изменчивость. 5.1. Сезонное распределение. Из 268 верифицированных случаев шквалов 91% приходится на тёплый период (апрель – сентябрь) с максимумом в июне – июле (48% всех случаев). Это соответствует максимальному проникновению сухих воздушных масс в среднюю тропосферу в летние месяцы и максимальной конвективной активности. Холодный период (октябрь – март) даёт лишь 9% случаев, причём они практически исключительно связаны со Средиземноморским типом вторжений.

Сезонная динамика $DCAPE_{dry}$ следует сезону конвективной активности: медиана растёт от 380 Дж/кг в апреле до 1380 Дж/кг в июле, затем снижается. Параллельно увеличивается TD на 700 гПа: от 9°C (апрель) до 22°C (июль), что отражает снижение влаги в континентальных воздушных массах в разгар лета. Это означает, что наиболее опасный период для нисходящих потоков данного типа июнь – август.

Таблица 5

Сезонная статистика $DCAPE_{dry}$ и дефицита точки росы TD_{700} в выборке шквалов (N=268)

Месяц	N случаев	Медиана $DCAPE_{dry}$, Дж/кг	Медиана TD_{700} , °C	Медиана V_{max} , м/с
Апрель	14	380	9,2	26,4
Май	29	680	13,8	27,1
Июнь	68	1180	20,4	31,2
Июль	61	1380	22,1	34,6
Август	52	1120	19,8	30,8
Сентябрь	24	640	12,1	27,3
Окт.–Март	20	310	8,8	25,7

5.2. Суточный ход. Суточный ход шквалов демонстрирует два пика: основной – в период 15:00–19:00 местного времени (52% случаев), связанный с максимумом дневного конвективного прогрева, и вторичный – в 22:00–02:00 (21% случаев), объясняемый развитием ночных мезомасштабных конвективных систем (МКС) и усилением низкоуровневых струйных потоков над сушей (Low-Level Jet, LLJ). Интересно, что ночные случаи (21% от общей выборки) характеризуются медианной $DCAPE_{dry}$ лишь незначительно ниже (970 против 1140 Дж/кг), что указывает на то, что ночные МКС в регионе нередко развиваются в обстановке сухих вторжений с высоким потенциалом нисходящих потоков, подпитываемых ночными LLJ.

6. Типичные случаи: синоптический анализ. 6.1. Случай 1: Шквал Каспийского типа, 15 июля 2021 г. Во второй половине дня 15 июля 2021 г. на территории Кабардино-Балкарии и Ставропольского края зафиксированы шквалы со скоростью ветра 38–42 м/с,

сопровождаясь уничтожением посевов и разрушением лёгких строений. К 12 UTC по данным зондирования Нальчик (12 UTC): SB-CAPE = 1480 Дж/кг, CIN = -72 Дж/кг, LCL = 1040 м, TD на 700 гПа = 28,4°C, RH на 700 гПа = 8%, DCAPE = 910 Дж/кг, DCAPE_{dry} = 1980 Дж/кг.

Карты реанализа ERA5 подтверждали присутствие мощного сухого вторжения Каспийского типа: суховей с RH < 10% на 700 гПа охватывал регион с востока. Зондирование имело классический вид: ниже 850 гПа – влажный (TD < 5°C), выше – крайне сухой воздух TD = 28–30°C. К-индекс = 16 легко пропустить угрозу. Однако DCAPE_{dry} = 1980 Дж/кг однозначно указывал на катастрофический потенциал нисходящих потоков. Прогноз на основе нового индекса, верно, идентифицировал ситуацию как чрезвычайно опасную.

6.2. Случай 2: Вводящий в заблуждение высокий CAPE без шквала, 22 июня 2019 г. Этот «контрпример» не менее ценен. 22 июня 2019 г. зондирование Ставрополь (12 UTC) показало: SB-CAPE = 3240 Дж/кг (максимально высокое значение), CIN = -15 Дж/кг (очень низкое подавление), LCL = 820 м. Казалось бы, идеальные условия для мощных гроз. Действительно, развились обширные грозовые системы. Однако шквалов ≥ 25 м/с зафиксировано не было: максимальная скорость ветра составила 12 м/с.

Дефицит точки росы TD на 700 гПа = 3,2°C это значит, что воздух на средних уровнях был практически насыщен влагой RH $\approx$ 88%. Испарительное охлаждение осадков было минимальным, что в сочетании с небольшим DCAPE = 180 Дж/кг и соответственно DCAPE_{dry} = 193 Дж/кг однозначно указывало на слабый потенциал нисходящих потоков. Несмотря на огромный CAPE, при котором все стандартные прогностические схемы предупреждали бы о «сильных грозах», однако реального ветрового ущерба не произошло. DCAPE_{dry} корректно обозначил ситуацию как «незначимую» с точки зрения шквала (193 Дж/кг << 620 Дж/кг). Данный случай демонстрирует, почему CAPE сам по себе не может служить достаточным предиктором приземного ветра.

7. Обсуждение. 7.1. Парадокс «сухой атмосферы» и оперативный прогноз. Результаты настоящего исследования ставят под сомнение распространённую в оперативной практике логику: «высокий CAPE – сильные грозы со шквалами». Применительно к нисходящим конвективным потокам это утверждение не только не работает, но может вводить в заблуждение. Как показано выше, «идеально неустойчивая» и влажная атмосфера с CAPE > 3000 Дж/кг генерирует грозы, но не шквалы, тогда как «умеренно неустойчивая» и крайне сухая атмосфера с CAPE $\sim$ 1500 Дж/кг и TD₇₀₀ > 25°C может породить катастрофический шквал.

Этот парадокс объясним с позиций физики. Высокая влажность во всей толще тропосферы означает, что испарительное охлаждение минимально, нисходящие потоки слабые, и весь CAPE реализуется в мощные восходящие потоки (ливни, грозы, крупный град). Высокая CAPE при сухой среде на среднем уровне в тропосфере реализуется в более умеренные восходящие потоки так как часть энергии теряется на вовлечение, но зато нисходящие потоки аномально мощны. Таким образом, профиль влаги определяет предпочтительный канал реализации атмосферной неустойчивости.

Таким образом, анализ угрозы опасного ветра должен начинаться с проверки влажностного профиля на средних уровнях. Предложенный индекс DCAPE_{dry} формализует алгоритм и даёт единственный легко вычисляемый параметр.

7.2. Сравнение с исследованиями для других регионов. Полученные результаты хорошо согласуются с международными данными. В архиве смерчопасных случаев США [19] также установлена слабая корреляция CAPE с интенсивностью приземного ветра при высокая корреляция дефицита точки росы в слое 700–500 гПа. Прогностические системы Национальной метеорологической службы США (NWS) в зонах сухого климата (юго-западные штаты) давно применяют дефицит точки росы на 700 гПа в качестве основного параметра прогноза «dry microbursts» [20, 21]. Различие состоит лишь в характере фонового климата: если в засушливых американских штатах сухая средняя тропосфера является нормой, на юге России она носит эпизодический характер, что требует специфического регионального анализа, проведённого в данной работе.

Исследования для Европы [22] демонстрируют аналогичную роль средиземноморских сухих вторжений в формировании опасных конвективных систем над Балканами и Центральной Европой. Предложенная концепция DCAPE_{dry} в принципе применима и для этих регионов, хотя калибровка коэффициента k может потребовать существенной адаптации к региону.

7.3. Ограничения и направления дальнейших исследований. Основным ограничением настоящей работы является опора на данные радиозондирования в стандартные синоптические сроки (00 и 12 UTC), которые могут не соответствовать моменту шквала. Учитывая суточный максимум конвективных событий в 15–19 ч местного времени, реальные профили влажности во время шквала могут отличаться от зондирований за 3–6 ч до события. Это источник случайной погрешности, снижающей возможно реальную прогностическую способность DCAPE_{dry}.

Следующим шагом является валидация индекса в конвективно-разрешающих модельных прогнозах (WRF с шагом 1 км) и в почасовых продуктах ERA5. Перспективным направлением представляется также включение дополнительных параметров – высоты нулевой изотермы (определяющей расстояние, на котором осадки начинают испаряться) и горизонтального сдвига ветра (влияющего на интенсивность линии шквалов) – в многопараметрический индекс с применением методов машинного обучения.

Заключение. На основе 34-летнего архива аэрологических данных, реанализа ERA5 и спутниковых профилей влажности для 11 станций юга России проведён систематический анализ роли сухих вторжений в средней тропосфере в формировании опасных нисходящих конвективных потоков. Полученные результаты позволяют сформулировать следующие основные выводы:

- ♦ Стандартный CAPE является неудовлетворительным предиктором интенсивности приземного шквала конвективного происхождения: коэффициент ранговой корреляции между CAPE и максимальной скоростью ветра составляет $\rho = 0,09$ (статистически незначимо), тогда как для DCAPE_{dry} $\rho = 0,74$ ($p < 0,001$). Принципиальным диагностическим параметром является дефицит точки росы в слое 700–500 гПа: в дни со шквалами он в 2,4 раза выше, чем в контрольных грозовых ситуациях (18,3 против 7,7°C на 700 гПа).

- ♦ Разработан новый интегральный индекс $DCAPE_{dry} = DCAPE \cdot (1 + 0,042 \cdot \overline{\Delta T D_{mid}})$, физически обоснованный как нелинейная комбинация потенциала нисходящего потока и интенсивности испарительного охлаждения в сухом слое. При пороге ≥ 620 Дж/кг индекс обеспечивает POD = 84%, FAR = 19%, CSI = 70%, AUC = 0,89 в задаче диагностики шквала, превосходя стандартный DCAPE (AUC = 0,68) и K-индекс (AUC = 0,63).

- ♦ Идентифицированы два основных типа сухих вторжений, наиболее значимых для региона: Средиземноморский (54,9% случаев, медиана Vmax = 28,4 м/с) и Каспийский (35,8% случаев, медиана Vmax = 34,2 м/с). Каспийский тип, характеризующийся экстремально низкой RH на 700 гПа (медиана 12%), порождает наиболее интенсивные шквалы и сопровождается наибольшими значениями DCAPE_{dry}.

- ♦ Орографические особенности Предкавказья (долинный фокусирующий эффект, фёновые зоны, конвергентные зоны на границе гор и равнин) систематически усиливают интенсивность нисходящих потоков, способствуя концентрации наиболее опасных шквалов в полосе от Нальчика до Ставрополя.

Внедрение DCAPE_{dry} в оперативную практику аэрологической диагностики позволит существенно улучшить заблаговременность и достоверность предупреждений об опасных конвективных явлениях в регионе, особенно в условиях «сухой конвекции» с высоким CIN и умеренным CAPE, которые стандартные схемы систематически недооценивают.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Чернокульский А.В., Шихов А.Н., Ажигов И.О., Ерошкина Н.А., Корнев Д.П., Быков А.В., Калинин Н.А., Курганский М.В., Павлюков Ю.Б., Спрыгин А.А., Ярынич Ю.И. Шквалы и смерчи на европейской части России 15 мая 2021 г.: диагностика и моделирование // Метеорология и гидрология. – 2022. – № 11. – С. 71-90.

2. Proctor F.H. Numerical simulations of an isolated microburst. Part I: Dynamics and structure // Journal of the Atmospheric Sciences. – 1988. – Vol. 45, No. 21. – P. 3137-3160. – doi: 10.1175/1520-0469(1988)045<3137: NSOAIM>2.0.CO;2.
3. Cholette M., Milbrandt J.A., Morrison H., Paquin-Ricard D., Jacques D. Combining Triple-Moment Ice With Prognostic Liquid Fraction in the P3 Microphysics Scheme: Impacts on a Simulated Squall Line // Journal of Advances in Modeling Earth Systems. – 2023. – Vol. 15, Issue 4. – e2022MS003328. – doi: 10.1029/2022MS003328.
4. Sun Y., Gao Q., Li H., Wang M. Evaluating Simulated Microphysics of Stratiform and Convective Precipitation in a Squall Line Event Using Polarimetric Radar Observations // Remote Sensing. – 2023. – Vol. 15, Issue 6. – P. 1507. – doi: 10.3390/rs15061507.
5. Srivastava R.C. A Model of intense Downdrafts Driven by the Melting and Evaporation of Precipitation // Journal of the Atmospheric Sciences. – 1987. – Vol. 44, No. 13. – P. 1752-1774. – doi: 10.1175/1520-0469(1987)044<1752: AMOIDD>2.0.CO;2.
6. Lin J., Qian T., Bechtold P., Grell G., Zhang G. J., Zhu P., Freitas S.R., Barnes H., Han J. Atmospheric convection // Atmosphere-Ocean. – 2022. – Vol. 60, No. 3-4. – P. 422-476. – doi: 10.1080/07055900.2022.2082915.
7. Ларина О.Г., Бадахова Г.Х. Эволюция и траектории перемещения южных циклонов и их фронтальных разделов в условиях Северного Кавказа // Академическая публицистика. – 2022. – № 6-1. – С. 621-632.
8. Наурызбаева Ж.К., Холопцев А.В., Кенжина Ж.Х., Рахматулла Н.Е. Моделирование межгодовых изменений суммарной продолжительности гроз северной части Прикаспия, с учетом вариаций атмосферных осадков // Гидрометеорология и экология. – 2025. – № 4 (119). – С. 159-176.
9. Prein A.F. Thunderstorm straight line winds intensify with climate change // Nature Climate Change. – 2023. – Vol. 13, No. 12. – P. 1353-1359.
10. Wakimoto R.M. Convectively driven high wind events // Severe Convective Storms / ed. C.A. Doswell III. – Boston: AMS, 2001. – P. 255-298.
11. Hoskins B.J., McIntyre M.E., Robertson A.W. On the use and significance of isentropic potential vorticity maps // Quarterly Journal of the Royal Meteorological Society. – 1985. – Vol. 111, No. 470. – P. 877-946. – doi: 10.1002/qj.49711147002.
12. Taszarek M., Brooks H.E., Czernecki B. Sounding-derived parameters associated with convective hazards in Europe // Monthly Weather Review. – 2017. – Vol. 145, No. 4. – P. 1507-1524. – doi: 10.1175/MWR-D-16-0384.1.
13. Susskind J., Barnett C.D., Blaisdell J.M. Retrieval of atmospheric and surface parameters from AIRS/AMSU/HSB data in the presence of clouds // IEEE Transactions on Geoscience and Remote Sensing. – 2003. – Vol. 41, No. 2. – P. 390-409. – doi: 10.1109/TGRS.2002.808236.
14. Pougetchev N., August T., Calbet X. et al. IASI temperature and water vapor retrievals – error assessment and validation // Atmospheric Chemistry and Physics. – 2009. – Vol. 9, No. 17. – P. 6453-6458. – doi: 10.5194/acp-9-6453-2009.
15. Ташилова А.А., Кешева Л.А., Теунова Н.В. Уменьшение водных ресурсов региона как следствие деградации ледников Северного Кавказа под воздействием климатических изменений // Международный научно-исследовательский журнал. – 2023. – № 1 (127). – С. 93-98.
16. Adzhieva A.A., Shapovalov V.A., Boldyreff A.S. Development of thunderstorm monitoring technologies and algorithms by integration of radar, sensors and satellite images // Proc. SPIE 10424, Remote Sensing of Clouds and the Atmosphere XXII. – 2017. – 104240H (20 October 2017). – doi: 10.1117/12.2299289.
17. Adzhiev A.K., Adzhieva A.A. Spatial and temporal variations of thunderstorm activity in the Northern Caucasus // Russian Meteorology and Hydrology. – 2009. – Vol. 34, No. 12. – P. 789-793.
18. Rotunno R., Klemp J.B., Weisman M.L. A theory for strong, long-lived squall lines // Journal of the Atmospheric Sciences. – 1988. – Vol. 45, No. 3. – P. 463-485. – DOI: 10.1175/1520-0469(1988)045<0463: ATFSLL>2.0.CO;2.
19. Romanic D., Taszarek M., Brooks H. Convective environments leading to microburst, macroburst and downburst events across the United States // Weather and Climate Extremes. – 2022. – Vol. 37. – P. 100474.
20. Atkins N.T., Wakimoto R.M. Wet microburst activity over the southeastern United States: Implications for forecasting // Weather and Forecasting. – 1991. – Vol. 6, No. 4. – P. 470-482. – DOI: 10.1175/1520-0434(1991)006<0470: WMAOTS>2.0.CO;2.
21. Williams I.N., Fieweger G.N. Mechanisms of projected changes in thunderstorm downburst environments across the United States // Geophysical Research Letters. – 2026. – Vol. 53, No. 1. – P. e2025GL117925.
22. Púčik T., Groenemeijer P., Rýva D., Kolář M. Proximity soundings of severe and nonsevere thunderstorms in Central Europe // Monthly Weather Review. – 2015. – Vol. 143, No. 12. – P. 4805-4821. – DOI: 10.1175/MWR-D-15-0008.1.

