

ОБЗОРНАЯ СТАТЬЯ

REVIEW PAPER

doi: 10.17586/2226-1494-2025-25-3-373-386

УДК 004.8

**Объяснимость и интерпретируемость — важные аспекты безопасности
решений, принимаемых интеллектуальными системами
(обзорная статья)**

Денис Николаевич Бирюков¹, Андрей Сергеевич Дудкин²

^{1,2} Военно-космическая академия имени А.Ф. Можайского, Санкт-Петербург, 197198, Российская Федерация

¹ Biryukov.D.N@yandex.ru, <https://orcid.org/0000-0003-1300-2470>

² andry-ll@mail.ru, <https://orcid.org/0000-0003-0283-9048>

Аннотация

Вопросы доверия к решениям, принимаемыми (формируемыми) интеллектуальными системами, становятся все более актуальными. Представлен систематический обзор методов и инструментов объяснимого искусственного интеллекта (Explainable Artificial Intelligence, XAI), направленных на преодоление разрыва между сложностью нейронных сетей и потребностью в интерпретируемости результатов для конечных пользователей. Проведен теоретический анализ различий между объяснимостью и интерпретируемостью в контексте искусственного интеллекта, а также их роли в обеспечении безопасности решений, принимаемых интеллектуальными системами. Показано, что объяснимость подразумевает способность системы генерировать понятные человеку обоснования, тогда как интерпретируемость сосредоточена на пассивной понятности внутренних механизмов. Предложена классификация методов XAI на основе их подхода (предварительный/последующий анализ: ante hoc/post hoc) и масштаба объяснений (локальный/глобальный). Рассмотрены популярные инструменты, такие как Local Interpretable Model Agnostic Explanations, Shapley Values и интегрированные градиенты, с оценкой их сильных сторон и ограничений применимости. Даны практические рекомендации по выбору методов для различных областей и сценариев. Обсуждается архитектура интеллектуальной системы, построенной на основе модели В.К. Финна, и адаптированной к современным требованиям к обеспечению «прозрачности» решений, где ключевыми компонентами являются информационная среда, решатель задач и интеллектуальный интерфейс. Рассмотрена проблема компромисса между точностью моделей и их объяснимостью: прозрачные модели («стеклянные ящики», например, деревья решений) уступают в производительности глубоким нейронным сетям, но обеспечивают большую бесспорность принятия решений. Приведены примеры методов и программных пакетов для объяснения и интерпретации данных и моделей машинного обучения. Показано, что развитие XAI связано с интеграцией нейро-символических подходов, объединяющих возможности глубокого обучения с логической интерпретируемостью.

Ключевые слова

искусственный интеллект, нейронные сети, глубокое обучение, «черный ящик», объяснимость, интерпретируемость, XAI

Ссылка для цитирования: Бирюков Д.Н., Дудкин А.С. Объяснимость и интерпретируемость – важные аспекты безопасности решений, принимаемых интеллектуальными системами (обзорная статья) // Научно-технический вестник информационных технологий, механики и оптики. 2025. Т. 25, № 3. С. 373–386. doi: 10.17586/2226-1494-2025-25-3-373-386

**Explainability and interpretability are important aspects in ensuring
the security of decisions made by intelligent systems
(review article)**

Denis N. Biryukov¹, Andrey S. Dudkin²

^{1,2} Mozhaisky Military Aerospace Academy, Saint Petersburg, 197198, Russian Federation

¹ Biryukov.D.N@yandex.ru, <https://orcid.org/0000-0003-1300-2470>

² andry-ll@mail.ru, <https://orcid.org/0000-0003-0283-9048>

© Бирюков Д.Н., Дудкин А.С., 2025

Abstract

The issues of trust in decisions made (formed) by intelligent systems are becoming more and more relevant. A systematic review of Explicable Artificial Intelligence (XAI) methods and tools aimed at bridging the gap between the complexity of neural networks and the need for interpretability of results for end users is presented. A theoretical analysis of the differences between explainability and interpretability in the context of artificial intelligence as well as their role in ensuring the security of decisions made by intelligent systems is carried out. It is shown that explainability implies the ability of a system to generate justifications understandable to humans, whereas interpretability focuses on the passive clarity of internal mechanisms. A classification of XAI methods is proposed based on their approach (preliminary/subsequent analysis: ante hoc/post hoc) and the scale of explanations (local/global). Popular tools, such as Local Interpretable Model Agnostic Explanations, Shapley Values, and integrated gradients, are considered, with an assessment of their strengths and limitations of applicability. Practical recommendations are given on the choice of methods for various fields and scenarios. The architecture of an intelligent system based on the V.K. Finn model and adapted to modern requirements for ensuring “transparency” of solutions, where the key components are the information environment, the problem solver and the intelligent interface, are discussed. The problem of a compromise between the accuracy of models and their explainability is considered: transparent models (“glass boxes”, for example, decision trees) are inferior in performance to deep neural networks, but provide greater certainty of decision-making. Examples of methods and software packages for explaining and interpreting machine learning data and models are provided. It is shown that the development of XAI is associated with the integration of neuro-symbolic approaches combining deep learning capabilities with logical interpretability.

Keywords

artificial intelligence, neural networks, deep learning, black box, explainability, interpretability, XAI

For citation: Biryukov D.N., Dudkin A.S. Explainability and interpretability are important aspects in ensuring the security of decisions made by intelligent systems (review article). *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2025, vol. 25, no. 3, pp. 373–386 (in Russian). doi: 10.17586/2226-1494-2025-25-3-373-386

Введение

На современном этапе развития наибольшую популярность получили результаты в области моделирования глубокого обучения (Deep Learning, DL), дающие прирост в скорости и точности формируемых решений. Однако у этих результатов есть и обратная сторона — сложность моделей, что зачастую приводит к невозможности понимания того, как и за счет чего были получены те или иные решения, какие входные данные являются наиболее важными и критичными, и какие параметры модели наиболее уязвимы.

Еще предстоит много сделать, чтобы разработать подходы и инструменты, позволяющие укрепить доверие к системам и моделям DL. Возможно, в ряде случаев в обмен на понимание сути придется полагаться скоростью и/или точностью формируемых решений. Работа в данном направлении уже ведется, но пока что наблюдается отсутствие единого тезауруса и разнообразие во взглядах на подходы к решению задач в обозначенной области исследований.

Не заканчиваются споры о том, какую систему можно считать интеллектуальной и что же такое искусственный интеллект (ИИ), но мало кто спорит с тем, что важным качеством интеллектуальной системы (ИС) является ее способность объяснить потребителю на понятном ему языке, почему и как она пришла к тому или иному решению. Только в этом случае пользователь сможет оценить результат, полученный ИС, и обоснованно принять решение. Ввиду этого видится важным рассмотреть различные взгляды на «объяснимость» и «интерпретируемость» применительно к моделям машинного обучения и к ИС в целом.

Общая модель ИС

Вопрос, связанный с тем, какую систему можно считать интеллектуальной, остается открытым до сих

пор. Ввиду этого, предлагается под ИС в общем случае понимать такую систему, которая соответствует положениям, сформулированным В.К. Финном [1, 2].

ИС обладают специфической архитектурой, допускающей определенные вариации. Схематически эта архитектура может быть представлена следующим образом [2, 3] (рисунок):

ИС = (1) решатель задач + (2) информационная среда + (3) интеллектуальный интерфейс.

(1) Решатель задач = (1.1) рассуждатель + (1.2) вычислитель + (1.3) синтезатор.

(1.1) Рассуждатель реализует синтез и взаимодействие познавательных процедур, образующих автоматизированное рассуждение, область применимости которого является класс задач, решаемых посредством формализованной эвристики. Логическим средством формализации этой эвристики и являются рассуждения. Важно понимать, что правдоподобные рассуждения, формализующие эвристику решения задач, адекватные цели применения ИС, являются основным инструментом ее решателя задач, реализуемым в рассуждателе [1].

Существуют два типа рассуждателей.

Рассуждатели первого типа применимы к неизменяющемуся множеству исходных высказываний, характеризующих «замкнутый мир», а их логическим средством являются дедуктивные выводы (возможны варианты, когда используются неклассические логики).