REFERENCES

1. Chernokul'skiy A.V., Shikhov A.N., Azhigov I.O., Eroshkina N.A., Korenev D.P., Bykov A.V., Kalinin N.A., Kurganskiy M.V., Pavlyukov Yu.B., Sprygin A.A., Yarynich Yu.I. Shkvaly i smerchi na evropeyskoy chasti Rossii 15 maya 2021 g.: diagnostika i modelirovanie [Shkansky shafts and tornadoes in the European part of Russia on May 15, 2021: diagnostics and modeling], *Meteorologiya i gidrologiya* [Meteorology and Hydrology], 2022, No. 11, pp. 71-90.
2. Proctor F.H. Numerical simulations of an isolated microburst. Part I: Dynamics and structure, *Journal of the Atmospheric Sciences*, 1988, Vol. 45, No. 21, pp. 3137-3160. doi: 10.1175/1520-0469(1988)045<3137: NSOAIM>2.0.CO;2.
3. Cholette M., Milbrandt J.A., Morrison H., Paquin-Ricard D., Jacques D. Combining Triple-Moment Ice With Prognostic Liquid Fraction in the P3 Microphysics Scheme: Impacts on a Simulated Squall Line, *Journal of Advances in Modeling Earth Systems*, 2023, Vol. 15, Issue 4. e2022MS003328. doi: 10.1029/2022MS003328.
4. Sun Y., Gao Q., Li H., Wang M. Evaluating Simulated Microphysics of Stratiform and Convective Precipitation in a Squall Line Event Using Polarimetric Radar Observations, *Remote Sensing*, 2023, Vol. 15, Issue 6, pp. 1507. doi: 10.3390/rs15061507.
5. Srivastava R.C. A Model of intense Downdrafts Driven by the Melting and Evaporation of Precipitation, *Journal of the Atmospheric Sciences*, 1987, Vol. 44, No. 13, pp. 1752-1774. doi: 10.1175/1520-0469(1987)044<1752: AMOIDD>2.0.CO;2.
6. Lin J., Qian T., Bechtold P., Grell G., Zhang G. J., Zhu P., Freitas S.R., Barnes H., Han J. Atmospheric convection, *Atmosphere-Ocean*, 2022, Vol. 60, No. 3-4, pp. 422-476. doi: 10.1080/07055900.2022.2082915.
7. Larina O.G., Badakhova G.Kh. Evolyutsiya i traektorii peremeshcheniya yuzhnykh tsiklonov i ikh frontal'nykh razdelov v usloviyakh Severnogo Kavkaza [Evolution and trajectories of the movement of southern cyclones and their frontal sections in the conditions of the North Caucasus], *Akademicheskaya publitsistika* [Academic Journalism], 2022, No. 6-1, pp. 621-632.
8. Naurozbaeva Zh.K., Kholoptsev A.V., Kenzhina Zh.Kh., Rakhmatulla N.E. Modelirovanie mezhdugodovykh izmeneniy summarnoy prodolzhitel'nosti groz severnoy chasti Prikaspiya, s uchedom variatsiy atmosferykh osadkov [Modeling of interannual changes in the total duration of thunderstorms in the northern part of the Caspian Sea, taking into account variations in precipitation], *Gidrometeorologiya i ekologiya* [Hydrometeorology and Ecology], 2025, No. 4 (119), pp. 159-176.
9. Prein A.F. Thunderstorm straight line winds intensify with climate change, *Nature Climate Change*, 2023, Vol. 13, No. 12, pp. 1353-1359.
10. Wakimoto R.M. Convectively driven high wind events, *Severe Convective Storms*, ed. C.A. Doswell III. Boston: AMS, 2001, pp. 255-298.
11. Hoskins B.J., McIntyre M.E., Robertson A.W. On the use and significance of isentropic potential vorticity maps, *Quarterly Journal of the Royal Meteorological Society*, 1985, Vol. 111, No. 470, pp. 877-946. doi: 10.1002/qj.49711147002.
12. Taszarek M., Brooks H.E., Czernecki B. Sounding-derived parameters associated with convective hazards in Europe, *Monthly Weather Review*, 2017, Vol. 145, No. 4, pp. 1507-1524. doi: 10.1175/MWR-D-16-0384.1.
13. Susskind J., Barnett C.D., Blaisdell J.M. Retrieval of atmospheric and surface parameters from AIRS/AMSU/HSB data in the presence of clouds, *IEEE Transactions on Geoscience and Remote Sensing*, 2003, Vol. 41, No. 2, pp. 390-409. doi: 10.1109/TGRS.2002.808236.
14. Pougetchev N., August T., Calbet X. et al. IASI temperature and water vapor retrievals – error assessment and validation, *Atmospheric Chemistry and Physics*, 2009, Vol. 9, No. 17, pp. 6453-6458. doi: 10.5194/acp-9-6453-2009.
15. Tashilova A.A., Kesheva L.A., Teunova N.V. Umen'shenie vodnykh resursov regiona kak sledstvie degradatsii lednikov Severnogo Kavkaza pod vozdeystviem klimaticheskikh izmeneniy [Reduction of water resources of the region as a result of degradation of glaciers of the North Caucasus under the influence of climatic changes], *Mezhdunarodnyy nauchno-issledovatel'skiy zhurnal* [International Research Journal], 2023, No. 1 (127), pp. 93-98.
16. Adzhieva A.A., Shapovalov V.A., Boldyreff A.S. Development of thunderstorm monitoring technologies and algorithms by integration of radar, sensors and satellite images, *Proc. SPIE 10424, Remote Sensing of Clouds and the Atmosphere XXII*, 2017, 104240H (20 October 2017). doi: 10.1117/12.2299289.
17. Adzhiev A.K., Adzhieva A.A. Spatial and temporal variations of thunderstorm activity in the Northern Caucasus, *Russian Meteorology and Hydrology*, 2009, Vol. 34, No. 12, pp. 789-793.
18. Rotunno R., Klemp J.B., Weisman M.L. A theory for strong, long-lived squall lines, *Journal of the Atmospheric Sciences*, 1988, Vol. 45, No. 3, pp. 463-485. DOI: 10.1175/1520-0469(1988)045<0463: ATFSLL>2.0.CO;2.

19. Romanic D., Taszarek M., Brooks H. Convective environments leading to microburst, macroburst and downburst events across the United States, *Weather and Climate Extremes*, 2022, Vol. 37, pp. 100474.
20. Atkins N.T., Wakimoto R.M. Wet microburst activity over the southeastern United States: Implications for forecasting, *Weather and Forecasting*, 1991, Vol. 6, No. 4, pp. 470-482. DOI: 10.1175/1520-0434(1991)006<0470: WMAOTS>2.0.CO;2.
21. Williams I.N., Fieweger G.N. Mechanisms of projected changes in thunderstorm downburst environments across the United States, *Geophysical Research Letters*, 2026, Vol. 53, No. 1, pp. e2025GL117925.
22. Púčik T., Groenemeijer P., Rýva D., Kolář M. Proximity soundings of severe and nonsevere thunderstorms in Central Europe, *Monthly Weather Review*, 2015, Vol. 143, No. 12, pp. 4805-4821. DOI: 10.1175/MWR-D-15-0008.1.

Эндреев Эльяс Маликович – Высокогорный геофизический институт; e-mail: eem507@bk.ru; г. Нальчик, Россия; тел.: 89287064453; аспирант.

Шаповалов Виталий Александрович – Высокогорный геофизический институт; e-mail: vet555_83@mail.ru; г. Нальчик, Россия; тел.: 89280758034; с.н.с.; д.ф.-м.н.

Аджиева Аида Анатольевна – Кабардино-Балкарский государственный аграрный университет имени В.М. Кокова; e-mail: aida-adzhieva@mail.ru; тел.: 89604279686; г. Нальчик, Россия; кафедра высшей математики и информатики; д.ф.-м.н.; профессор.

Куповых Геннадий Владимирович – Южный федеральный университет; e-mail: kupovykh@sfedu.ru; г. Таганрог, Россия; тел.: +78634371636; кафедра физико-математических основ инженерного образования; зав. кафедрой; д.ф.-м.н.; профессор.

Гривцов Владимир Владиславович – Южный федеральный университет; e-mail: gvv@sfedu.ru; г. Таганрог, Россия; тел.: +78634371794; кафедра инженерной графики и компьютерного дизайна; к.т.н.; доцент.

Andreyev Elias Malikhovich – Vysokogorny Geophysical Institute; e-mail: eem507@bk.ru; Nalchik, Russia; phone: +79287064453; postgraduate student.

Shapovalov Vitaly Aleksandrovich – Vysokogorny Geophysical Institute; e-mail: vet555_83@mail.ru; Nalchik, Russia; phone: +79280758034; dr. of phis. and math. sc.; senior researcher.

Adzhieva Aida Anatolyevna – V.M. Kokov Kabardino-Balkarian State Agrarian University, Nalchik, KBR, Russia; e-mail: aida-adzhieva@mail.ru; phone: +79604279686; the Department of Higher Mathematics and Computer Science; dr. of phis. and math. sc.; professor.

Kupovkh Gennady Vladimirovich – Southern Federal University; e-mail: kupovykh@sfedu.ru; Taganrog, Russia; phone: +78634371636; the Department of Physical and Mathematical Foundations of Engineering Education; head of department; dr. of phis. and math. sc.; professor.

Grivtsov Vladimir Vladislavovich – Southern Federal University; e-mail: gvv@sfedu.ru; Taganrog, Russia; phone: +78634371794; the Department of Engineering Graphics and Computer Design; cand. of eng. sc.; associate professor.

УДК 004.021

DOI 10.18522/2311-3103-2026-4-272-282

А.В. Боженюк, И.А. Дубчак, О.В. Косенко

ПРИМЕНЕНИЕ НЕЧЕТКОГО МЕТОДА TOPSIS ДЛЯ ПРИНЯТИЯ РЕШЕНИЙ С ИСПОЛЬЗОВАНИЕМ СИНТАКСИЧЕСКИ НЕЗАВИСИМЫХ ЛИНГВИСТИЧЕСКИХ ПЕРЕМЕННЫХ

Неопределенные и неточные данные типичны для многокритериальных задач, поэтому теория нечетких множеств является адекватным инструментом для их решения. Цель данной работы – применить метод TOPSIS в нечеткой среде с использованием синтаксически независимых лингвистических переменных. Из-за частого появления нечетких понятий в данных для принятия решений, четких значений недостаточно для моделирования реальных ситуаций. В предлагаемом подходе оценка каждой альтернативы и вес каждого критерия выражаются синтаксически независимыми лингвистическими переменными, произвольные значения которых определяются через

семантику базовых терминов. Рассматривается подход к вычислению произвольных (не базовых) значений лингвистических переменных, используемых экспертом. Базовые значения используемых синтаксически независимых лингвистических переменных задаются треугольными нечеткими числами. Предлагаемый подход хорошо подходит для решения задачи группового принятия решений в нечеткой среде, когда эксперт или группа экспертов не связана ограниченным количеством нечетких оценок. Здесь в качестве весовых коэффициентов важности критериев и оценки качественных критериев рассматриваются нечеткие переменные, которые являются значениями синтаксически независимых лингвистических переменных «важность критерия» и «оценка по критерию». В работе предлагается использовать метод центра тяжести для вычисления расстояния между двумя треугольными нечеткими числами. В соответствии с концепцией нечеткого TOPSIS определяется коэффициент близости для определения порядка ранжирования всех альтернатив путем вычисления расстояний до нечеткого положительного идеального решения (PIS) и до нечеткого отрицательного идеального решения (NIS). В работе также приводится пример для иллюстрации предлагаемого подхода.

Многокритериальное принятие решений; нечеткое множество; нечеткое число; лингвистическая переменная; треугольное нечеткое число; нечеткий метод TOPSIS; положительное идеальное решение (PIS); отрицательное идеальное решение (NIS).

A.V. Bozhenyuk, I.A. Dubchak, O.V. Kosenko

APPLICATION OF TOPSIS FUZZY METHOD FOR DECISION MAKING USING SYNTACTICALLY INDEPENDENT LINGUISTIC VARIABLES

Uncertain and imprecise data are typical for multicriteria problems, making fuzzy set theory an adequate tool for their solution. The objective of this paper is to apply the TOPSIS method in a fuzzy environment using syntactically independent linguistic variables. Due to the frequent occurrence of fuzzy concepts in decision-making data, crisp values are insufficient for modelling real-world situations. In the proposed approach, the evaluation of each alternative and the weight of each criterion are expressed by syntactically independent linguistic variables, whose arbitrary values are determined through the semantics of the base terms. An approach to calculating arbitrary (non-base) values of linguistic variables used by an expert is considered. The base values of the syntactically independent linguistic variables used are specified by triangular fuzzy numbers. The proposed approach is well suited for solving group decision-making problems in a fuzzy environment, where an expert or group of experts is not constrained by a limited number of fuzzy assessments. Here, fuzzy variables representing the values of the syntactically independent linguistic variables "criterion importance" and "criterion score" are considered as weighting coefficients for criterion importance and evaluation of qualitative criteria. The paper proposes using the centre of gravity method to calculate the distance between two triangular fuzzy numbers. In accordance with the fuzzy TOPSIS concept, a proximity coefficient is defined to determine the ranking order of all alternatives by calculating the distances to the fuzzy positive ideal solution (PIS) and the fuzzy negative ideal solution (NIS). An example is provided to illustrate the proposed approach.

Multicriteria decision making, fuzzy set, fuzzy number, triangular fuzzy number, linguistic variable, fuzzy TOPSIS method, positive ideal solution (PIS), negative ideal solution (NIS).

Введение. Принятие решений – это процесс поиска наилучшего варианта из всех возможных альтернатив. Практически во всех таких задачах повсеместно присутствует множество критериев для оценки альтернатив. То есть, для многих таких задач лицо, принимающее решения (ЛПР), хочет решить задачу многокритериального принятия решений (МКПР). Задача МКПР может быть кратко выражена в матричном виде как:

$$D = \begin{matrix} & Cr_1 & Cr_2 & \dots & Cr_n \\ \begin{matrix} A_1 \\ A_2 \\ \dots \\ A_m \end{matrix} & \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1n} \\ x_{21} & x_{22} & \dots & x_{2n} \\ \dots & \dots & \dots & \dots \\ x_{m1} & x_{m2} & \dots & x_{mn} \end{bmatrix} \end{matrix},$$

$$W = [w_1, w_2, \dots, w_n].$$

Здесь $A_1, A_2, \dots, A_m$ – возможные альтернативы, из которых ЛПР, должен выбрать, $Cr_1, Cr_2, \dots, Cr_n$ – критерии, по которым измеряется эффективность альтернатив, x_{ij} – оценка альтернативы A_i относительно критерия Cr_j ; w_j – вес критерия Cr_j .

В классических методах МКПР оценки и веса критериев известны точно [1–3]. Один из известных классических методов решения является метод TOPSIS [2]. Данный метод опирается на идею, согласно которой наилучшая альтернатива должна характеризоваться минимальным удалением от положительного идеального решения (PIS) и максимальным удалением от отрицательного идеального решения (NIS). В классической версии TOPSIS показатели эффективности и весовые коэффициенты критериев представляются в виде точных числовых значений. Однако во многих практических ситуациях использование четких данных оказывается недостаточным для адекватного описания реальных условий. Поскольку человеческие суждения, в том числе предпочтения, по своей природе нечетки, и люди не всегда могут выразить свои предпочтения точным числом, более обоснованным с практической точки зрения является применение лингвистических оценок вместо числовых значений. Иными словами, предполагается, что оценки альтернатив и веса критериев в задаче задаются с помощью лингвистических переменных [4–7]. В работах [8–10] концепция TOPSIS была расширена для разработки методологии МКПР в нечеткой среде. Учитывая нечеткость данных решения, ЛПР использует значения лингвистических переменных для оценки весов всех критериев и оценок каждой альтернативы относительно каждого критерия. В результате получается нечеткая матрица решений и далее – взвешенная нормализованная нечеткая матрица. При этом ЛПР использует ограниченное, заранее определенное множество значений лингвистических переменных. Согласно логике метода TOPSIS, сначала формируются нечеткое положительное идеальное решение (FPIS) и нечеткое отрицательное идеальное решение (FNIS). Далее в настоящей работе предлагается использовать метод вершин для расчета расстояния между двумя треугольными нечеткими оценками. После этого для каждой альтернативы вычисляются расстояния до FPIS и FNIS, а также коэффициент близости, на основе которого производится ранжирование всех альтернатив. Чем выше значение коэффициента близости, тем ближе альтернатива к FPIS и дальше от FNIS одновременно.

В данной работе для оценки весов всех критериев и оценок каждой альтернативы относительно каждого критерия предлагается использовать синтаксически независимые лингвистические переменные, что позволяет более адекватно описывать нечеткость экспертных параметров.

Для разработки метода TOPSIS на основе синтаксически независимых лингвистических переменных данная статья организована следующим образом. Во втором разделе представлены основные определения и обозначения нечеткого числа и лингвистической переменной. В разделе 3 представлен метод TOPSIS в контексте синтаксически независимых лингвистических переменных. Затем предложенный метод иллюстрируется примером. В заключение приводятся некоторые выводы.

1. Основные понятия и определения. Рассмотрим основные определения нечетких множеств согласно [5, 6, 11–14], которые будут использоваться в данной статье.

Определение 1. Нечеткое множество $\tilde{A}$ в универсальном множестве X характеризуется функцией принадлежности $\mu_{\tilde{A}}(x)$, которая сопоставляет каждому элементу $x \in X$ действительное число из интервала $[0, 1]$. Значение функции $\mu_{\tilde{A}}(x)$ называется степенью принадлежности x к $\tilde{A}$ [13].

Определение 2. Нечеткое множество $\tilde{A}$ универсального множества X является *выпуклым* тогда и только тогда, когда для всех $x_1, x_2 \in X$ выполняется:

$$\mu_{\tilde{A}}(\lambda x_1 + (1 - \lambda)x_2) \geq \min(\mu_{\tilde{A}}(x_1), \mu_{\tilde{A}}(x_2)),$$

где параметр $\lambda \in [0, 1]$.

Определение 3. Нечеткое множество $\tilde{A}$ универсального множества X называется *нормальным* нечетким множеством, если $(\exists x \in X)[\mu_{\tilde{A}}(x) = 1]$.

Определение 4. Нечеткое число – это нечеткое подмножество на универсальном множестве X , которое является одновременно выпуклым и нормальным.