Рассуждатели второго типа реализуют формализованные эвристики для решения классов задач, исходными данными которых являются изменяемые и пополняемые множества высказываний (под изменением высказываний понимается пересмотр его истинностного значения, соответствующие базы фактов называют эпистемическими). Формализованные эвристики этого типа осуществляют синтез познавательных процедур, включающих эмпирическую индукцию, основанную на установлении сходства фактов, и абдукцию, по-

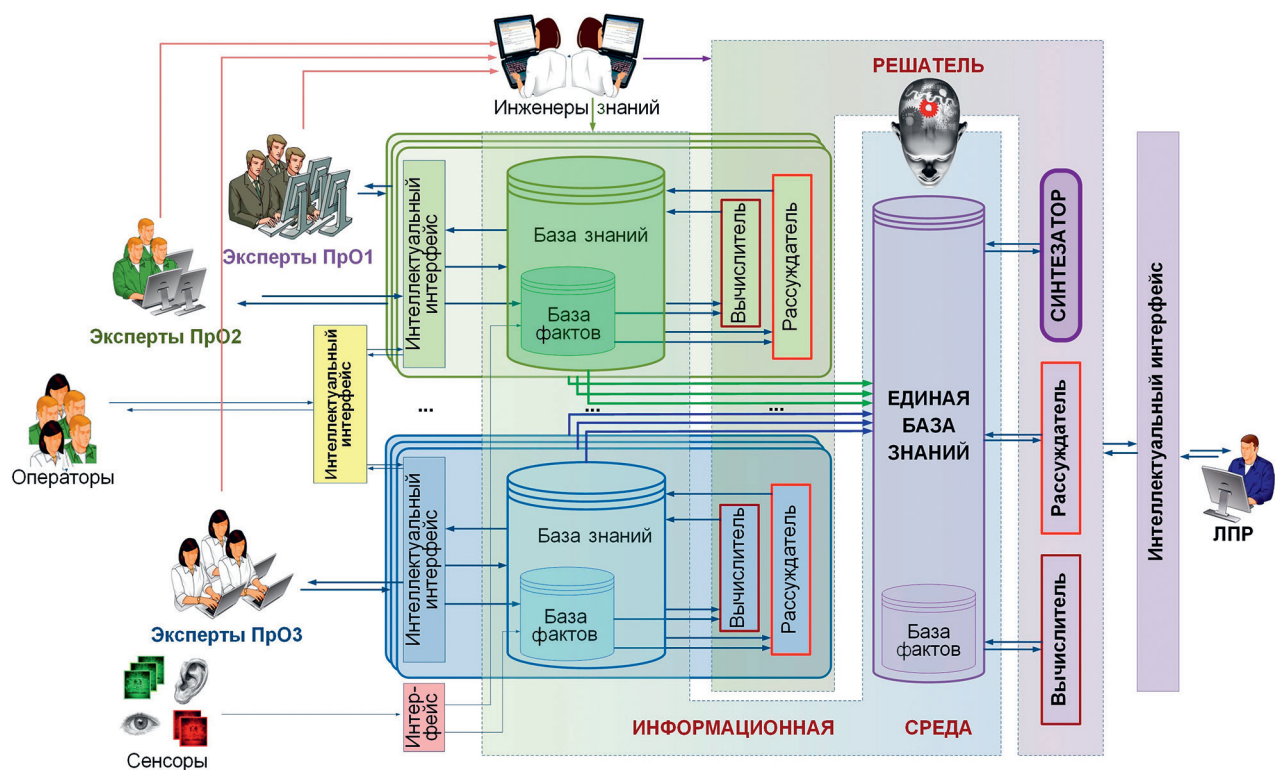


Рисунок. Структурно-функциональная схема интеллектуальной системы.

ПрО1 — предметная область 1; ПрО2 — предметная область 2; ПрО3 — предметная область 3; ЛПР — лицо, принимающее решения

Figure. Structural and functional diagram of the intelligent system.

Subject area 1 — ПрО1; subject area 2 — ПрО2; subject area 3 — ПрО3; the decision maker — ЛПР

средством которой объясняется начальное состояние базы фактов и, в случае необходимости реализуется ее пополнение (рассуждатель второго типа применимы к «открытым мирам»). Рассуждения рассуждателей второго типа называют когнитивными (правдоподобными) рассуждениями [2].

(1.2) Вычислитель применяется к числовым данным, используя численные методы, релевантные целям ИС (например, различные статистические методы анализа данных).

(1.3) Синтезатор выбирает стратегии, адекватные не только цели ИС, но и состоянию базы факторов (БФ), и результатам предыдущих применений решателя задач.

Если через Γ — обозначить множество правил вывода, содержащих как правила достоверного вывода, так и правила правдоподобных выводов, посредством которых осуществляются различные правдоподобные рассуждения, а через Σ — множество вычислительных процедур, то их комбинирование осуществляет синтезатор.

(2) Информационная среда = (2.1) БФ + (2.2) база знаний (БЗ).

(2.1) БФ представляет рассматриваемую предметную область («замкнутый мир» или «открытый мир»; в первом случае БФ не изменяется, во втором — возможно ее пополнение в соответствии с результатами, полученными решателем задач, и желаниями пользователя ИС как человеко-машинной системы).

Наличие БФ как подсистемы ИС создает возможность осуществления машинного обучения, а следовательно, расширения БЗ.

(2.2) БЗ — подсистема представления знаний.

Обычно выделяют три типа знаний для ИС: декларативные, процедурные и концептуальные [1].

(3) Интеллектуальный интерфейс включает в себя диалог (наилучший вариант — диалог на естественном языке), демонстрацию как результатов работы ИС, так и процесса их получения, графическое представление результатов, обучение пользователя работе с ИС, поддержку интерактивного режима работы ИС.

Рассуждения и вычисления, представление знаний и интерфейс являются практическими реализациями принципов функционирования ИС. Посредством этих компонент функционирования ИС осуществляется интеллектуальная обработка данных.

Для того чтобы повысить степень доверия к ИС, как к системе, способствующей принятию решений лицом, принимающим решения (ЛПР), необходимо иметь возможность понимать то, каким образом и почему ИС пришла к тому или иному решению [4, 5]. Для этого необходимо, чтобы была возможность объяснить, каким образом решатель задач породил решение, а интеллектуальный интерфейс позволил бы интерпретировать сформированное решение. При этом следует учесть наличие разного уровня компетенций как ЛПР, так и инженеров по знаниям.

Доверие к нейронным сетям: вызовы и стратегии обеспечения прозрачности

В последние годы в области проектирования, разработки и применения ИС наметилась тенденция к разработке «белых ящиков» — моделей, поведение которых можно объяснить и которые являются более понятными для людей. Для этого в ряде случаев предлагается разбивать ИС (решатель задач) на модули, каждый из которых может быть интерпретирован человеком, или изначально строить модели с локальной прозрачностью. Однако на данный момент «белый ящик» все еще остается достаточно нишевой идеей, а основным методом обучения всех моделей ИИ все еще является «черный ящик» (ЧЯ) (см. ИИ от Alpha Zero, ChatGPT, GPT-3, GPT-4 и др.). Быстрота и простота ЧЯ в настройке делают его более выгодным для компаний, которые ищут более быстрых результатов. Также следует отметить, что сейчас для решения большинства задач предпочтительнее использовать подходы ЧЯ, так как они показывают более высокую эффективность, чем интерпретируемые модели. Даже если интерпретируемые модели выдают хорошие результаты, все равно приходится идти на компромисс между их возможностями и пониманием. Эта проблема, по всей видимости, все еще будет актуальной достаточно продолжительное время.

Видится, что ситуацию может изменить только введение ограничений со стороны регуляторов, основанных на понимании того, что решения, принимаемые ИС, функционирующей на базе модели, относящейся к типу ЧЯ, могут содержать опасность. Предотвратить или хотя бы снизить возможную опасность может только четкое понимание того, как и почему ИС пришла к тому или иному выводу. К сожалению, подавляющее большинство популярных на данный момент инструментов и моделей не позволяют этого сделать: большинство систем ИИ работает как классические модели ЧЯ, т. е. разработчики ориентируются на вход и выход модели, не обращая внимание на процессы внутри ИС, которая занимается поиском результата. Они просто получают результаты и сравнивают их с желаемым результатом, не пытаясь понять, как ИС это достигает.

Для понимания того, как ИС (ИИ) приходит к тому или иному заключению, применяют методы и инструменты, разработанные для «объяснения» и «интерпретации» решений, однако при анализе классификаций данных методов можно заметить, что зачастую методы «объяснения» и методы «интерпретации» разделяют по-разному. Ввиду этого обозначим на начальном этапе различие между «объяснимостью» и «интерпретируемостью».

Объяснимость и интерпретируемость ИИ являются одними из важнейших качеств, которыми должен обладать современный ИИ. Объяснимость ИИ означает, что его принятие решений должно быть понятным и прозрачным для людей. Это качество особенно важно в областях, где принимаемые ИИ решения имеют серьезные последствия, например, в медицине, финансах, правоохранительных органах или в военной сфере.