Определение 5. Треугольное нечеткое число $\tilde{n}$ есть тройка $\tilde{n} = (a_{\tilde{n}}, b_{\tilde{n}}, c_{\tilde{n}})$, функция принадлежности $\mu_{\tilde{n}}(x)$ которой определяется как [11]:

$$\mu_{\tilde{n}}(x) = \begin{cases} 0, & x < a_{\tilde{n}}, \\ \frac{x - a_{\tilde{n}}}{b_{\tilde{n}} - a_{\tilde{n}}}, & a_{\tilde{n}} \leq x \leq b_{\tilde{n}}, \\ \frac{x - c_{\tilde{n}}}{b_{\tilde{n}} - c_{\tilde{n}}}, & b_{\tilde{n}} \leq x \leq c_{\tilde{n}}, \\ 0, & x > c_{\tilde{n}}. \end{cases}$$

Определение 6. Трапецевидное нечеткое число $\tilde{m}$ может быть определено четверкой $\tilde{m} = (a_{\tilde{m}}, b_{\tilde{m}}, c_{\tilde{m}}, d_{\tilde{m}})$, определяемая функцией принадлежности:

$$\mu_{\tilde{m}}(x) = \begin{cases} 0 & \text{if } x < a_{\tilde{m}}, \\ \frac{x - a_{\tilde{m}}}{b_{\tilde{m}} - a_{\tilde{m}}} & \text{if } a_{\tilde{m}} \leq x \leq b_{\tilde{m}}, \\ \frac{d_{\tilde{m}} - x}{d_{\tilde{m}} - c_{\tilde{m}}} & \text{if } c_{\tilde{m}} \leq x \leq d_{\tilde{m}}, \\ 0 & \text{if } x > d_{\tilde{m}}. \end{cases}$$

Здесь $a_{\tilde{m}} \leq b_{\tilde{m}} \leq c_{\tilde{m}} \leq d_{\tilde{m}}$.

Очевидно, что треугольные нечеткие числа являются частными случаями трапецевидных нечетких чисел, у которых $b_{\tilde{m}} = c_{\tilde{m}}$.

Определение 7. Матрица $\tilde{D}$ называется *нечеткой матрицей*, если хотя бы один элемент в $\tilde{D}$ является нечетким числом [10].

Определение 8. Нечеткой переменной [5] называется тройка $\langle \alpha, X, \tilde{C}_\alpha \rangle$. Здесь α – наименование нечеткой переменной; X – область ее определения; $\tilde{C}_\alpha = \{ \langle \mu_{\tilde{C}}(x)/x \rangle \}$ – нечеткое множество с функцией принадлежности $\mu_{\tilde{C}}$ на X , описывающее ограничения на возможные значения нечеткой переменной α .

Определение 9. Лингвистическая переменная – это переменная, значения которой являются лингвистическими терминами [6]. Формально лингвистической переменной называется пятерка $\langle \beta, T, X, G, M \rangle$. Здесь β – наименование лингвистической переменной; T – множество ее значений (терм-множество), представляющих собой наименования нечетких переменных. Множество T называется базовым терм-множеством лингвистической переменной; G – синтаксическая процедура (грамматика), позволяющая генерировать новые термы. Множество $T^* = T \cup G(T)$ называется расширенным терм-множеством лингвистической переменной; M – семантическая процедура, позволяющая превратить каждое новое значение лингвистической переменной, образуемое процедурой G , в нечеткую переменную, иными словами, приписать ему семантику путем формирования соответствующего нечеткого множества.

Лингвистические переменные, у которых процедура образования новых значений G определяется указанным выше способом, были названы *синтаксически зависимыми* [15]. Наряду с ними существует класс лингвистических переменных, у которых процедура образования новых значений зависит не от множества базовых значений T , а от области определения X , т. е. $G = G(X)$. Такие лингвистические переменные в работе [15] были названы *синтаксически независимыми*.

2. Определение значений синтаксически независимой лингвистической переменной. Пусть на числовой оси X заданы базовые значения синтаксически независимой лингвистической переменной β , у которой базовое терм-множество T есть семейство положительных нечетких треугольных чисел. Обозначим через $\tilde{m}$ – «примерно m » и $\tilde{n}$ – «примерно n » – два *соседних* треугольных нечетких числа из терм-множества T , соответственно $\tilde{m} = (a_{\tilde{m}}, b_{\tilde{m}}, c_{\tilde{m}})$ и $\tilde{n} = (a_{\tilde{n}}, b_{\tilde{n}}, c_{\tilde{n}})$ таких, что $b_{\tilde{m}} < b_{\tilde{n}}$.

Обозначим через $l_{\tilde{m}}^L = b_{\tilde{m}} - a_{\tilde{m}}$, $l_{\tilde{m}}^R = c_{\tilde{m}} - b_{\tilde{m}}$, $l_{\tilde{n}}^L = b_{\tilde{n}} - a_{\tilde{n}}$, $l_{\tilde{n}}^R = c_{\tilde{n}} - b_{\tilde{n}}$ их левые и правые разбросы.

Пусть «примерно k » – произвольное (не входящее в T) значение лингвистической переменной β , с функцией принадлежности $\mu_{\tilde{k}}(x)$ в виде нечеткого треугольного числа $\tilde{k} = (a_{\tilde{k}}, b_{\tilde{k}}, c_{\tilde{k}})$. При этом выполняется $a_{\tilde{k}} \leq b_{\tilde{k}} \leq c_{\tilde{k}}$. Обозначим через $l_{\tilde{k}}^L = b_{\tilde{k}} - a_{\tilde{k}}$,

$l_k^R = c_k - b_k$. Для определения значений a_k и c_k в работе [15] был предложен подход, согласно которому чем ближе значение b_k к $b_{\tilde{m}}$ и дальше от $b_{\tilde{n}}$, тем значения l_k^L и l_k^R ближе к значениям $l_{\tilde{m}}^L$ и $l_{\tilde{m}}^R$ и дальше от значений $l_{\tilde{n}}^L$ и $l_{\tilde{n}}^R$ соответственно:

$$l_k^L = \frac{b_{\tilde{n}} - b_k}{b_{\tilde{n}} - b_{\tilde{m}}} \times l_{\tilde{m}}^L + (1 - \frac{b_{\tilde{n}} - b_k}{b_{\tilde{n}} - b_{\tilde{m}}}) \times l_{\tilde{n}}^L; \quad l_k^R = \frac{b_{\tilde{n}} - b_k}{b_{\tilde{n}} - b_{\tilde{m}}} \times l_{\tilde{m}}^R + (1 - \frac{b_{\tilde{n}} - b_k}{b_{\tilde{n}} - b_{\tilde{m}}}) \times l_{\tilde{n}}^R. \quad (1)$$

Тогда $a_k = b_k - l_k^L$ и $c_k = b_k + l_k^R$. Это схематично показано на рис. 1:

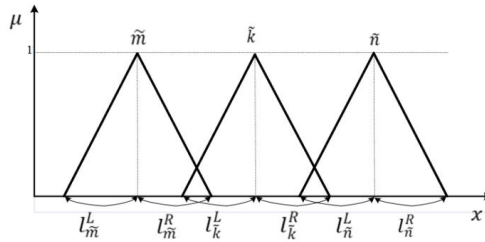


Рис. 1. Определение треугольной функции принадлежности $\mu_k(x)$

Эта же методика используется и для вычисления левой и правой границ нечетких трапециевидных чисел:

Пусть $\tilde{m} = (a_{\tilde{m}}, b_{\tilde{m}}, c_{\tilde{m}}, d_{\tilde{m}})$, $\tilde{n} = (a_{\tilde{n}}, b_{\tilde{n}}, c_{\tilde{n}}, d_{\tilde{n}})$ – два неотрицательных трапециевидных нечетких числа, являющиеся соседними базовыми значениями синтаксически независимой лингвистической переменной β . Обозначим через $l_{\tilde{m}}^L = b_{\tilde{m}} - a_{\tilde{m}}$, $l_{\tilde{m}}^R = d_{\tilde{m}} - c_{\tilde{m}}$, $l_{\tilde{n}}^L = b_{\tilde{n}} - a_{\tilde{n}}$, $l_{\tilde{n}}^R = d_{\tilde{n}} - c_{\tilde{n}}$.

Пусть $\tilde{k} = (a_{\tilde{k}}, b_{\tilde{k}}, c_{\tilde{k}}, d_{\tilde{k}})$ – произвольное значение лингвистической переменной β , с функцией принадлежности $\mu_k(x)$, и при этом выполняются неравенства: $b_{\tilde{m}} \leq b_{\tilde{k}} \leq b_{\tilde{n}}$ и $c_{\tilde{m}} \leq c_{\tilde{k}} \leq c_{\tilde{n}}$. Тогда значения $l_k^L = b_{\tilde{k}} - a_{\tilde{k}}$ и $l_k^R = d_{\tilde{k}} - c_{\tilde{k}}$ вычисляются аналогично, как и для треугольных нечетких чисел:

$$l_k^L = \frac{b_{\tilde{n}} - b_k}{b_{\tilde{n}} - b_{\tilde{m}}} \times l_{\tilde{m}}^L + (1 - \frac{b_{\tilde{n}} - b_k}{b_{\tilde{n}} - b_{\tilde{m}}}) \times l_{\tilde{n}}^L; \quad l_k^R = \frac{c_{\tilde{n}} - c_k}{c_{\tilde{n}} - c_{\tilde{m}}} \times l_{\tilde{m}}^R + (1 - \frac{c_{\tilde{n}} - c_k}{c_{\tilde{n}} - c_{\tilde{m}}}) \times l_{\tilde{n}}^R.$$

3. Определение расстояний между нечеткими числами. Существует множество функций измерения расстояния между нечеткими числами [16]. Рассмотрим метод измерения расстояния на основе центра тяжести [17, 18]. Центр тяжести *COG* трапециевидного нечеткого числа $\tilde{m} = (a_{\tilde{m}}, b_{\tilde{m}}, c_{\tilde{m}}, d_{\tilde{m}})$ определится следующим образом [19]:

$$COG(\tilde{m}) = (\alpha_{\tilde{m}}, \beta_{\tilde{m}}), \quad (2)$$

где

$$\alpha_{\tilde{m}} = \frac{1}{3} \left[a_{\tilde{m}} + b_{\tilde{m}} + c_{\tilde{m}} + d_{\tilde{m}} - \frac{d_{\tilde{m}} \cdot c_{\tilde{m}} - a_{\tilde{m}} \cdot b_{\tilde{m}}}{(d_{\tilde{m}} + c_{\tilde{m}}) - (a_{\tilde{m}} + b_{\tilde{m}})} \right], \quad (3)$$

$$\beta_{\tilde{m}} = \frac{1}{3} \left[1 + \frac{c_{\tilde{m}} - b_{\tilde{m}}}{(d_{\tilde{m}} + c_{\tilde{m}}) - (a_{\tilde{m}} + b_{\tilde{m}})} \right]. \quad (4)$$

Следствие 1. Для любого треугольного нечеткого числа $\tilde{n} = (a_{\tilde{n}}, b_{\tilde{n}}, c_{\tilde{n}})$ как частный случай трапециевидного нечеткого числа, у которого $b_{\tilde{m}} = c_{\tilde{m}}$ его центр тяжести можно определить как:

$$\alpha_{\tilde{n}} = \frac{1}{3} [a_{\tilde{n}} + b_{\tilde{n}} + c_{\tilde{n}}], \quad (5)$$

$$\beta_{\tilde{n}} = \frac{1}{3}. \quad (6)$$

Пусть $\tilde{m} = (a_{\tilde{m}}, b_{\tilde{m}}, c_{\tilde{m}}, d_{\tilde{m}})$, $\tilde{n} = (a_{\tilde{n}}, b_{\tilde{n}}, c_{\tilde{n}}, d_{\tilde{n}})$ – два трапециевидных нечетких числа. Обозначим через $(l_{\tilde{m}}^L, l_{\tilde{m}}^R)$, $(l_{\tilde{n}}^L, l_{\tilde{n}}^R)$ – их левые и правые разбросы, через $(\alpha_{\tilde{m}}, \beta_{\tilde{m}})$, $(\alpha_{\tilde{n}}, \beta_{\tilde{n}})$ – их центры тяжести. Тогда мера расстояния между двумя указанными трапециевидными нечеткими числами определяется как [17]:

$$d(\tilde{m}, \tilde{n}) = \max\{|\alpha_{\tilde{m}} - \alpha_{\tilde{n}}|, |\beta_{\tilde{m}} - \beta_{\tilde{n}}|, |l_{\tilde{m}}^L - l_{\tilde{n}}^L|, |l_{\tilde{m}}^R - l_{\tilde{n}}^R|\}. \quad (7)$$

Следствие 2. Пусть $\tilde{m} = (a_{\tilde{m}}, b_{\tilde{m}}, c_{\tilde{m}})$ и $\tilde{n} = (a_{\tilde{n}}, b_{\tilde{n}}, c_{\tilde{n}})$ – два треугольных нечетких числа, тогда

$$d(\tilde{m}, \tilde{n}) = \max\{|\alpha_{\tilde{m}} - \alpha_{\tilde{n}}|, |l_{\tilde{m}}^L - l_{\tilde{n}}^L|, |l_{\tilde{m}}^R - l_{\tilde{n}}^R|\}. \quad (8)$$

4. Предложенный подход. В этом разделе предлагается подход к расширению нечеткого метода TOPSIS с использованием синтаксически независимых лингвистических переменных. Этот подход хорошо подходит для решения задачи группового принятия решений в нечеткой среде, когда эксперт или группа экспертов не связана ограниченным количеством нечетких оценок. Здесь в качестве весовых коэффициентов важности критериев и оценки качественных критериев рассматриваются нечеткие переменные, которые являются значениями синтаксически независимых лингвистических переменных β_1 – «важность критерия» и β_2 – «оценка по критерию».

Весовые коэффициенты важности каждого критерия можно получить либо путем прямого присвоения, либо косвенно с помощью попарных сравнений [20]. Здесь предлагается, чтобы лица, принимающие решения, использовали лингвистические переменные (показанные в табл. 1 и 2) для оценки важности критериев и рейтингов альтернатив относительно различных критериев.

Лицам, принимающим решения, часто бывает сложно присвоить точную оценку альтернативе по рассматриваемым критериям. В этой ситуации задача *нечеткого* многокритериального принятия решений может быть выражена в матричном виде следующим образом:

$$\tilde{D} = \begin{matrix} & Cr_1 & Cr_2 & \dots & Cr_n \\ \begin{matrix} A_1 \\ A_2 \\ \dots \\ A_m \end{matrix} & \begin{bmatrix} \tilde{x}_{11} & \tilde{x}_{12} & \dots & \tilde{x}_{1n} \\ \tilde{x}_{21} & \tilde{x}_{22} & \dots & \tilde{x}_{2n} \\ \dots & \dots & \dots & \dots \\ \tilde{x}_{m1} & \tilde{x}_{m2} & \dots & \tilde{x}_{mn} \end{bmatrix} \end{matrix},$$

$$\tilde{W} = [\tilde{w}_1, \tilde{w}_2, \dots, \tilde{w}_n].$$

Здесь $\tilde{x}_{ij}$ и $\tilde{w}_j$ – произвольные значения (базовые и/или не базовые) лингвистических переменных, вычисляемые рассмотренным выше способом. В дальнейшем будем считать, что они описываются треугольными нечеткими числами: $\tilde{x}_{ij} = (a_{ij}, b_{ij}, c_{ij})$ и $\tilde{w}_j = (w_{j1}, w_{j2}, w_{j3})$.

Применим линейное масштабирование для преобразования различных шкал критериев в сопоставимую шкалу. Таким образом получаем *нормализованную нечеткую матрицу решений*:

$$\tilde{R} = [\tilde{r}_{ij}]_{m \times n}. \quad (9)$$

В работах [8, 10] элементы матрицы $\tilde{R}$ предложено определять следующим образом. Пусть B и C – это подмножества критериев *выгоды* и критериев *затрат* соответственно, тогда:

$$\tilde{r}_{ij} = \left(\frac{a_{ij}}{c_j^*}, \frac{b_{ij}}{c_j^*}, \frac{c_{ij}}{c_j^*} \right), \quad Cr_j \in B; \quad (10)$$

$$\tilde{r}_{ij} = \left(\frac{a_j^-}{c_{ij}}, \frac{a_j^-}{b_{ij}}, \frac{a_j^-}{a_{ij}} \right), \quad Cr_j \in C; \quad (11)$$

Здесь $c_j^* = \max_i c_{ij}$ если $Cr_j \in B$, а $a_j^- = \min_i a_{ij}$ если $Cr_j \in C$.

Учитывая различную важность каждого критерия, строится взвешенная нормализованная нечеткая матрица решений следующим образом:

$$\tilde{V} = [\tilde{v}_{ij}]_{m \times n}, \quad \tilde{v}_{ij} = \tilde{r}_{ij} \times \tilde{w}_j, \quad i = 1, 2, \dots, m, \quad j = 1, 2, \dots, n. \quad (12)$$

Здесь операция $\times$ для треугольных нечетких чисел $\tilde{x} = (x_1, x_2, x_3)$ и $\tilde{y} = (y_1, y_2, y_3)$ определяется как [9]: $\tilde{x} \times \tilde{y} = (x_1 \times y_1, x_2 \times y_2, x_3 \times y_3)$.

Согласно взвешенной нормализованной нечеткой матрице решений, элементы $\tilde{v}_{ij}$ являются нормализованными треугольными нечеткими числами. Тогда можно определить нечеткое положительно-идеальное решение A^* и нечеткое отрицательно-идеальное решение A^- как:

$$A^* = (\tilde{v}_1^*, \tilde{v}_2^*, \dots, \tilde{v}_n^*), A^- = (\tilde{v}_1^-, \tilde{v}_2^-, \dots, \tilde{v}_n^-), \quad (13)$$

где согласно [8, 9] величины $\tilde{v}_j^* = \max_{i=\overline{1,n}} v_{ij}$ и $\tilde{v}_j^- = \min_{i=\overline{1,n}} v_{ij}$, $j=1,2,\dots,n$,

Расстояние каждой альтернативы от A^* и A^- рассчитывается как:

$$d_i^* = \sum_{j=1}^n d(\tilde{v}_{ij}, \tilde{v}_j^*), \quad i = 1, 2, \dots, m, \quad (14)$$

$$d_i^- = \sum_{j=1}^n d(\tilde{v}_{ij}, \tilde{v}_j^-), \quad i = 1, 2, \dots, m. \quad (15)$$

Для определения порядка ранжирования всех альтернатив после вычисления d_i^* и d_i^- каждой альтернативы A_i ($i=1, 2, \dots, m$) определяется коэффициент близости каждой альтернативы:

$$CC_i = \frac{d_i^-}{d_i^* + d_i^-}, \quad i = 1, 2, \dots, m. \quad (17)$$

В итоге алгоритм нечеткого многокритериального принятия решений методом TOPSIS с использованием синтаксически независимых лингвистических переменных будет иметь вид:

Шаг 1: Определяются базовые значений синтаксически независимых лингвистических переменных критериев оценки.