Если решения ИИ непонятны или необоснованны, это может привести к серьезным проблемам и негативным последствиям.

Интерпретируемость ИИ подразумевает, что люди могут понимать, как ИИ принимает свои решения и на основе каких данных. Это может быть полезно для того, чтобы проверять работу ИИ и корректировать его ошибки и недостатки. Интерпретируемость ИИ также может помочь улучшить его работу, например, путем обучения ИИ новым способам обработки данных или принятия решений. Объяснимость же означает способность объяснить, как работает ИС, используя понятный человеку язык и концепции. Это означает, что ИС должна быть способна объяснить свое решение на основе входных данных и ее параметров.

Таким образом, интерпретация и объяснимость — два разных понятия, причем объяснимость является более широким концептом, который включает в себя не только интерпретацию, но и способность объяснить работу ИС в целом.

Интерпретируемое машинное обучение, применяемое в системах ИИ, не является подмножеством Explainable Artificial Intelligence (XAI) или наоборот [6]. В то время как первая направлена на создание прогностических моделей «стеклянного ящика» («белого»/«прозрачного»), вторая стремится понять ЧЯ, используя объяснительную модель, суррогатную модель, подход атрибуции, важность релевантности или другие статистические данные. Существует опасение, что определения, подходы и методы не совпадают, что приводит к непоследовательной классификации систем и моделей DL для интерпретации и объяснения.

Несмотря на недавние успехи в XAI, до сих пор неясно, как конкретная глубокая нейронная сеть (Deep Neural Network, DNN) приходит к тому или иному решению, насколько она уверена в своем решении, можем ли мы доверять или не доверять ей, и когда ее нужно корректировать. Основная причина этого заключается в том, что DNN приводят к сложным моделям ЧЯ, т. е. когда известны только входные признаки и выходные прогнозы, что затрудняет понимание природы обучения в рамках их структуры. Вопросы, над которыми нам следует подумать: «Можем ли мы доверять решениям, принятым моделями глубокого обучения?» или «Как модель глубокого обучения, принимает решение?».

Использование DL многообещающе, поскольку оно может обрабатывать сложные наборы данных и моделировать сильно нелинейные внутренние представления данных.

Растущий интерес исследователей к XAI привел к тому, что исследовательское сообщество DL сосредоточилось на методах интерпретируемости и объяснимости DNN. Основная проблема с объяснимостью, согласно DARPA [7], состоит в том, чтобы обеспечить достаточное обоснование выводов AI/ML, чтобы пользователи знали, почему вывод был сделан или нет, чтобы пользователь знал, когда алгоритм будет успешным или неудачным, и когда можно доверять. Чтобы укрепить доверие к системам и моделям DL [8, 9], заинтересованные стороны должны получить представление

о процессе принятия решений в системе или модели, точно узнав, почему и как система, модель и алгоритм DL пришли к определенному решению.

Можно предположить, что два ключевых слова «интерпретировать» и «объяснять» никогда не бывают полностью взаимозаменяемыми в языковом употреблении, особенно в техническом и научном языках [10]. Хотя и некоторые считают [11], что нет четких различий между двумя основными концепциями ХАИ — интерпретацией и объяснением.

Исследования в области моделирования DL традиционно были сосредоточены на улучшении качества, алгоритмов или скорости прогнозирования модели нейронной сети [12]. В то же время DNN обычно рассматриваются как модели ЧЯ из-за их многослойной нелинейности и глубоко вложенной структуры, которые часто критикуются как непрозрачные и непонятные для человека [13]. Поскольку они обучены, а не запрограммированы напрямую, может быть трудно понять, как именно они приходят к своим решениям.

Термин ЧЯ относится к модели, которая принимает последовательность входных данных запроса и выдает соответствующие выходные данные, скрывая при этом внутренние состояния, такие как архитектура модели [14]. Существуют методы объяснения ЧЯ, которые пытаются объяснить существующие модели DL без учета внутренней структуры модели. Этот класс методов объяснения не зависит от модели и может быть легко интегрирован в модели DL, от деревьев решений до сложных нейронных сетей.

Методы объяснения ЧЯ также называют постфактум (post hoc-методами), потому что их можно использовать для опроса моделей DL после обучения и развертывания, не зная процедур обучения. Объяснения, полученные таким образом, не гарантируют, что они будут удобными для человека, полезными или осуществимыми, и могут быть опасными в решениях с высокими ставками.

Противоположностью ЧЯ является «стеклянный ящик» [15, 16], который по своей сути прозрачен. Модули или задачи интерпретации заранее интегрируются в архитектуру и алгоритмы DL называются ante hoc. Преимущество этого подхода заключается в том, что специалисты-практики могут переводить модели в DL, обнаруживать ошибки в данных и/или маркировке и, в некоторых случаях, редактировать решения моделей, когда они не соответствуют значениям или знаниям предметной области. Этот тип подхода решает проблему компромисса между точностью и интерпретацией, которая является проблемой ЧЯ и post hoc-модели.

Растет интерес к пониманию того, как эти модели DL достигают своих успешных прогнозных задач. Работа по объяснению этих сетей ЧЯ была сосредоточена на понимании того, как фиксированная модель DL приводит к конкретному прогнозу [17]. Проблемы использования моделей DL для объяснения их решений в основном связаны со следующим: отсутствием прозрачности [16, 18]; отсутствием объяснимости [19–21]; большой сложностью и потребностью в больших вычислительных ресурсах [22]; недостаточной

устойчивостью к атакам со стороны противника [23] и неспособностью объяснить решения и действия так, чтобы ЛПР могли их понять [24].

Целью ХАИ является создание набора методов, которые создают более объяснимые модели, сохраняя при этом высокий уровень решения задач поиска, обучения, планирования и рассуждений за счет оптимизации и повышения точности. Основная идея ХАИ состоит в том, чтобы вскрыть алгоритм работы ЧЯ, объяснить, как ЧЯ принимал решения, и учесть выполняемые шаги и модели. Эти функции ЧЯ могут быть чрезвычайно сложными для понимания ЛПР.

Таким образом:

- цель ХАИ [25] состоит в том, чтобы сделать поведение системы более понятным для людей, предоставляя объяснения;
- задача ХАИ состоит в том, чтобы дать объяснения, которые были бы полными и интерпретируемыми [26].

Системы ХАИ должны быть в состоянии предоставлять исторически масштабируемые объяснения того, что система делала, что она делает сейчас и что произойдет дальше, а также раскрывать ключевую информацию, на которую она реагирует. При этом ХАИ предполагает, что существует много типов конечных пользователей или групп конечных пользователей.

ХАИ: теоретические основы объяснимости

Поскольку каждое решение DNN представляет собой комбинацию тысяч нейронов и весов, взаимодействующих друг с другом, объяснение зачастую оказывается очень сложной задачей. Потому существует большая потребность в возможности задать системе дополнительные вопросы, чтобы понять, почему система пришла к тому или иному прогнозу/решению.

Внутренней целью алгоритмов объяснения является облегчение человеческого понимания [27]. Если ЛПР поймет объяснения, оно будет более склонно доверять системам DL и применять их. Безопасность системы DL зависит от объяснений, честности, безопасности/конфиденциальности и отладки модели [28].

Однозначного и принятого всем научным сообществом определения «объяснимости» применительно к DL и ИИ в целом пока нет. «Объяснимость — это совокупность признаков интерпретируемой области, которые в данном примере способствовали принятию решения». Объяснения могут быть полными или частичными [25]. Полностью объяснимые модели обеспечивают полное объяснение и являются прозрачными [29]. Модели, которые можно объяснить лишь частично, раскрывают только важные части процесса рассуждений.

В научном исследовании научное объяснение должно включать как минимум две части [30]: объект, подлежащий объяснению, и содержание объяснения.

Выделяют четыре типа объяснений: аналитические (дидактические) высказывания, кейсы, визуализация, альтернативный выбор.

Некоторые исследователи считают [31], что визуальные объяснения должны удовлетворять двум критери-

ям: они должны различать классы и точно описывать конкретный экземпляр изображения.

Текстовые объяснения [32] представляют собой высказывания на естественном языке, которые словесно формулируются или описываются. Текстовые пояснения могут быть основаны либо на шаблонах [27], либо на правилах [33].

Объяснения на примерах [16, 34] выделяют частные экземпляры.

Объяснения упрощением [35] включают разделение сложного пространства признаков на более простые, объяснимые области.

Атрибуция релевантности признаков [36] присваивает оценку важности каждому признаку для конкретного ввода.