Шаг 2: Задается семантика базовых значений синтаксически независимых лингвистических переменных для весовых коэффициентов важности критериев и лингвистических оценок альтернатив относительно критериев.

Шаг 3: Задаются нечеткие веса $\{\tilde{w}_j\}$ критериев $\{Cr_j\}$ и нечеткие оценки $\{\tilde{x}_{ij}\}$ альтернатив по соответствующим критериям ($i = 1, 2, \dots, m$, $j = 1, 2, \dots, n$); вычисляются семантики используемых (не базовых) значений $\{\tilde{w}_j\}$ и $\{\tilde{x}_{ij}\}$.

Шаг 4: Строится нечеткая матрица решений.

Шаг 5: Строится нормализованная нечеткая матрица решений.

Шаг 6: Строится взвешенная нормализованная нечеткая матрица решений.

Шаг 7: Определяется нечеткое положительно-идеальное решение A^* и нечеткое отрицательно-идеальное решение A^- .

Шаг 8: Вычисляется расстояние каждой альтернативы от A^* и A^- соответственно.

Шаг 9: Вычисляется коэффициент близости каждой альтернативы.

Шаг 10: Все альтернативы ранжируются в порядке уменьшения коэффициента близости.

5. Числовой пример: выбор перевозчика для транспортно-логистической компании. Торговая компания (заказчик) проводит тендер на выбор логистического оператора (перевозчика) для доставки товаров по стране на следующий год. Здесь альтернативами (A_i) и критериями выбора (Cr_j) являются соответственно:

A_1 – Перевозчик №1;

A_2 – Перевозчик №2;

A_3 – Перевозчик №3.

Cr_1 – тариф (стоимость перевозки);

Cr_2 – время доставки (скорость);

Cr_3 – финансовая стабильность;

Cr_4 – качество сервиса (сохранность груза, информирование).

На шаге 1 были определены базовые значения переменных β_1 – «важность критерия» и β_2 – «оценка по критерию», задаваемые на множествах $X_1 = [0,1]$ и $X_2 = [0,10]$ соответственно. В качестве базовых значений переменной β_1 были выбраны значения «малая», «средняя», «большая» – $\{\alpha_{11}, \alpha_{12}, \alpha_{13}\}$; а переменной β_2 – «плохая», «удовлетворительная», «хорошая», «отличная» – $\{\alpha_{21}, \alpha_{22}, \alpha_{23}, \alpha_{24}\}$.

На шаге 2 была определена семантика (значения треугольных нечетких чисел) базовых значений переменных β_1 и β_2 , приведенная в табл. 1 и 2.

Таблица 1

Семантика базовых значений переменной β_1

α_{1i}	значение
α_{11} – «малая»	$\tilde{0} = (-0.3, 0, 0.3)$
α_{12} – «средняя»	$\tilde{0.5} = (0.2, 0.5, 0.8)$
α_{13} – «большая»	$\tilde{1} = (0.6, 1.0, 1.4)$

Таблица 2

Семантика базовых значений переменной β_2

α_{2i}	значение
α_{21} – «плохая»	$\tilde{0} = (-1, 0, 1)$
α_{22} – «удовлетворительная»	$\tilde{3} = (0, 3, 6)$
α_{23} – «хорошая»	$\tilde{7} = (5, 7, 9)$
α_{24} – «отличная»	$\tilde{10} = (9, 10, 11)$

Здесь соответствующие параметры крайних базовых значений ($\tilde{0}, \tilde{1}, \tilde{0}, \tilde{10}$) выходят за пределы областей определения $X_1 = [0, 1]$ и $X_2 = [0, 10]$. Это сделано искусственно для того, чтобы в дальнейшем корректно вычислялись параметры произвольных значений β_1 и β_2 .

На шаге 3 была определена матрица $\tilde{D}$ нечетких оценок и вектор весов $\tilde{W}$ критериев $\{Cr_j\}$:

$$\tilde{D} = \begin{matrix} & Cr_1 & Cr_2 & Cr_3 & Cr_4 \\ \begin{matrix} A_1 \\ A_2 \\ A_3 \end{matrix} & \left| \begin{matrix} \tilde{8} & \tilde{5} & \tilde{3} & \tilde{4} \\ \tilde{7} & \tilde{7} & \tilde{10} & \tilde{8} \\ \tilde{9} & \tilde{9} & \tilde{3} & \tilde{5} \end{matrix} \right| \end{matrix}, \quad \tilde{W} = [\tilde{0.7}, \tilde{0.8}, \tilde{0.2}, \tilde{1.0}].$$

Вычисляем семантику используемых не базовых значений переменных β_1 и β_2 :

Так для нечеткого числа $\tilde{0.7}$ определяем левый и правый разбросы:

$$l_{0.7}^L = \frac{b_1 - b_{0.7}}{b_1 - b_{0.5}} \times l_{0.5}^L + \left(1 - \frac{b_1 - b_{0.7}}{b_1 - b_{0.5}}\right) \times l_1^L = \frac{1 - 0.7}{1 - 0.5} \times 0.3 + \left(1 - \frac{0.3}{0.5}\right) \times 0.4 = 0.34; \quad l_{0.7}^R = l_{0.7}^L = 0.34.$$

Получаем: $\tilde{0.7} = (0.36, 0.7, 1.04)$.

Аналогично находим параметры для остальных нечетких треугольных чисел: $\tilde{0.8} = (0.44, 0.8, 1.16)$, $\tilde{0.2} = (-0.1, 0.2, 0.5)$, $\tilde{8} = (6.33, 8, 9.66)$, $\tilde{9} = (7.67, 9, 10.33)$, $\tilde{5} = (2.5, 5, 7.5)$, $\tilde{4} = (1.25, 4, 6.25)$.

На шаге 4 матрица $\tilde{D}$ и вектор весов $\tilde{W}$ примут вид:

$$\tilde{D}' = \begin{matrix} & Cr_1 & Cr_2 & Cr_3 & Cr_4 \\ \begin{matrix} A_1 \\ A_2 \\ A_3 \end{matrix} & \left| \begin{matrix} (6.33, 8, 9.66) & (2.50, 5, 7.50) & (0.0, 3, 6.0) & (1.25, 4, 6.75) \\ (5.0, 7, 9.00) & (5.0, 7, 9.0) & (9.0, 10, 11) & (6.33, 8, 9.66) \\ (7.67, 9, 10.33) & (7.67, 9, 10.33) & (0.0, 3, 6.0) & (2.5, 5, 7.5) \end{matrix} \right| \end{matrix},$$

$$\tilde{W}' = [(0.32, 0.7, 1.04), (0.44, 0.8, 1.16), (-0.1, 0.2, 0.5), (0.6, 1, 1.4)].$$

Учитывая, что критерии Cr_1 и Cr_2 являются критериями затрат, а критерии Cr_3 и Cr_4 – критериями выгоды, на шаге 5 определяем значения $a_1^- = 5.0$, $a_2^- = 2.5$, $c_3^* = 11$, $c_4^* = 9.66$.

Нормализованная нечеткая матрица решений примет вид:

$$\tilde{R} = \begin{matrix} & Cr_1 & Cr_2 & Cr_3 & Cr_4 \\ \begin{matrix} A_1 \\ A_2 \\ A_3 \end{matrix} & \left| \begin{matrix} (0.52, 0.63, 0.79) & (0.33, 0.50, 1.00) & (0.0, 0.27, 0.55) & (0.13, 0.41, 0.70) \\ (0.56, 0.71, 1.00) & (0.28, 0.38, 0.50) & (0.82, 0.91, 1.0) & (0.66, 0.83, 1.0) \\ (0.48, 0.65, 0.65) & (0.24, 0.28, 0.33) & (0.0, 0.27, 0.55) & (0.26, 0.52, 0.78) \end{matrix} \right| \end{matrix}.$$

На шаге 6 определяем взвешенную нормализованную нечеткую матрицу решений:

$$\tilde{V} = \begin{matrix} & Cr_1 & Cr_2 & Cr_3 & Cr_4 \\ \begin{matrix} A_1 \\ A_2 \\ A_3 \end{matrix} & \left| \begin{matrix} (0.16, 0.44, 0.80) & (0.15, 0.40, 1.16) & (0.0, 0.05, 0.28) & (0.08, 0.41, 0.98) \\ (0.18, 0.50, 1.04) & (0.12, 0.29, 0.53) & (-0.08, 0.18, 0.5) & (0.40, 0.83, 1.4) \\ (0.15, 0.39, 0.68) & (0.11, 0.22, 0.38) & (0.0, 0.05, 0.28) & (0.16, 0.52, 1.01) \end{matrix} \right| \end{matrix}.$$

На шаге 7 определяем нечеткое положительно-идеальное решение A^* и нечеткое отрицательно-идеальное решение A^- :

$$A^* = ((0.18, 0.50, 1.04), (0.15, 0.40, 1.16), (-0.08, 0.18, 0.5), (0.40, 0.83, 1.4)),$$

$$A^- = ((0.15, 0.39, 0.68), (0.11, 0.22, 0.38), (0.0, 0.05, 0.28), (0.08, 0.41, 0.98)).$$

На шаге 8 вычисляем расстояние каждой альтернативы от A^* и A^- соответственно. Используя выражение (8), получаем:

$$d_{11}^* = |v_{11} - v_1^*| = \max \left\{ \frac{1}{3}(1.72 - 1.40), |0.32 - 0.28|, |0.54 - 0.36| \right\} = \max \{0.11, 0.04, 0.18\} = 0.18;$$

$$d_{11}^- = |v_{11} - v_1^-| = \max \left\{ \frac{1}{3}(1.22 - 1.40), |0.24 - 0.28|, |0.29 - 0.36| \right\} = \max \{0.06, 0.04, 0.07\} = 0.07;$$

$$d_{12}^* = |v_{12} - v_2^*| = 0.0; d_{12}^- = |v_{12} - v_2^-| = 0.59;$$

$$d_{13}^* = |v_{13} - v_3^*| = 0.21; d_{13}^- = |v_{11} - v_3^-| = 0.0;$$

$$d_{14}^* = |v_{11} - v_4^*| = 0.39; d_{14}^- = |v_{11} - v_1^-| = 0.0;$$

$$d_{21}^* = |v_{21} - v_1^*| = 0.0; d_{21}^- = |v_{21} - v_1^-| = 0.17;$$

$$d_{22}^* = |v_{22} - v_1^*| = 0.52; d_{22}^- = |v_{22} - v_2^-| = 0.08;$$

$$d_{23}^* = |v_{23} - v_3^*| = 0.0; d_{23}^- = |v_{23} - v_3^-| = 0.21;$$

$$d_{24}^* = |v_{24} - v_4^*| = 0.0; d_{24}^- = |v_{24} - v_4^-| = 0.39;$$

$$d_{31}^* = |v_{31} - v_1^*| = 0.25; d_{31}^- = |v_{31} - v_1^-| = 0.0;$$

$$d_{32}^* = |v_{32} - v_2^*| = 0.60; d_{32}^- = |v_{32} - v_2^-| = 0.0;$$

$$d_{33}^* = |v_{33} - v_3^*| = 0.21; d_{33}^- = |v_{33} - v_3^-| = 0.0;$$

$$d_{34}^* = |v_{34} - v_4^*| = 0.31; d_{34}^- = |v_{34} - v_4^-| = 0.08.$$

На шаге 9 вычисляем коэффициент близости каждой альтернативы:

$$d_1^* = \sum_{j=1}^4 d_{1j}^* = 0.78; d_1^- = \sum_{j=1}^4 d_{1j}^- = 0.66; d_2^* = \sum_{j=1}^4 d_{2j}^* = 0.52; d_2^- = \sum_{j=1}^4 d_{2j}^- = 0.85; d_3^* = \sum_{j=1}^4 d_{3j}^* = 1.37; d_3^- = \sum_{j=1}^4 d_{3j}^- = 0.08.$$

Откуда, коэффициент близости каждой альтернативы будет равен:

$$CC_1 = \frac{d_1^-}{d_1^* + d_1^-} = \frac{0.66}{0.78 + 0.66} \approx 0.46;$$

$$CC_2 = \frac{d_2^-}{d_2^* + d_2^-} = \frac{0.85}{0.52 + 0.85} \approx 0.62;$$

$$CC_3 = \frac{d_3^-}{d_3^* + d_3^-} = \frac{0.08}{1.37 + 0.08} \approx 0.056.$$

На шаге 10 согласно коэффициенту близости, порядок ранжирования трех вариантов следующий: $A_2 > A_3 > A_1$. Очевидно, что лучшим выбором является вариант A_2 . Вариант A_2 лучше, чем A_3 , но оба значительно лучше варианта A_1 .

Закключение. Неопределенные и неточные данные характерны для многокритериальных задач, и поэтому теория нечетких множеств является адекватным инструментом для их решения. В данной работе предложен подход использования синтаксически независимых лингвистических переменных в процедуре TOPSIS для решения задачи многокритериального принятия решений в нечеткой среде.

Хотя предложенный в данной статье метод иллюстрируется на примере задачи выбора перевозчика для транспортно-логистической компании, он также может быть применен к таким задачам, как выбор информационных проектов, персонального отбора, выбор наилучшего пути и многим другим областям принятия управленческих решений [21].

Подтверждения. Работа выполнена при поддержке Министерства транспорта Российской Федерации в рамках проекта № 103-00001-26-00 «Разработка и исследование методологии интеллектуального геоинформационного моделирования транспортных процессов в условиях неполноты информации».

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Dyer J.S., Fishburn P.C., Steuer R.E., Zionts S., Deb K. Multiple Criteria Decision Making, Multiattribute Utility Theory: Recent Accomplishments and What Lies Ahead // *Management Science*. – 2008. Vol. 54 (7). – P. 1336-1349. DOI: 10.1287/mnsc.1070.0838.
2. Hwang C.L., Yoon K. Multiple Attributes Decision Making Methods and Applications. – Springer, Berlin Heidelberg, 1981. – DOI: 10.1007/978-3-642-48318-9.
3. Teghem J., Delhaye C., Kunsch P.L. An interactive decision support system (IDSS) for multicriteria decision aid // *Mathematical and Computer Modeling*. – 1989. – Vol. 12 (10-11). – P.1311-1320. DOI: 10.1016/0895-7177(89)90370-1.
4. Bellman R.E., Zadeh L.A. Decision-making in a fuzzy environment // *Management Science*. – 1970. – Vol. 17 (4). – P. 141-164. DOI: 10.1287/mnsc.17.4.B141.
5. Борисов А.Н., Алексеев А.В., Крумберг О.А., Меркурьева Г.В., Попов В.А. Модели принятия решений на основе лингвистической переменной. – Рига: Зинатне, 1982. – 256 с.
6. Zadeh L.A. The concept of a linguistic variable and its application to approximate reasoning-I // *Information Sciences*. – 1970. – Vol. 8 (3). – P. 199-249. – DOI: 10.1016/0020-0255(75)90036-5.
7. Herrera F., Herrera-Viedma E., Verdegay J.L. A model of consensus in group decision making under linguistic assessments // *Fuzzy Sets and Systems*. – 1996. – Vol. 78. – P. 73-87. – DOI: 10.1016/0165-0114(95)00107-7.
8. Nădăban S., Dzitac S., Dzitac I. Fuzzy TOPSIS: A General View // *Procedia Computer Science*. – 2016. – Vol. 91. – P. 823-831. – DOI: 10.1016/j.procs.2016.07.088.
9. Yang T., Hung C.C. Multiple-attribute decision making methods for plant layout design problem // *Robotics and Computer-Integrated Manufacturing*. – 2007. – Vol. 23, No. 1. – P. 126-137. – DOI: 10.1016/j.rcim.2005.12.002.
10. Chen C.T. Extensions of the TOPSIS for group decision-making under fuzzy environment // *Fuzzy Sets and Systems*. – 2000. – Vol. 114, No. 1. – P. 1-9. – DOI: 10.1016/S0165-0114(97)00377-1.
11. Kaufmann A., Gupta M.M. Introduction to Fuzzy Arithmetic: Theory and Applications. – New York: Van Nostrand Reinhold, 1985. – 361 p.
12. Kaufmann A. Introduction to the theory of fuzzy subsets. – New York: Academic Press. – 1975. – 416 p.
13. Zimmermann H.J. Fuzzy Set Theory and its Applications. – Kluwer Academic Publishers. Boston Dordrecht London. – 1991. – 525 p. DOI: 10.1007/978-94-015-7949-0.
14. Buckley J.J. Fuzzy hierarchical analysis // *Fuzzy Sets and Systems*. – 1985. – Vol. 17 (3). – P. 233-247. – DOI: 10.1016/0165-0114(85)90090-9.
15. Малышев Н.Г., Берштейн Л.С., Боженюк А.В. Нечеткие модели для экспертных систем в САПР. – М.: Энергоатомиздат. – 1991. – 136 с.
16. Zwick R., Carlstein E., Budescu D.V. Measures of similarity among fuzzy concepts: A comparative analysis // *International Journal of Approximate Reasoning*. – 1987. – Vol. 1, No. 2. – P. 221-242. – DOI: 10.1016/0888-613X(87)90015-6.
17. Ebadi M.J., AhmadianHosseini A., Salahshour S., Suleiman M., Ismail F., Shahryari B. A New Distance Measure for Trapezoidal Fuzzy Numbers // *Mathematical Problems in Engineering*. – 2013. – DOI: 10.1155/2013/424186.
18. Liu G., Wang X. A Trapezoidal Fuzzy Number-Based VIKOR Method with Completely Unknown Weight Information // *Symmetry*. – 2023. – Vol. 15, No. 2. – 559. – DOI: 10.3390/sym15020559.
19. Wang Y.G., Yang J.B., Xu D.L., Chin K.S. On the centroids of fuzzy numbers // *Fuzzy Sets and Systems*. – 2006. – Vol. 157 (7). – P. 919-926. – DOI: 10.1016/j.fss.2005.11.006.
20. Hsu H.M., Chen C.T. Aggregation of fuzzy opinions under group decision making // *Fuzzy Sets and Systems*. – 1996. – Vol. 79 (13). – P. 279-285. – DOI: 10.1016/0165-0114(95)00185-9.
21. Anbarkhan S.H. A Fuzzy-TOPSIS-Based Approach to Assessing Sustainability in Software Engineering: An Industry 5.0 Perspective // *Sustainability*. – 2023. – Vol. 15 (18). – 13844. – DOI: 10.3390/su151813844.

REFERENCES

1. Dyer J.S., Fishburn P.C., Steuer R.E., Zionts S., Deb K. Multiple Criteria Decision Making, Multiattribute Utility Theory: Recent Accomplishments and What Lies Ahead, *Management Science*, 2008, Vol. 54, No. 7, pp. 1336-1349. DOI: 10.1287/mnsc.1070.0838.
2. Hwang C.L., Yoon K. Multiple Attributes Decision Making Methods and Applications. Springer, Berlin Heidelberg, 1981. DOI: 10.1007/978-3-642-48318-9.

3. Teghem J., Delhaye C., Kunsch P.L. An interactive decision support system (IDSS) for multicriteria decision aid, *Mathematical and Computer Modeling*, 1989, Vol. 12, No. 10-11, pp. 1311-1320. DOI: 10.1016/0895-7177(89)90370-1.
4. Bellman R.E., Zadeh L.A. Decision-making in a fuzzy environment, *Management Science*, 1970, Vol. 17, No. 4, pp. 141-164. DOI: 10.1287/mnsc.17.4.B141.
5. Borisov A.N., Alekseev A.V., Krumberg O.A., Merkuryeva G.V., Popov V.A. Modeli prinyatiya resheniy na osnove lingvisticheskoy peremennoy [Models of decision-making based on a linguistic variable]. Riga: Zinatne, 1982, 256 p.
6. Zadeh L.A. The concept of a linguistic variable and its application to approximate reasoning–I, *Information Sciences*, 1970, Vol. 8, No. 3, pp. 199-249. DOI: 10.1016/0020-0255(75)90036-5.
7. Herrera F., Herrera-Viedma E., Verdegay J.L. A model of consensus in group decision making under linguistic assessments, *Fuzzy Sets and Systems*, 1996, Vol. 78, pp. 73-87. DOI: 10.1016/0165-0114(95)00107-7.
8. Nădăban S., Dzitac S., Dzitac I. Fuzzy TOPSIS: A General View, *Procedia Computer Science*, 2016, Vol. 91, pp. 823-831. DOI: 10.1016/j.procs.2016.07.088.
9. Yang T., Hung C.C. Multiple-attribute decision making methods for plant layout design problem, *Robotics and Computer-Integrated Manufacturing*, 2007, Vol. 23, No. 1, pp. 126-137. DOI: 10.1016/j.rcim.2005.12.002.
10. Chen C.T. Extensions of the TOPSIS for group decision-making under fuzzy environment, *Fuzzy Sets and Systems*, 2000, Vol. 114, No. 1, pp. 1-9. DOI: 10.1016/S0165-0114(97)00377-1.
11. Kaufmann A., Gupta M.M. Introduction to Fuzzy Arithmetic: Theory and Applications. New York: Van Nostrand Reinhold, 1985.
12. Kaufmann A. Introduction to the theory of fuzzy subsets. New York: Academic Press, 1975.
13. Zimmermann H.J. Fuzzy Set Theory and its Applications. Kluwer Academic Publishers, Boston Dordrecht London, 1991. DOI: 10.1007/978-94-015-7949-0.
14. Buckley J.J. Fuzzy hierarchical analysis, *Fuzzy Sets and Systems*, 1985, Vol. 17, No. 3, pp. 233-247. DOI: 10.1016/0165-0114(85)90090-9.
15. Malyshev N.G., Bershtein L.S., Bozhenyuk A.V. Nechetkie modeli dlya expertnykh system v SAPR [Fuzzy models for expert systems in CAD]. Moscow, Energoatomizdat, 1991, 136 p.
16. Zwick R., Carlstein E., Budescu D.V. Measures of similarity among fuzzy concepts: A comparative analysis, *International Journal of Approximate Reasoning*, 1987, Vol. 1, No. 2, pp. 221-242. DOI: 10.1016/0888-613X(87)90015-6.
17. Ebadi M.J., AhmadianHosseini A., Salahshour S., Suleiman M., Ismail F., Shahryari B. A New Distance Measure for Trapezoidal Fuzzy Numbers, *Mathematical Problems in Engineering*, 2013. DOI: 10.1155/2013/424186.
18. Liu G., Wang X. A Trapezoidal Fuzzy Number-Based VIKOR Method with Completely Unknown Weight Information, *Symmetry*, 2023, Vol. 15, No. 2, 559. DOI: 10.3390/sym15020559.
19. Wang Y.G., Yang J.B., Xu D.L., Chin K.S. On the centroids of fuzzy numbers, *Fuzzy Sets and Systems*, 2006, Vol. 157, No. 7, pp. 919-926. DOI: 10.1016/j.fss.2005.11.006.
20. Hsu H.M., Chen C.T. Aggregation of fuzzy opinions under group decision making, *Fuzzy Sets and Systems*, 1996, Vol. 79, No. 13, pp. 279-285. DOI: 10.1016/0165-0114(95)00185-9.
21. Anbarkhan S.H. A Fuzzy-TOPSIS-Based Approach to Assessing Sustainability in Software Engineering: An Industry 5.0 Perspective, *Sustainability*, 2023, Vol. 15, No. 18, 13844. DOI: 10.3390/su151813844.

Боженюк Александр Витальевич – Южный федеральный университет; e-mail: avbozhenyuk@sfedu.ru; г. Таганрог, Россия; тел.: +79198799621; кафедра информационно-аналитических систем безопасности имени профессора Берштейна Леонида Самойловича; профессор.

Дубчак Ирина Александровна – Российский университет транспорта; e-mail: iri-dubchak@yandex.ru; тел.: +79166652310; зам. директора.

Косенко Олеся Валентиновна – Южный федеральный университет; e-mail: ovkosenko@sfedu.ru; г. Таганрог, Россия; тел.: +79287740818; кафедра систем автоматического управления; доцент.

Bozhenyuk Alexander Vitalievich – Southern Federal University; e-mail: avbozhenyuk@sfedu.ru; Taganrog, Russia; phone: +79198799621; the Department of Information and Analytical Security Systems named after Professor Bershtein Leonid Samoylovich; professor.

Dubchak Irina Alexandrovna – Russian University of Transport (MIIT); e-mail: iri-dubchak@yandex.ru; Moscow, Russia; phone: +79166652310; deputy director.

Kosenko Olesia Valentinovna – Southern Federal University; e-mail: ovkosenko@sfedu.ru; Taganrog, Russia; phone: +79287740818; the Department of Automatic Control Systems; associate professor.

Е.А. Титенко, Т.М. Белова, Л.А. Лисицин, И.И. Пузанов, А.А. Положенцев

МАТЕМАТИЧЕСКАЯ МОДЕЛЬ И СХЕМА УПРАВЛЕНИЯ ПРОДУКЦИЯМИ
ДЛЯ РЕАЛИЗАЦИИ ПАРАЛЛЕЛЬНЫХ ВЫЧИСЛЕНИЙ

Актуальность. Эффективные математические модели организации параллельных вычислений используют принципы одновременного выполнения правил над независимыми фрагментами данных, а также методы обработки символической информации как объединяющей категории для различных типов данных. Продукционные системы в нотации А.А. Маркова (мл.) имеют необходимый потенциал для создания таких моделей, но стандартная схема управления в них является последовательной, что ограничивает применение таких продукционных систем в прикладных задачах. В связи с этим актуальной является задача синтеза модифицированных продукционных систем и схем управления в них, направленных на распараллеливание вычислительного процесса обработки символической информации. **Целью исследования** является уменьшение времени работы модифицированной продукционной модели за счет двунаправленной обработки цепочек символов. **Метод решения** основан на применении принципов двунаправленного грамматического разбора из теории синтаксического анализа, компиляции в алгоритмических продукционных системах и состоит в модификации стандартной схемы продукционных вычислений А.А. Маркова. Это достигается введением продукций с однократным срабатыванием, двунаправленным доступом к обрабатываемой строке, декомпозицией исходной системы на секции, введением набора охранных условий для разрешения конфликтов. Продукционная модель реализована для типовых задач символической обработки: разметка слова и обращения слова. **Результаты.** Разработана структура модифицированной продукционной системы, преодолевающая ограничение стандартной схемы А.А. Маркова – последовательный характер вычислений. В результате проведенного исследования была создана формальная модель последовательно-параллельной обработки слов, разработана модифицированная схема управления и показана корректность вычислений. Моделирование на типовых задачах показало уменьшение времени обработки слов практически в 2 раза, что было достигнуто, в том числе, за счет исключения обязательного возврата на первую продукцию в схеме управления и агрегации продукций в секции. **Заключение.** Проведенное исследование подтвердило возможность и эффективность трансформации последовательных продукционных систем А.А. Маркова в последовательно-параллельные продукционные модели с выдачей корректного результата. Перспективная область применения – методы и программно-аппаратные средства в однородных вычислительных системах для высокопроизводительной обработки символической информации.

Продукционная система; челнок; двунаправленная обработка; символическая информация; продукция; конфликт данных.

E.A. Titenko, T.M. Belova, L.A. Lisitsin, I.I. Puzanov, A.A. Polozhenev

A MATHEMATICAL MODEL AND PRODUCT MANAGEMENT SCHEME
FOR IMPLEMENTING PARALLEL PRODUCT COMPUTING

Relevance. Effective mathematical models for organizing parallel computing utilize principles of simultaneous rule execution on independent data fragments, as well as methods for processing symbolic information as a unifying category for various data types. Production systems in A.A. Markov Jr.'s notation offer the necessary potential for creating such models, but their standard control scheme is sequential, limiting their application in applied problems. Therefore, synthesizing modified production systems and their control schemes aimed at parallelizing the computational process of processing symbolic information is a pressing issue. **The goal** of this study is to reduce the execution time of a modified production model through bidirectional processing of symbol strings. **The solution method** is based on the application of principles of bidirectional grammatical parsing from the theory of syntactic analysis and compilation in algorithmic production systems and consists of modifying A.A. Markov's standard production computation scheme. This is achieved by introducing single-shot productions, bidirectional access to the string being processed, decomposing the original system into sections, and introducing a set of guard conditions for conflict resolution. The production model is implemented for typical symbolic processing tasks: word tagging and word reversal. **Results.** The structure of a modified production system was developed that overcomes the limitation of A.A. Markov's standard scheme—the sequential nature of computations. As a result of the study, a formal model of serial-parallel processing was created, a new control scheme was developed, and the correctness of the computations was demonstrated. Simulation on typical tasks showed a reduction in word processing time by almost half, achieved, in part, by eliminating the

mandatory return to the first production in the control scheme and aggregating productions into sections.
Conclusion. *The study confirmed the feasibility and effectiveness of transforming A.A. Markov's sequential production systems into serial-parallel production models that produce correct results. A promising area of application is methods and software and hardware in homogeneous computing systems for high-performance processing of symbolic information.*

Production system; shuttle; bidirectional processing; symbolic data; production; data conflict.

Введение. Создание математических моделей параллельных вычислений для вычислительных систем (ВС) и их элементов представляет один из значимых путей повышения производительности их работы, особенно при решении задач со сложноорганизованными данными и нерегулярной структурной организацией, задач искусственного интеллекта (ИИ) [1], задач глобальной оптимизации [2], адаптивного управления [3] и др. Известно, что традиционная (фон-неймановская) модель вычислений и архитектуры ВС [4, 5], базирующаяся на этой модели, в том числе суперЭВМ, кластерные ВС и др., характеризуются падением производительности при масштабировании системы в силу значительного роста накладных расходов на организацию параллельных вычислений. Возникающий разрыв между традиционными моделями параллельных вычислений (сети Петри, модели с общей памятью, модели с распределенной памятью и др.) и архитектурами ВС не позволяет эффективно решать прикладные задачи большой размерности, особенно проблемно-поискового, переборного характера.

По мнению ведущих ученых в области создания высокопроизводительных ВС и интеллектуальных систем (Luger D.F., Klein W.B., Garty M., Johnson D., Воеводин В.В., Воеводин Вл.В., Каляев И.А., Левин И.И., Курейчик В.В., Гайдук А.Р., Бурцев В.С., Стемповский А.Л., Огнев И.В., Страбыкин Д.А., Махортов С.Д., Жданов А.А., Рыбина Г.В., Сафронова Е.А., Борисов А.В. и др.) модели символьных вычислений имеют собственный потенциал для организации параллельных вычислительных действий [4]. Они используют для обработки данные в высшей форме агрегации – символьные, реляционные и др. данные, что позволяет с единых позиций вести как числовую, так и символьную (строковую) обработку наборов данных.

Современный научно-технический прогресс делает акцент в высокопроизводительных вычислениях на задачи генерации кодовых последовательностей, параллельного вывода в динамических экспертных системах, задачи структурного распознавания образов [6], комбинаторные поисковые задачи, в том числе символьной комбинаторики [7], задачи интеллектуального анализа многомерных данных, машинного обучения и понимания текста, задачи дискретной оптимизации, генерации и оценки символьных структур с заданными свойствами [8], информационно-поисковые задачи, задачи сортировки, сравнения, ранжирования данных [9, 10] др. Для класса таких задач существенная доля вычислений приходится на операции сравнения по образцу, модификации данных, структурного преобразования наборов данных.

Аппарат продукционных систем (ПС) имеет необходимые условия для эффективного решения такого класса задач. Как известно, продукционные системы (алгоритмического или исчислительного типа) осуществляют обработку данных на основе вычислительной схемы «условие → действие» [11, 12]. С точки зрения организации вычислений ПС имеют следующие особенности:

- ◆ однородность состава правил;
- ◆ потенциальный параллелизм правил;
- ◆ структурная наращиваемость;
- ◆ гибкие управляющие схемы;
- ◆ минимальный унифицированный набор базовых операций.

В работе выбран и развивается аппарат алгоритмических продукционных систем А.А. Маркова [13]. Созданная А.А. Марковым теория алгорифмов, а также фундаментальные исследования в области конструктивной математики и логики послужили основой для применения алгоритмических продукционных систем в решении множества исследовательских и прикладных задач обработки символьной информации, в том числе

обработки знаний [1]. Важнейшая особенность продукционных алгоритмов А.А. Маркова – равносильность таким признанным абстрактным системам как продукционная система Э. Поста, рекурсивные функции А. Чёрча, машина А. Тьюринга, алгоритмические комплексы А.Н. Колмогорова и др. [14, 15]. Данные абстрактные системы задают *любой* вычислительный процесс, но их отличает преимущественно последовательная схема управления, что не позволяет реализовывать эффективные по времени вычислительные процессы, в том числе последовательно-параллельные алгоритмы.

В связи с этим актуальными являются вопросы создания продукционных моделей, а также схем управления продукциями для повышения временной эффективности работы ПС.

Постановка задачи. Объектами обработки ПС являются конечные дискретные объекты, построенные из букв рабочего алфавита по определенным правилам, в виде цепочек символов (слов) [16]. Слово – это одномерный дискретный объект, формируемый в рабочем алфавите A мощностью w букв (символов) на основе следующего набора правил конкатенации:

- а) $S = S \circ \xi_1$ – правая конкатенация;
- б) $S = \xi_2 \circ S$ – левая конкатенация;
- в) $S = \xi_2 \circ S \circ \xi_1$ – двойная конкатенация.

где $\circ$ – операция конкатенации, ξ_1, ξ_2 – алфавитные переменные ($\xi_1 \in A, \xi_2 \in A$), S – начало индукции (например, пустое слово λ) [9].

Тогда слово S длиной l символов, порожаемое правилами конкатенации, имеет следующую структуру

$$S = \lambda \circ \xi_1 \circ \xi_2 \circ \dots \xi_i \circ \dots \xi_{l-1} \circ \xi_l \circ \lambda = \xi_1 \xi_2 \dots \xi_i \dots \xi_{l-1} \xi_l. \quad (1)$$

Особенность указанных правил конкатенации заключается в том, что они позволяют детализировать структуру порожаемого слова S как конкатенацию трех частей: собственного префикса слова S_{pref} , тела S_{body} и собственного суффикса слова S_{suff} ,

$$S = S_{pref} \circ S_{body} \circ S_{suff} = S_{pref} S_{body} S_{suff}. \quad (2)$$

Продукция – это математическое выражение вида

$$O \rightarrow P, \quad (3)$$

где O – левая часть (образец, шаблон); P – правая часть (подстановка, модификатор).

В записи продукции присутствует метасимвол « $\rightarrow$ ». Он обозначает направление обработки левой и правой частей продукции и порядок выполнения базовых операций: поиск по образцу и модификация слова [13].