Можно различать системы объяснения с самоанализом, которые объясняют, как модель определяет свой окончательный вывод, и системы объяснения с обоснованием, которые производят предложения, подробно описывающие, как визуальные свидетельства согласуются с выводом системы.

Визуальные объяснения выделяют области DNN, которые описывают интересующие классы, или, в более общем смысле, визуализируют поведение модели [37].

В общем случае хорошее объяснение должно быть, по крайней мере, верным и интерпретируемым [38].

Достоверное объяснение — точная характеристика поведения модели, в то время как интерпретируемое объяснение легко понять специалисту-человеку [39].

Существует два типа объяснений: что научилась делать ИС и как она это делает?

Вопрос «что» относится к внешним свойствам преобразований, например, является ли они инвариантными по отношению к входным данным.

Вопрос «как» относится к внутреннему функционированию, т. е. к тому, как скрытые элементы обрабатывают информацию для достижения полученного результата.

Имеют место пять общих требований к полезному объяснению системы DL [40]: точность, понятность, достаточность, низкая ресурсоемкость и эффективность.

Не без оснований считается [11], что объяснимость модели ML обычно обратно пропорциональна ее производительности. Часто модели DL являются самыми мощными, но наименее объяснимыми. Деревья решений являются наиболее объяснимыми, но с наименьшей точностью [25]. Также считается, что для моделей ЧЯ, чем выше точность прогноза, тем ниже объяснительная способность модели [26, 41].

Объяснения можно классифицировать также, как и интерпретацию по нескольким основаниям. Так, например, объяснения могут быть глобальными и локальными.

Глобальные объяснения пытаются сосредоточиться на модели в целом. В этом случае вся модель может быть объяснена, и можно проследить рассуждения от входных данных до каждого возможного результата. Этот подход позволяет получить лучшее представление обо всей модели. Это может быть, например, визуализация

распределения веса в DNN или визуализация глубоких слоев сети, распространяющихся по сети.

Локальное объяснение пытается описать отдельные результаты, например, объяснить каждое предсказание. Цель состоит в том, чтобы объяснить, почему ЧЯ делает конкретное предсказание на основе локальных признаков, например пикселей. Эти объяснительные методы можно использовать для небольшой части сети, например, при рассмотрении одного фильтра в глубокой сети. Локальная объяснимость имеет дело с ситуацией, в результате которой можно понять только причины конкретного решения.

Также объяснимые методы подразделяют на post hoc- и ante hoc-методы.

Ante hoc (также известный как внутренний) объяснение (до этого события) решений модели ЧЯ заранее включается в архитектуру модели или в концептуальные ограничения. Ante hoc-системы предоставляют объяснения, которые идут от начала модели или ввода к выводу [40].

Post hoc (после события) объяснения решения модели ЧЯ могут быть даны постфактум. Post hoc-объяснения применяются для моделей, которые сложно объяснить. Потому post hoc-методы предполагают создание второй модели (эксплейнера), которая предоставляет объяснения. Это могут быть визуальные объяснения, текстовые пояснения, локальные пояснения, пояснения на примере, пояснения по упрощению и атрибуция релевантности признаков.

Основная проблема с post hoc-объяснениями заключается в том, что они могут не точно определять, как работает модель DL. Тем не менее, они предоставляют пользователям полезную информацию. Post hoc-модели позволяют объяснить процесс с точки зрения его результата, например, путем определения того, какая часть входных данных отвечает за конечный результат.

Основное различие между post hoc- и ante hoc-методами заключается в компромиссе между точностью модели и точностью объяснения.

Также методы объяснимости могут зависеть от модели DL, которую необходимо объяснить. Методы объяснимости, специфичные для модели, ограничены определенным классом моделей. Они пытаются понять модель DL, анализируя внутренние компоненты сети и то, как они взаимодействуют, исследуя функции активации или активацию обратного прохода на входе. Объяснимость, специфичная для модели, может быть связана с конкретным типом модели ЧЯ или входными данными.

Объяснители ЧЯ, не зависящие от модели, обычно требуют доступа только к функции предсказания модели, в то время как объяснители «стеклянного ящика» обычно требуют доступа к внутренним компонентам модели.

Методы, не зависящие от модели, можно разделить на три категории: упрощение модели, оценка атрибуции (релевантность признаков) и методы визуализации.

Независимые от модели объяснения (post hoc-объяснители) предназначены для применения к любой модели DL посредством получения некоторой информации из ее процедуры прогнозирования.