Эти операции поиска и модификации можно представить в виде логической схемы

$$\frac{S = S_{pref} \circ O \circ S_{suff}, O \rightarrow P}{S = S_{pref} \circ P \circ S_{suff}}. \quad (4)$$

Сама алгоритмическая система А.А. Маркова представляется линейным списком из f продукций вида (3), заданных в алфавите A

$$\mathfrak{P}: \left\{ \begin{array}{l} O_1 \rightarrow P_1 \\ \dots \\ O_i \rightarrow P_i \\ \dots \\ O_f \rightarrow P_f \end{array} \right. \quad (5)$$

Согласно А. А. Маркову стандартная схема управления продукциями имеет последовательно-возвратный вид [13]. В ней на каждом шаге работы активной является единственная продукция из списка (5), после срабатывания, которой осуществляется *обязательный* возвратный переход на первую продукцию в списке (5).

Циклические возвраты с i -й продукции на первую продукцию в схеме $\mathfrak{Z}$ необходимы для организации вычислительных процессов, но служат источником избыточных временных затрат. Действительно, линейная последовательность k сработавших продукций вида $\{i_1, i_2, \dots, i_k\}$ имеет детализированный состав

$$\{(1 \dots i_1 - 1), i_1, (1 \dots i_2 - 1), i_2, \dots, (1 \dots i_k - 1), i_k\}. \quad (6)$$

В (6) в круглых скобках показаны упорядоченные наборы проверяемых, но не сработавших продукций, так как их образцы не были найдены в структуре слова S [13].

В связи с этим необходима такая эквивалентная схема $\mathfrak{Z}^*$, которая имеет сокращенное число обязательных возвратов и возможность параллельного выполнения набора продукций, что приводит к уменьшению времени T работы алгоритмической ПС

$$T(\mathfrak{Z}^*(O_1 \rightarrow P_1, O_2 \rightarrow P_2, \dots, O_f \rightarrow P_f)) \rightarrow \min. \quad (7)$$

Методы решения. Как известно, стандартная схема управления алгоритмической ПС по А.А. Маркову [13] содержит:

- ♦ стартовую продукцию (продукция с пустым образцом);
- ♦ рабочие (текущие) продукции;
- ♦ заключительную продукцию.

Кроме букв алфавита продукции в левых или правых частях могут содержать челноки или алфавитные (символьные) переменные. Эти специальные символы указывают на продукцию в составе $\mathfrak{Z}$ и на локальную область обработки в слове S , что упорядочивает проверочно-вычислительные шаги.

Для алгоритмических ПС по А.А. Маркову стандартная схема управления продукциями приведена на рис. 1, где 1/0 – обозначают переходы при срабатывании или несрабатывании продукций. Продукционная обработка завешается, если ни одна из продукций вида (3) не применима к обрабатываемому слову, т.е. после n отрицательных проверок вхождения образца в слово. Кроме того, возможен досрочный выход из стандартной схемы управления в алгоритмической ПС. Он реализуется так называемой заключительной продукцией, срабатывание которой принудительно завершает вычислительный процесс в любой позиции схемы (5).

В целом, данные особенности приводят к избыточным временным затратам работы алгоритмической ПС.

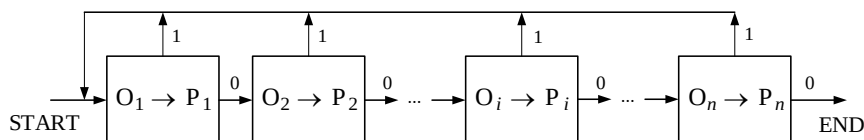


Рис. 1. Стандартная схема передачи управления по А.А. Маркову

Для уменьшения времени работы ПС предлагаются следующие структурные приемы:

- ♦ ввести в состав $\mathfrak{Z}^*$ тип продукции с однократным срабатыванием;
- ♦ считать слово S конструктивным объектом, имеющим 2 стартовых (начальных) символа с возможностью введения двух челноков и одновременной двунаправленной обработки;
- ♦ агрегировать продукции в смысловые секции и управлять в схеме не продуктами, а секциями;
- ♦ разграничить фазы (этапы) вычислительного процесса и допускать параллельное выполнение продукций в составе секций.

Первый структурный прием состоит во введении третьего типа продукций вида

$$O \rightarrow *P. \quad (8)$$

Продукция с однократным срабатыванием (8) безальтернативно передает управление на следующую продукцию, что характерно для классических процедурных арифметико-логических вычислений. Это введение позволяет расположить стартовую продук-

ционную секцию непосредственно в начале схемы $\mathfrak{F}^+$ ПС, тогда как для ПС со стандартной схемой управления характерно расположение стартовых продукций в конечной части схемы (5) для присоединения челноков в начало слова S [13].

Второй структурный прием основан на переносе принципа двунаправленной обработки цепочек символов из теории синтаксического анализа и методов грамматического разбора [17, 18]. Данные методы направлены на грамматический разбор и установление соответствия анализируемой цепочки (записи программы в нотации (Бэкуса-Наура) синтаксису языка программирования. В этом случае попеременно или параллельно реализуются два вычислительных процесса грамматического разбора анализируемой цепочки с ее двух сторон на основе правил грамматики.

Действительно, в соответствии с первыми двумя базовыми правилами конкатенации в слове S допускаются 2 стартовых символа. Слово S будет по-прежнему считаться конструктивным объектом, над которым заданы два параллельных вычислительных процесса, основанных на встречном продвижении двух челноков.

В соответствии с двумя структурными приемами стартовые продукции будут иметь вид

$$\lambda \rightarrow * \alpha, \beta * \leftarrow \lambda, \quad (9)$$

где α, β – челноки, вводимые в слово S слева и справа соответственно.

Однократное срабатывание продукций (9) приводит к преобразованию слова S к виду $S = \alpha \xi_1 \xi_2 \dots \xi_i \dots \xi_{l-1} \xi_l \beta$ и дальнейшим параллельным встречным процессам символьных преобразований с помощью челноков α, β .

Корректность организации 2-х встречных параллельных процессов обеспечивается контролем расстояния сближения двух челноков. Два встречных параллельных процесса являются независимыми до тех пор, пока челночки α, β не станут конфликтовать между собой за общие символы слова S . Устранение конфликта основано на синтезе специального охранного условия, обнаруживающего сближение челноков α, β .

Третий и четвертый структурные приемы состоят в укрупнении базового объекта ПС и исключении обязательного возвратного цикла на первую продукцию в составе $\mathfrak{F}^+$ за счет контроля проверки и разграничения фаз вычислительных процессов на основе алгебры продукций. При отсутствии зависимостей по данным между продукциями каждая фаза вычислений является обособленной, что позволяет передавать управление на продукцию в пределах секции или на следующую секцию.

Введенные структурные приемы позволили получить модифицированную схему управления (рис. 2). Новизна модифицированной схемы управления состоит в сокращении количества возвратных переходов за счет агрегации продукций в секции и осуществлении «ближних» переходов между секциями.

Далее рассматриваются две модифицированные схемы ПС:

- ♦ модифицированная схема с одновременным встречным движением двух челноков (модификация 1);
- ♦ модифицированная схема с одновременным встречным продвижением двух челноков с текущими буквами (модификация 2).

Модификация-1 ПС с параллельной рабочей секцией со встречным движением двух челноков

$$\left\{ \begin{array}{l} \lambda \rightarrow * \alpha \\ \beta * \leftarrow \lambda \\ \left\| \begin{array}{l} \alpha \xi_i \rightarrow \xi_i \alpha \\ \beta \xi_j \leftarrow \xi_j \beta \end{array} \right\| (\alpha \dots \beta) \leq h \\ \alpha \beta \rightarrow \bullet \lambda \\ \alpha \xi_i \beta \rightarrow \bullet \xi_i \end{array} \right. \quad (10)$$

Первая и вторая продукций являются стартовыми, они отвечают за фазу разметки первого и последнего символов обрабатываемого слова S челноками α (слева) и β (справа). Данные правила срабатывают один раз.

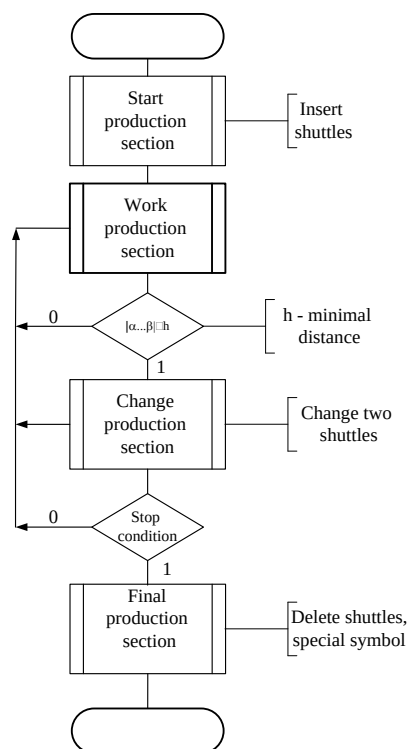


Рис. 2. Модифицированная схема управления ПС

Третья и четвертая продукция представляют собой рабочую секцию, в которой параллельно исполняются 2 продукции. Они отвечают за фазу продвижения, анализа (поиск вхождения) и обработку локальных групп символов, ассоциированных с челноками α и β . Поскольку челноки распределены по длине слова S , то продукции рабочей секции являются независимыми между собой. Следовательно, возможно их одновременное применение с передачей управления в пределах рабочей секции. При каждом срабатывании 2-х продукций из рабочей секции челноки продвигаются навстречу другу. Параллельная обработка разнесенных групп символов выполняется за несколько итераций, количество которых определяется длиной слова S и его свойством быть четным или нечетным по длине n .

При этом слово S выступает общим информационным ресурсом. Поэтому одновременная работа продукций рабочей секции потенциально ведет к возникновению конфликта при сближении челноков на минимальное расстояние h , не позволяющее детерминировано применить рабочие продукции. В зависимости от свойства четности/нечетности длины слова S минимальное расстояние h составляет 1 или 2 символа соответственно. Параллельная работа продукций из рабочей секции предваряется проверкой охранного условия вида

$$(|\alpha...\beta|) \leq h, \quad (11)$$

где h – минимальное расстояние между челноками α и β .

Вводимое охранное условие с проверкой минимального расстояния h , позволяет обнаружить конфликтные ситуации и среагировать на них, например, передать управление с рабочей секции на следующие продукции в модифицированной ПС.

Пятая и шестая продукции являются заключительными. Они прекращают обработку слова S при достижении челноками α и β середины слова S .

Модификация-1 ПС реализует два встречных процесса обхода символов слова с элементарным продвижением челноков и потенциальной обработкой символов слова S . Челноки α и β встречаются в середине слова, что служит индикатором прекращения работы.

Модификация-2 ПС с параллельной рабочей секцией при встречном продвижении челноков вместе с символами слова S имеет вид:

$$\left\{ \begin{array}{l} \lambda \rightarrow * \alpha \\ \beta^* \leftarrow \lambda \\ \left\| \begin{array}{l} \alpha \xi_i \xi_j \rightarrow \xi_j \alpha \xi_i \\ \xi_i \beta \xi_k \leftarrow \xi_k \xi_i \beta \end{array} \right\| (\alpha \dots \beta) = h \\ \alpha \xi_i \xi_j \beta \rightarrow \xi_j \beta \alpha \xi_i \\ \alpha \xi_i \xi_k \xi_j \beta \rightarrow \xi_j \beta \xi_k \alpha \xi_i \\ \beta \xi_i \mid \xi_i \alpha \rightarrow \bullet \xi_i \mid \xi_i \end{array} \right. \quad (12)$$

Первая и вторая продукции являются стартовыми, они также отвечают за фазу разметки граничных символов слова S .

Третья и четвертая продукция представляют собой рабочую секцию, в которой параллельно выполняются 2 продукции. Они отвечают за фазу продвижения вместе с челноками α и β двух символов слова S , при этом встречное перемещение символов должно пройти через все слово S . Значит, количество итераций рабочей секции пропорционально длине n самого слова. Поскольку челноки α и β распределены по длине слова S , то возможно их одновременное движение с передачей управления в пределах рабочей секции. В общем случае, рабочая секция выполняется $n/2$ раз, тогда как в традиционной схеме управления количество обходов по слову S равно n .

Одновременная работа продукций рабочей секции также потенциально ведет к возникновению конфликта при сближении челноков на минимальное расстояние h . В зависимости от свойства четности/нечетности длины слова S минимальное расстояние h составляет 2 или 3 символа соответственно, так как длина условий в продукциях рабочей секции равна 3 символам.

Соответственно параллельная работа продукций из рабочей секции также предваряется охранной проверкой вида

$$(|\alpha \dots \beta|) \leq h. \quad (13)$$

Пятая и шестая продукции составляют обменную секцию, необходимую для разрешения (снятия) конфликтной ситуации при минимальном расстоянии между челноками.

Седьмая продукция является заключительной, она имеет составной образец и прекращает обработку слова S при достижении челноками границ слова.

Таким образом, преимущество двух модифицированных схем ПС – уменьшенное время обхода символов слова S за счет параллельного движения двух челноков и реализации двух самостоятельных процессов от границ слова.

Разрешение конфликта в обеих модификациях ПС основывается на алгебре продукций, трактующей части продукций как отдельные конструктивные объекты, над которыми выполняются операции пересечения и объединения, дополнения строк [19]. Разрешение конфликта осуществляется за счет продукций, имеющих в проверочной части расширенную группу символов вместе челноками. Само разрешение конфликта сводится к взаимному обмену местами челноков вместе с прикрепленными к ним символами, что позволяет далее перейти на рабочую секцию и продолжить вычисления.

Для модификации-1 ПС конфликт проявляется сближением челноков α , β на минимальное расстояние, кратное длине образца рабочей продукции, что не позволяет корректно завершить вычисления (рис. 3).

Нижняя строка в схемах (рис. 3) показывает структуру конфликтных слов – охранных условий, при обнаружении которых дальнейшая работа рабочей секции не выполняется. Вместо этого выполняется шаг обмена челноков своими позициями, что моделирует 1 шаг одновременного срабатывания рабочих продукций.

Для модификации-2 ПС конфликт также проявляется при сближении челноков α , β на минимальное расстояние, кратное длине образца рабочей продукции (рис. 4).

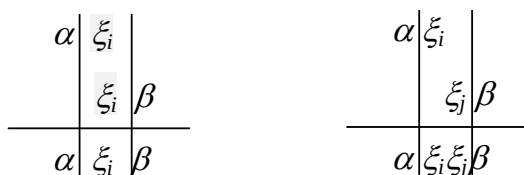


Рис. 3. Схемы выявления конфликта для модификации-1

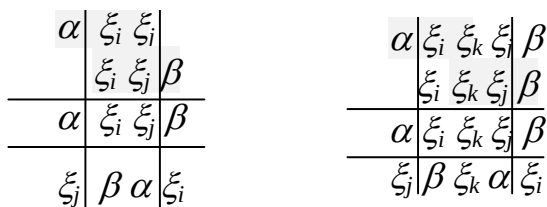


Рис. 4. Схемы разрешения конфликта для модификации-2 ПС

Две нижние строки (рис. 4), отделенные горизонтальными линиями представляют собой запись обменной продукции: верхняя строка – это охранный условный по минимальному расстоянию между челноками (образец), нижняя строка описывает взаимный обмен челноками с прикрепленными к ним символами.

Таким образом, созданы модель и схема управления алгоритмическими продукционными системами в нотации А.А. Маркова, основанные на переносе принципа двунаправленной обработки цепочек из теории синтаксического анализа, компиляции и грамматического разбора в алгоритмические продукционные вычисления. Созданные модель и схема управления отличаются возможностью реализации последовательно-параллельных процессов преобразования с помощью двух одновременно движущихся челноков, указывающих на независимые локальные фрагменты слов для их одновременной обработки с обеспечением корректности вычислений – контроле минимального расстояния между челноками.

Результаты и обсуждение. Для проверки эффективности созданных теоретических средств (модель, схема управления) рассматриваются типовые задачи обработки СИ [12, 13]:

- ♦ разметка слова;
- ♦ обращение слова.

Задача разметки заключается в «раскраске» маркерным символом всех букв слова. Данная задача решается за один проход по слову. Выполняется конкатенация заданного маркерного символа с каждой буквой слова, при этом порядок следования букв не изменяется.

Схематически работу алгоритма разметки можно описать

$$I(S = s_1 s_2 \dots s_j \dots s_n, +) = +s_1 + s_2 \dots + s_j \dots + s_n +, \quad (14)$$

где «+» – маркерный символ.

Алгоритм разметки с стандартной схемой управления А.А. Маркова имеет вид

$$\begin{cases} \alpha \xi \rightarrow \xi + \alpha & ; \text{work production} \\ \alpha \rightarrow \bullet + & ; \text{stop production} \\ \lambda \rightarrow + \alpha & ; \text{start production} \end{cases} \quad (15)$$

Данный алгоритм основан на введении слева в слово S челнока α и его побуквенном движении вправо с разметкой заданным символом «+».

Алгоритм имеет достаточно высокую эффективность, поскольку только первая и последняя буквы слова размечаются последней (стартовой) и второй (финальной) продукциями. Оставшиеся $n-2$ буквы слова S размечаются первой (рабочей) продукцией, т.е. без потерь времени на возвратные переходы. Тем не менее, срабатывания стартовой и заключительной продукций содержат холостые срабатывания.

Модификация ПС с двумя челноками состоит из 3-х секций (*start*, *work*, *final*) и имеет вид

$$\left\{ \begin{array}{l} \lambda \rightarrow * (+\alpha) \\ (\beta+)^* \leftarrow \lambda \\ \left\| \begin{array}{l} \alpha \xi_1 \rightarrow \xi_1 + \alpha \\ \beta + \xi_2 \leftarrow \xi_2 \beta \end{array} \right\| (\alpha \dots \beta) \leq 1 \\ + \alpha \beta \rightarrow \bullet \\ \alpha \xi \beta \rightarrow \bullet \xi \end{array} \right. \quad (16)$$

Первая (*start*) секция содержит 2 однократные продукции, которые последовательно размечают за 2 шага первую и последнюю буквы слова *S*. Далее управление передается на рабочую (*work*) секцию, в пределах которой осуществляется одновременное встречное движение челноков α, β и разметка двух букв одновременно на каждом шаге. За количество шагов равное $(n/2 - 2)$ осуществляется разметка оставшихся букв слова *S*. Если охранный условие $(\alpha \dots \beta) \leq 2$ выполняется, тогда управление передается на третью (*final*) секцию, которая завершает работу ПС. Третья секция состоит из двух продукции, завершающих разметку для слов четной или нечетной длины, и удаляет ставшие ненужными челноки из слова *S*.

В табл. 1, 2 показаны трассировки алгоритма разметки слова для $n=4, n=7$.

Таблица 1

Трассировка алгоритма разметки слова для $n=4$

Обрабатываемое слово <i>S</i>	Номер сработавшей продукции	Номер шага алгоритма	Примечание
$\xi_1 \xi_2 \xi_3 \xi_4$		-	Исходное слово
$+ \alpha \xi_1 \xi_2 \xi_3 \xi_4 \beta +$	1,2	1,2	Разметка челноками α, β
$+ \xi_1 + \alpha \xi_2 \xi_3 \beta + \xi_4 +$	3,4	3	Параллельная разметка
$+ \xi_1 + \xi_2 + \alpha \beta + \xi_3 + \xi_4 +$	3,4	4	Активация охранный условия
$+ \xi_1 + \xi_2 + \xi_1 + \xi_3 + \xi_4 +$	5	5	удаление челноков – результат

Таблица 2

Трассировка алгоритма разметки слова для $n=7$

Обрабатываемое слово <i>S</i>	Номер сработавшей продукции	Номер шага алгоритма	Примечание
$\xi_1 \xi_2 \xi_3 \xi_4 \xi_5 \xi_6 \xi_7$		-	Исходное слово
$+ \alpha \xi_1 \xi_2 \xi_3 \xi_4 \xi_5 \xi_6 \xi_7 \beta +$	1,2	1,2	Разметка челноками α, β
$+ \xi_1 + \alpha \xi_2 \xi_3 \xi_4 \xi_5 \xi_6 \beta + \xi_7 +$	3,4	3	Параллельная разметка
$+ \xi_1 + \xi_2 + \alpha \xi_3 \xi_4 \xi_5 \beta + \xi_6 + \xi_7 +$	3,4	4	Параллельная разметка
$+ \xi_1 + \xi_2 + \xi_3 + \alpha \xi_4 \beta + \xi_5 + \xi_6 + \xi_7 +$	3,4	5	Активация охранный условия
$+ \xi_1 + \xi_2 + \xi_3 + \xi_4 + \xi_5 + \xi_6 + \xi_7 +$	6	6	удаление челноков – результат

Задача обращения слова имеет следующую формулировку. Пусть в рабочем алфавите *A* задано слово *S* длиной *n* букв $S = \xi_1 \xi_2 \dots \xi_i \xi_{i+1} \dots \xi_n$, где $\xi_i \in A$ ($i=1-n$).

Результат работы алгоритма обращения I представляется как

$$I(S) = \tilde{S} = \xi_n \dots \xi_{i+1} \xi_i \dots \xi_2 \xi_1. \quad (17)$$

Обращение слова (строковая инверсия) осуществляется циклическими перемещениями текущей буквы вместе с челноком α вправо до конца слова. В процессе инверсии i -й символ меняет свою позицию на $(n-i-1)$ -ую позицию. Ядро алгоритма обращения $I(S) = \tilde{S}$ составляют 2 продукции

$$\begin{cases} \alpha \xi \eta \rightarrow \eta \alpha \xi \\ \lambda \rightarrow \alpha \end{cases}. \quad (18)$$

Вторая продукция – стартовая, она размечает граничный символ слова S . Продукция $\alpha \xi \eta \rightarrow \eta \alpha \xi$ выполняет движение вправо текущего символа ξ до конца слова и размечает его метасимволом α . Далее вторая продукция $\lambda \rightarrow \alpha$ повторно осуществляет левую конкатенацию челнока α к слову S . Тогда многократные применения этих текущих продукций приводят к инверсному расположению символов исходного слова.

Трассировка ядра алгоритма $I(S) = \tilde{S}$ для $n=4$ приведена в табл. 3.

Таблица 3

Трассировка ядра алгоритма обращения

Обрабатываемое слово S	Номер сработавшей продукции	Номер шага алгоритма	Примечание
$\xi_1 \xi_2 \xi_3 \xi_4$	-	-	Исходное слово
$\alpha \xi_1 \xi_2 \xi_3 \xi_4$	2	1	Этап движения вправо
$\xi_2 \alpha \xi_1 \xi_3 \xi_4$	1	2	
$\xi_2 \xi_3 \alpha \xi_1 \xi_4$	1	3	
$\xi_2 \xi_3 \xi_4 \alpha \xi_1$	1	4	Разметка ξ_1
$\alpha \xi_2 \xi_3 \xi_4 \alpha \xi_1$	2	5	Этап движения вправо
$\xi_3 \alpha \xi_2 \xi_4 \alpha \xi_1$	1	6	
$\xi_3 \xi_4 \alpha \xi_2 \alpha \xi_1$	1	7	Разметка ξ_2
$\alpha \xi_3 \xi_4 \alpha \xi_2 \alpha \xi_1$	2	8	Этап движения вправо
$\xi_4 \alpha \xi_3 \alpha \xi_2 \alpha \xi_1$	1	9	Разметка ξ_3
$\alpha \xi_4 \alpha \xi_3 \alpha \xi_2 \alpha \xi_1$	2	10	Этап движения вправо, разметка ξ_4
$\alpha \alpha \xi_4 \alpha \xi_3 \alpha \xi_2 \alpha \xi_1$	2	11	Результат
....			

В ходе решения на шаге 11 фактически формируется слово и инверсным порядком букв, но результат «зашумлен» метасимволами α , расположенными между буквами инверсного слова. Значит, необходима секция удаления метасимволов α .

Полный алгоритм обращения слова со стандартной схемой управления имеет вид

$$\begin{cases} \alpha \xi \eta \rightarrow \eta \alpha \xi \\ \alpha \alpha \rightarrow \beta \alpha \\ \beta \alpha \rightarrow \beta \\ \beta \xi \rightarrow \xi \beta \\ \beta \rightarrow \bullet \lambda \quad ; \text{stop production} \\ \lambda \rightarrow \alpha \quad ; \text{start production} \end{cases} \quad (19)$$

Между рабочей продукцией ($\alpha \xi \eta \rightarrow \eta \alpha \xi$) и стартовой продукцией размещены продукции для удаления метасимволов α с помощью вводимого челнока β .

Каждый проход челнока α , вводимого стартовой продукцией, связан с избыточными затратами времени на холостые срабатывания первых 5 продуктов, поэтому данный алгоритм не является эффективным по времени по количеству проходов по слову и по числу шагов внутри каждого прохода.

Модификация ПС с двумя челноками для обращения слова состоит из 4-х секций (*start, work, change, final sections*), расположенных друг за другом

$$\left\{ \begin{array}{l} \lambda \rightarrow * \alpha_1 \\ \beta_1^* \leftarrow \lambda \\ \left\| \begin{array}{l} \alpha \xi_i \xi_{i''} \rightarrow \xi_{i''} \alpha \xi_i \\ \xi_j \beta \xi_{j''} \leftarrow \xi_{j''} \xi_j \beta \end{array} \right\| (\alpha \dots \beta) \leq 3 \\ \alpha \xi_i \xi_j \beta \rightarrow \xi_j \beta \alpha \xi_i \\ \alpha \xi_i \xi_k \xi_j \beta \rightarrow \xi_j \beta \xi_i \alpha \xi_k \\ \alpha_w \rightarrow \beta_{w+1} \\ \beta_w \rightarrow \alpha_{w+1} \\ \beta \alpha \mid \beta \xi_i \alpha \rightarrow \bullet \lambda \end{array} \right. \quad (20)$$

Продукции рабочей секции параллельно срабатывают, при этом вхождение образцов подчиняется соотношениям ($i < i''$) и ($j > j''$). Охранное условие контролирует сближение двух челноков на расстояние конфликта и передает управление на следующую секцию – обменную (*change section*). Две первые продукции из *change section* осуществляют взаимный обмен челноками и принадлежащими им символами, что разрешает конфликт и позволяет продолжить работу ПС.

Две последние продукции из *change sections* моделируют обмен челноками для следующего прохода по слову S (с индексацией номера цикла обхода w). После их срабатывания выполняется переход на рабочую секцию – *work section*.

В табл. 4 и 5 показана трассировка работы модификации ПС обращения слова при $n=6$ и $n=9$. Серым цветом выделены обнаруживаемые охранным условием потенциальные конфликты данных.

Таблица 4

Трассировка модификации ПС обращения слова при $n=6$

Обрабатываемое слово S	Номер сработавшей продукции	Номер шага алгоритма	Примечание
$\xi_1 \xi_2 \xi_3 \xi_4 \xi_5 \xi_6$	-	-	Исходное слово
$\alpha \xi_1 \xi_2 \xi_3 \xi_4 \xi_5 \xi_6 \beta$	1,2	1,2	Разметка ξ_1, ξ_6 цикл обхода слова, $w=1$
$\xi_2 \alpha \xi_1 \xi_3 \xi_4 \xi_6 \beta \xi_5$	3,4	3	Параллельная работа рабочих продуктов
$\xi_2 \xi_3 \alpha \xi_1 \xi_6 \beta \xi_4 \xi_5$	5	4	Охранное условие, обменная продукция
$\xi_2 \xi_3 \xi_6 \beta \alpha \xi_1 \xi_4 \xi_5$	3,4	5	Параллельная работа рабочих продуктов
$\xi_2 \xi_6 \beta \xi_3 \xi_4 \alpha \xi_1 \xi_5$	3,4	6	Параллельная работа рабочих продуктов
$\xi_6 \beta \xi_2 \xi_3 \xi_4 \xi_5 \alpha \xi_1$	7,8	7	Разметка ξ_2, ξ_5 , цикл обхода слова, $w=2$
↓ ↓			
$\xi_6 \alpha \xi_2 \xi_3 \xi_4 \xi_5 \beta \xi_1$	3,4	8	Параллельная работа рабочих продуктов
$\xi_6 \xi_3 \alpha \xi_2 \xi_5 \beta \xi_4 \xi_1$	5	9	Охранное условие, обменная продукция

$\xi_6 \xi_3 \xi_5 \beta \alpha \xi_2 \xi_4 \xi_1$	3,4	10	Параллельная работа рабочих продуктов
$\xi_6 \xi_5 \beta \xi_3 \xi_4 \alpha \xi_2 \xi_1$	7,8	11	Разметка ξ_3, ξ_4 , цикл обхода слова, $w=3$
↓ ↓			
$\xi_6 \xi_5 \alpha \xi_3 \xi_4 \beta \xi_2 \xi_1$	5	12	Охранное условие, обменная продукция
$\xi_6 \xi_5 \xi_4 \beta \alpha \xi_3 \xi_2 \xi_1$	9	13	Заключительная продукция
$\xi_6 \xi_5 \xi_3 \xi_4 \xi_2 \xi_1$	-	14	Результат

Таблица 5

Трассировка модификации ПС обращения слова при $n=9$

Обрабатываемое слово S	Номер сработавшей продукции	Номер шага алгоритма	Примечание
$\xi_1 \xi_2 \xi_3 \xi_4 \xi_5 \xi_6 \xi_7 \xi_8 \xi_9$	-	-	Исходное слово
$\alpha \xi_1 \xi_2 \xi_3 \xi_4 \xi_5 \xi_6 \xi_7 \xi_8 \xi_9 \beta$	1,2	1,2	Разметка ξ_1, ξ_9 цикл обхода слова, $w=1$
$\xi_2 \alpha \xi_1 \xi_3 \xi_4 \xi_5 \xi_6 \xi_7 \xi_9 \beta \xi_8$	3,4	3	Параллельная работа рабочих продуктов
$\xi_2 \xi_3 \alpha \xi_1 \xi_4 \xi_5 \xi_6 \xi_9 \beta \xi_7 \xi_8$	3,4	4	Параллельная работа рабочих продуктов
$\xi_2 \xi_3 \xi_4 \alpha \xi_1 \xi_5 \xi_9 \beta \xi_6 \xi_7 \xi_8$	6	5	Охранное условие, обменная продукция
$\xi_2 \xi_3 \xi_4 \xi_9 \beta \xi_5 \alpha \xi_1 \xi_6 \xi_7 \xi_8$	3,4	6	Параллельная работа рабочих продуктов
$\xi_2 \xi_3 \xi_9 \beta \xi_4 \xi_5 \xi_6 \alpha \xi_1 \xi_7 \xi_8$	3,4	7	Параллельная работа рабочих продуктов
$\xi_2 \xi_9 \beta \xi_3 \xi_4 \xi_5 \xi_6 \xi_7 \alpha \xi_1 \xi_8$	3,4	8	Параллельная работа рабочих продуктов
$\xi_9 \beta \xi_2 \xi_3 \xi_4 \xi_5 \xi_6 \xi_7 \xi_8 \alpha \xi_1$	7,8	9	Разметка ξ_2, ξ_8 , цикл обхода слова, $w=2$
↓ ↓			
$\xi_9 \alpha \xi_2 \xi_3 \xi_4 \xi_5 \xi_6 \xi_7 \xi_8 \beta \xi_1$	3,4	8	Параллельная работа рабочих продуктов
$\xi_9 \xi_3 \alpha \xi_2 \xi_4 \xi_5 \xi_6 \xi_8 \beta \xi_7 \xi_1$	3,4	10	Параллельная работа рабочих продуктов
$\xi_9 \xi_3 \xi_4 \alpha \xi_2 \xi_5 \xi_8 \beta \xi_6 \xi_7 \xi_1$	6	11	Охранное условие, обменная продукция
$\xi_9 \xi_3 \xi_4 \xi_8 \beta \xi_5 \alpha \xi_2 \xi_6 \xi_7 \xi_1$	3,4	12	Параллельная работа рабочих продуктов
$\xi_9 \xi_3 \xi_8 \beta \xi_4 \xi_5 \xi_6 \alpha \xi_2 \xi_7 \xi_1$	3,4	13	Параллельная работа рабочих продуктов
$\xi_9 \xi_8 \beta \xi_3 \xi_4 \xi_5 \xi_6 \xi_7 \alpha \xi_2 \xi_1$	7,8	14	Разметка ξ_3, ξ_7 , цикл обхода слова, $w=3$
↓ ↓			

$\xi_9 \xi_8 \alpha \xi_3 \xi_4 \xi_5 \xi_6 \xi_7 \beta \xi_2 \xi_1$	3,4	15	Параллельная работа рабочих продукций
$\xi_9 \xi_8 \xi_4 \alpha \xi_3 \xi_5 \xi_7 \beta \xi_6 \xi_2 \xi_1$	6	16	Охранное условие, обменная продукция
$\xi_9 \xi_8 \xi_4 \xi_7 \beta \xi_5 \alpha \xi_3 \xi_6 \xi_2 \xi_1$	3,4	17	Параллельная работа рабочих продукций
$\xi_9 \xi_8 \xi_7 \beta \xi_4 \xi_5 \xi_6 \alpha \xi_3 \xi_2 \xi_1$	7,8	18	Разметка ξ_4, ξ_6 , цикл обхода слова, $w=4$
↓ ↓			
$\xi_9 \xi_8 \xi_7 \alpha \xi_4 \xi_5 \xi_6 \beta \xi_3 \xi_2 \xi_1$	6	19	Охранное условие, обменная продукция
$\xi_9 \xi_8 \xi_7 \xi_6 \beta \xi_5 \alpha \xi_4 \xi_3 \xi_2 \xi_1$	9	20	Заключительная продукция
$\xi_9 \xi_8 \xi_7 \xi_6 \xi_5 \xi_4 \xi_3 \xi_2 \xi_1$	-	21	Результат

На рис. 6 показаны временные зависимости обращения слова стандартной и модифицированной ПС в зависимости от длины слова n . Моделирование показало уменьшение времени обработки слова практически в 2 раза, что подтверждает эффективность созданной математической модели и схемы управления производствами.

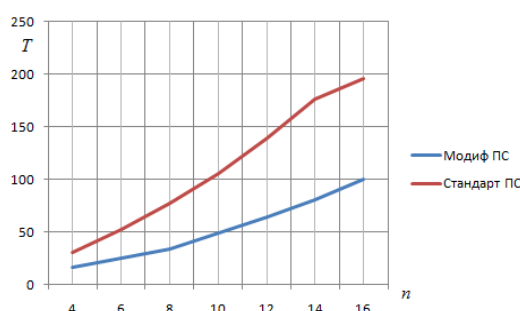


Рис. 6. Временные зависимости работы стандартной и модифицированной ПС для задачи обращения слова

Выводы. Созданы математическая модель и схема управления алгоритмической продукционной системой, направленные на организацию последовательно-параллельных вычислений. Известные типы продукции в математической модели дополнены продуктами с однократным срабатыванием, введена двунаправленная обработка слова с помощью двух встречно движущихся челноков, указывающих на независимые локальные фрагменты слов для их параллельной обработки. Корректность вычислений основана на алгебре продукции и обеспечивается проверкой введенного охранного условия, проверяющего сближение двух челноков на минимальное расстояние, что обеспечивает разрешение конфликта по данным между параллельно выполняющимися продуктами. Введены оригинальные обменные продукты, позволяющие реализовать одновременную работу двух продукции при минимальном расстоянии между двумя челноками.

Созданная схема управления, отличается объединением продукции в смысловые секции. Выделены последовательно располагаемые стартовая, рабочая, обменная и финальная секции. Вместо обязательного возвратного перехода в начало продукционной системы переходы осуществляются внутри секции или на предыдущую секцию, что, в целом, обеспечивает уменьшение времени работы продукционной модели.

Полученные теоретические конструкции (модифицированная математическая модель и схема управления продукциями) позволяют получить и использовать параллельные вычислительные модели для высокопроизводительной обработки битовых потоков, работы баз знаний, параллельных кодовых преобразователей информации, в системах информационной безопасности, в информационно-аналитических и советующих системах, использующих символическую информацию как носитель данных и знаний о предметной области [1, 20–22].

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Финн В.К. Искусственный интеллект: Методология, применения, философия. – 2-е изд. испр. и доп. – М.: ЛЕНАНД, 2021. – 468 с.
2. Гергель В.П., Козинев Е.А. Параллельный подход для решения вычислительно сложных задач многокритериальной оптимизации // Информационный бюллетень Омского научно-образовательного центра ОмГТУ и ИМ СО РАН в области математики и информатики: Матер. X Международной молодежной научно-практической конференции с элементами научной школы. – Т. 4, № 1. – Омск: Омский государственный технический университет, 2020. – С. 21-26.
3. Жданов А.А. Автономный искусственный интеллект: монография. – М.: БИНОМ. Лаборатория знаний, 2024. – 360 с.
4. Воеводин В.В., Воеводин Вл.В. Параллельные вычисления. – М.: BNV, 2004. – 608 с.
5. Добрица В.П., Титенко Е.А., Халин Ю.А., Катыхин А.И. Модели представления и обработки знаний в информационно-аналитических системах: монография. – Курск: ЗАО «Университетская книга», 2023. – 172 с.
6. Рыбина Г.В. Архитектуры современных интеллектуальных систем: синергия кибернетики и символического искусственного интеллекта, инструментальные средства и технологии для разработки интеллектуальных систем // Нейрокомпьютеры: разработка, применение. – 2024. – Т. 26, № 4. – С. 69-82.
7. Решетников И.А. Комбинаторика слов, фактординамика и нормальные формы // Чебышевский сборник. – 2021. – С. 202-235.
8. Левин И.И. Интеллектуальные многопроцессорные системы. Опыт создания и применения // Приоритетные направления развития науки и образования в условиях формирования технологического суверенитета: Матер. международной научно-практической конференции. – Ростов-на-Дону: ДГТУ, 2024. – С. 246-255.
9. Левин И.И., Алексеев К.Н., Гуленок А.А. Преобразование простейших сортирующих сетей к нечетно-четной сети Батчера // Известия ЮФУ. Технические науки. – 2024. – № 4 (240). – С. 50-63. – DOI 10.18522/2311-3103-2024-4-50-63.
10. Титенко Е.А., Талдыкин Е.В. Коммутационная модель параллельных сравнений элементов для производственных систем, управляемых потоком данных // Известия ЮФУ. Технические науки. – 2020. – № 7 (217). – С. 19-34.
11. Гурин А.Г., Григорьев А.В. Сравнительный анализ методов повышения эффективности производственных систем для работы с большими данными // Проблемы искусственного интеллекта. – 2025. – № 4.
12. Судоплатов С.В., Овчинникова Е.В. Математическая логика и теория алгоритмов: учебник и практикум для вузов. – 5-е изд., стер. – М.: Изд-во Юрайт, 2026. – 207 с.
13. Марков А.А. Избранные труды. Т. 2: Теория алгоритмов и конструктивная математика, математическая логика, информатика и смежные вопросы. – М.: МЦНМО, 2007. – 650 с.
14. Малявко А.А. Формальные языки и компиляторы: учеб. пособие для вузов. – М.: Изд-во ЮРАЙТ, 2020. – 429 с.
15. Карлов Б.Н. О разрешимых и неразрешимых теориях языков // Вестник Тверского государственного университета. Серия: Прикладная математика. – 2024. – № 4. – С. 30-39.
16. Гашиков С.Б. Теория алгоритмов и вычислений. – СПб.: Лань, 2023. – 168 с.
17. Пентус А.Е., Пентус М.Р. Математическая теория формальных языков: учеб. пособие. – М.: ИНТУИТ, IPR Медиа, 2020.
18. Ляхова Е.Г., Сахаров Ю.А., Коваленко Н.Ю., Филичева О.С. Искусственный интеллект в языковом пространстве: монография / сост. Ю.А. Сахаров. – М.: ИП Воронцов М.Ю., 2025. – 152 с.
19. Титенко Е.А., Ватутин Э.И., Титенко М.А. [и др.]. Аппаратно-ориентированный метод ускоренного поиска вхождений образца на основе структурно-процедурных вычислений // Известия ЮФУ. Технические науки. – 2024. – № 5 (241). – С. 68-78.
20. Пекунов В.В. Индукция правил трансформации естественно-языковой постановки задачи в смысловую модель порождения решающей программы // Программные системы и вычислительные методы. – 2020. – № 3. – С. 29-39.

21. Титенко Е.А., Ватутин Э.И., Чернецкая И.Е., Трокоз Д.А., Титенко М.А., Пузанов И.И. Мультиконвейерные каналные схемы для обработки унарных кодов в высокопроизводительных вычислительных системах // Вестник Рязанского государственного радиотехнического университета. – 2025. – № 93. – С. 61-71.
22. Атакищев О.И., Грибулин В.Г., Ананьев В.Е., Титенко Е.А. Реконфигурационные метаграмматики для описания и моделирования многоэтапных комплексных атак // Известия ЮФУ. Технические науки. – 2022. – № 5 (229). – С. 189-201.

REFERENCES

1. Finn V.K. *Iskusstvennyy intellekt: Metodologiya, primeneniya, filosofiya* [Artificial intelligence: Methodology, applications, and philosophy]. 2nd ed. corr. and add. Moscow: LENAND. 2021, 468 p.
2. Gergel' V.P., Kozinov E.A. *Parallelnyy podkhod dlya resheniya vychislitel'no slozhnykh zadach mnogokriterial'noy optimizatsii* [Parallel approach for solving computationally complex multicriteria optimization problems], *Informatsionnyy byulleten' Omskogo nauchno-obrazovatel'nogo tsentra OmGTU i IM SO RAN v oblasti matematiki i informatiki: Mater. X Mezhdunarodnoy molodezhnoy nauchno-prakticheskoy konferentsii s elementami nauchnoy shkoly* [Information Bulletin of the Omsk Scientific and Educational Center of OmskSTU and the Institute of Mathematics of the Siberian Branch of the Russian Academy of Sciences in Mathematics and Computer Science: Proceedings of the X International Youth Scientific and Practical Conference with Elements of a Scientific School], Vol. 4, No. 1. Omsk: Omskiy gosudarstvennyy tekhnicheskii universitet, 2020, pp. 21-26.
3. Zhdanov A.A. *Avtonomnyy iskusstvennyy intellekt: monografiya* [Autonomous artificial intelligence: monograph]. Moscow: BINOM. Laboratoriya znaniy, 2024, 360 p.
4. Voevodin V.V., Voevodin V.V. *Parallelnyye vychisleniya* [Parallel computing]. Moscow: BHV, 2004, 608 p.
5. Dobrica V.P., Titenko E.A., Khalin Yu.A., Katykhin A.I. *Modeli predstavleniya i obrabotki znaniy v informacionno-analiticheskikh sistemakh: monografiya* [Models of knowledge representation and processing in information and analytical systems: monograph]. Kursk: ZAO «Universitetskaya kniga», 2023, 172 p.
6. Rybina G.V. *Arkhitektury sovremennykh intellektual'nykh sistem: sinergiya kibernetiki i simvol'nogo iskusstvennogo intellekta, instrumental'nye sredstva i tekhnologii dlya razrabotki intellektual'nykh sistem* [Architectures of modern intelligent systems: synergy of cybernetics and symbolic artificial intelligence, tools and technologies for the development of intelligent systems], *Neyrokomp'yutery: razrabotka, primeneniye* [Neurocomputers: Development, Application], 2024, Vol. 26, No. 4, pp. 69-82.
7. Reshetnikov I.A. *Kombinatorika slov, faktordinamika i normal'nye formy* [Combinatorics of words, factor dynamics, and normal forms], *Chebyshevskiy sbornik* [Chebyshevsky Collection], 2021, pp. 202-235.
8. Levin I.I. *Intellektual'nye mnogoprocessornyye sistemy. Opyt sozdaniya i primeneniya* [Intelligent multiprocessor systems. Experience of creation and application], *Prioritetnye napravleniya razvitiya nauki i obrazovaniya v usloviyakh formirovaniya tekhnologicheskogo suvereniteta: Mater. mezhdunarodnoy nauchno-prakticheskoy konferentsii* [Priority Directions for the Development of Science and Education in the Context of the Formation of Technological Sovereignty: Proceedings of the International Scientific and Practical Conference]. Rostov-on-Don: DGTU, 2024, pp. 246-255.
9. Levin I.I., Alekseev K.N., Gulenok A.A. *Preobrazovanie prosteyshikh sortiruyushchikh setey k nechetno-chetnoy seti Batchera* [Transformation of the simplest sorting networks to an odd-even Butcher network], *Izvestiya YuFU. Tekhnicheskie nauki* [Izvestiya SFedU. Engineering Sciences], 2024, No. 4 (240), pp. 50-63. DOI 10.18522/2311-3103-2024-4-50-63.
10. Titenko E.A., Taldykin E.V. *Kommutatsionnaya model' parallelnykh sravneniy elementov dlya produktsionnykh sistem, upravlyayemykh potokom dannykh* [Switching model of parallel comparisons of elements for data-flow-controlled production systems], *Izvestiya YuFU. Tekhnicheskie nauki* [Izvestiya SFedU. Engineering Sciences], 2020, No. 7 (217), pp. 19-34.
11. Gurin A.G., Grigor'ev A.V. *Sravnitel'nyy analiz metodov povysheniya effektivnosti produktsionnykh sistem dlya raboty s bol'shimi dannymi* [Comparative analysis of methods for improving the efficiency of production systems for working with big data], *Problemy iskusstvennogo intellekta* [Problems of Artificial Intelligence], 2025, No. 4.
12. Sudoplatov S.V., Ovchinnikova E.V. *Matematicheskaya logika i teoriya algoritmov: uchebnik i praktikum dlya vuzov* [Mathematical logic and theory of algorithms: textbook and workshop for universities]. 5th ed., reprinted. Moscow: Izd-vo Yurayt, 2026, 207 p.
13. Markov A.A. *Izbrannyye trudy. T. 2: Teoriya algorifmov i konstruktivnaya matematika, matematicheskaya logika, informatika i smezhnye voprosy* [Selected works. Vol. 2: Theory of algorithms and constructive mathematics, mathematical logic, computer science, and related topics]. Moscow: MCNMO, 2007, 650 p.
14. Malyavko A.A. *Formal'nye yazyki i kompilyatory: ucheb. posobie dlya vuzov* [Formal Languages and Compilers: manual for universities]. Moscow: Izd-vo YuRAYT, 2020, 429 p.

15. Karlov B.N. O razreshimyykh i nerazreshimyykh teoriyakh yazykov [On decidable and undecidable theories of languages], *Vestnik Tverskogo gosudarstvennogo universiteta. Seriya: Prikladnaya matematika* [Bulletin of Tver State University. Series: Applied Mathematics], 2024, No. 4, pp. 30-39.
16. Gashkov S.B. Teoriya algoritmov i vychisleniy [Theory of Algorithms and Computations]. Saint Petersburg: Lan', 2023, 168 p.
17. Pentus A.E., Pentus M.R. Matematicheskaya teoriya formal'nykh yazykov: ucheb. posobie [Mathematical theory of formal languages: a manual]. Moscow: INTUIT, IPR Media, 2020.
18. Lyakhova E.G., Sakharov Yu.A., Kovalenko N.Yu., Filicheva O.S. Iskusstvennyy intellekt v yazykovom prostranstve: monografiya [Artificial intelligence in the language space: monograph], comp. Yu.A. Sakharov. Moscow: IP Voroncov M.Yu., 2025, 152 p.
19. Titenko E.A., Vatutin E.I., Titenko M.A. [et al.]. Apparato-orientirovannyi metod uskorenogo poiska vkhozheniy obrazca na osnove strukturno-procedurnykh vychisleniy [Hardware method for accelerated search for pattern occurrences based on structural-procedural calculations], *Izvestiya YuFU. Tekhnicheskie nauki* [Izvestiya SFedU. Engineering Sciences], 2024, No. 5 (241), pp. 68-78.
20. Pekunov V.V. Indukciya pravil transformatsii estestvenno-yazykovoy postanovki zadachi v smyslovuyu model' porozhdeniya reshayushchey programmy [Induction of rules for transforming a natural-language statement of a problem into a semantic model for generating a solver], *Programmnye sistemy i vychislitel'nye metody* [Software Systems and Computational Methods], 2020, No. 3, pp. 29-39.
21. Titenko E.A., Vatutin E.I., Cherneckaya I.E., Trokoz D.A., Titenko M.A., Puzanov I.I. Mul'tikonveyernye kanal'nye skhemy dlya obrabotki unarnykh kodov v vysokoproizvoditel'nykh vychislitel'nykh sistemakh [Multi-pipeline channel schemes for processing unary codes in high-performance computing systems], *Vestnik Ryazanskogo gosudarstvennogo radiotekhnicheskogo universiteta* [Bulletin of Ryazan State Radio Engineering University], 2025, No. 93, pp. 61-71.
22. Atakishchev O.I., Gribulin V.G., Anan'ev V.E., Titenko E.A. Rekonfiguratsionnye metagrammatiki dlya opisaniya i modelirovaniya mnogoetapnykh kompleksnykh atak [Reconfiguration metagrammars for describing and modeling multi-stage complex attacks], *Izvestiya YuFU. Tekhnicheskie nauki* [Izvestiya SFedU. Engineering Sciences], 2022, No. 5 (229), pp. 189-201.

Титенко Евгений Анатольевич – Юго-Западный государственный университет; e-mail: johntit@mail.ru; г. Курск, Россия; к.т.н.; доцент кафедры программной инженерии.

Белова Татьяна Михайловна – Юго-Западный государственный университет; e-mail: tm_belova@mail.ru; г. Курск, Россия; к.т.н.; доцент кафедры программной инженерии.

Пузанов Иван Иванович – Юго-Западный государственный университет; e-mail: ivan.puzanov2000@yandex.ru; г. Курск, Россия; аспирант.

Положенцев Алексей Анатольевич – Юго-Западный государственный университет; e-mail: polojencev135@mail.ru; г. Курск, Россия; студент.

Лисицин Леонид Александрович – Юго-Западный государственный университет; e-mail: leo_263@mail.ru; г. Курск, Россия; к.т.н.; доцент кафедры программной инженерии.

Titenko Evgeny Anatolyevich – South-West State University; e-mail: johntit@mail.ru; Kursk, Russia; cand. of eng. sc.; associate professor of the Department of Program Engineering.

Belova Tatyana Mikhailovna – South-West State University; e-mail: tm_belova@mail.ru; Kursk, Russia; cand. of eng. sc.; associate professor of the Department of Program Engineering.

Puzanov Ivan Ivanovich – South-West State University; e-mail: ivan.puzanov2000@yandex.ru; Kursk, Russia; graduate student.

Polozhenets Aleksey Anatolyevich – South-West State University; e-mail: polojencev135@mail.ru; Kursk, Russia; student.

Lisitsin Leonid Alexandrovich – South-West State University; e-mail: leo_263@mail.ru; Kursk, Russia; cand. of eng. sc.; associate professor of the Department of Program Engineering.

ПРАВИЛА ОФОРМЛЕНИЯ РУКОПИСЕЙ

С полными требованиями к рукописи, перечнем сопроводительных документов и образцом оформления статьи можно ознакомиться на сайте журнала:

https://izv-tn.tti.sfedu.ru/index.php/izv_tn/authorrequirements.

1. Объём статьи – от 12 до 20 страниц формата А4, включая рисунки, таблицы и список литературы (для обзорных статей – от 20 до 30 страниц). Основной текст статьи должен быть набран шрифтом Times New Roman размером 10 пунктов с одинарным интервалом.

2. Названию статьи предшествует индекс УДК, соответствующий заявленной теме.

3. Текст статьи начинается с названия статьи (на русском и английском языках), фамилии, имени и отчества автора (полностью) и снабжается аннотацией на русском и английском языках объёмом **200-250 слов**.

Аннотация должна содержать: актуальность исследования (постановка проблемы); цель и задачи работы; используемые методы и подходы; – основные результаты (с конкретикой, без общих формулировок); выводы и значимость работы.

4. Статья должна быть структурирована и включать: введение, постановку задачи, методику исследования, результаты и их обсуждение (или вычислительный эксперимент), выводы.

5. Все графические материалы (графики, диаграммы, фотографии) должны быть пронумерованы и снабжены подписями. Таблицы также нумеруются и имеют названия. Формулы должны быть набраны в редакторе формул.

6. Наличие библиографического списка на русском и английском языках обязательно. Все источники указываются в тексте в порядке возрастания. Должно быть не менее **20 источников** (для обзорных статей – не менее 40): 75% источников – не старше 10 лет; 25% источников – допускаются старше 10 лет (классические/базовые работы). Ссылки оформляются по ГОСТ Р 7.0.5-2008, в тексте заключаются в квадратные скобки.

7. Оригинальность текста статьи должна составлять не менее 70% по результатам проверки в системе антиплагиата. Объём самоцитирования не должен превышать 20% от общего объёма статьи. Автор обязан предоставить справку-отчёт системы антиплагиат (желательно **antiplagiat.ru**).

8. Статья сопровождается сведениями об авторе(ах) (фамилия, имя, отчество, учёное звание, должность, место работы, адрес, электронный адрес и номер телефона) на русском и английском языках, а также необходимыми сопроводительными документами в соответствии с требованиями журнала.

9. Срок рассмотрения статьи, соответствующей всем требованиям, составляет до 3 месяцев, с момента направления статьи на рецензирование.

10. Публикация в журнале для авторов бесплатна. Редакция не взимает плату с авторов за подготовку, размещение и печать материалов.

Статьи, не соответствующие требованиям журнала, возвращаются авторам на доработку до направления на рецензирование.