Чтобы повысить объяснимость моделей DL, необходимо принять во внимание, по крайней мере, несколько наиболее важных требований к объяснимости DL:

- причинность: способность метода прояснять взаимосвязь между входом и выходом в заданном контексте использования [16, 42, 43];
- корректируемость: способность метода вносить необходимые коррективы обратно в модель обучения [44];
- эффективность: способность метода поддерживать правильное принятие решений [45];
- производительность: способность метода поддерживать более быстрый наилучший вариант для принятия решений конечным пользователем [41];
- явность: способность метода немедленно и понятно давать объяснения [46, 47];
- добросовестность: способность метода давать объяснения, указывающие на истинные релевантные признаки [46, 47];
- достоверность: способность метода согласовываться с отображением ввода-вывода глубинной модели [48];
- информативность: способность метода предоставлять полезную информацию конечному пользователю посредством его вывода [16, 49];
- стабильность: согласованность метода предоставления аналогичных объяснений для аналогичных или соседних входных данных [46, 47];
- переносимость: способность метода обобщать и переносить новые знания на незнакомые ситуации [16, 48, 49];
- надежность: постоянство метода, позволяющего поддерживать небольшие возмущения входных данных, которые не меняют прогноз модели [46, 47, 50];
- убедительность: способность метода убеждать пользователей выполнять определенные действия [46, 47];
- исследуемость: способность метода проверять процесс обучения, который не сходится или не обеспечивает приемлемой производительности [46, 47].

Интерпретируемость моделей: методы и критерии оценки

По сравнению с системами ИИ, ориентированными на задачи, системы ХАИ предназначены для выполнения конкретных задач, которые приводят к объяснениям или созданию объяснительных моделей для решения проблем ЧЯ или систем Interface-mock-live, где интерпретации или понятные модели должны быть «стеклянными ящиками». Это может способствовать созданию доверительного ИИ и реализации удобных для человека методов ИИ, ориентированных на прозрачность, справедливость и подотчетность модели [51].

Некоторые авторы [17] определяют интерпретацию как «отображение абстрактного понятия, например, предсказанного класса, в область, которую человек может понять». Другие [25] определяют цель интерпретируемости как описание внутренностей системы способом, понятным человеку («интерпретируемость определяется как способность объяснить или пере-

дать значение в понятных человеку терминах»). Не существует математического определения интерпретируемости. Однако есть взгляды [52] на то, как можно измерить интерпретируемость. Нематематическое определение [43] гласит: «интерпретируемость — степень, в которой человек может понять причину решения».

Интерпретируемость относится к пассивному свойству модели и уровню, на котором конкретная модель имеет смысл для человека. Объяснимость же, напротив, можно рассматривать как активную сторону модели, которая относится к раскрытию внутренних функций модели. Другими словами, интерпретируемость — результат модели DL, а объяснимость — инструмент, который должен «открыть» этот результат. Чем выше интерпретируемость модели DL, тем легче кому-то понять, почему были сделаны те или иные решения или прогнозы. Модель более интерпретируема, чем другая модель, если ее решения легче понять человеку, чем решения последней. Отсюда можно сделать вывод, что понятие интерпретируемости имеет более широкую перспективу по сравнению с понятием объяснимости.

Выделяют [52] глобальную интерпретируемость и локальную интерпретируемость. Также важным аспектом является то, сколько времени пользователь должен или может потратить на понимание объяснения. Интерпретируемость модели, конечно, зависит и от опыта пользователя. Наряду с интерпретируемостью человеком, которая помогает людям понимать машины, существует также интерпретируемость машин. Это относится к тому, как машины «понимают» решения друг друга в рамках многоагентных ИС. Таким образом, глядя на цель DL с точки зрения моделирования, интерпретируемость тесно связана с ключевыми аудиториями конечных пользователей с точки зрения системы, а соответственно зависит от возможностей, заложенных в интеллектуальный интерфейс, через который ИС взаимодействует с ЛПР и инженерами по знаниям.

Глобальная интерпретируемость связана с пониманием того, как общая модель DL принимает решения, которые являются результатом прогноза. Это можно сделать, исследуя сложную структуру и параметры всей модели, какие входные шаблоны фиксируются и как они преобразуются для получения выходных данных [53]. Локальная интерпретируемость исследует локально причины поведения модели DL с учетом конкретного прогноза. Это достигается путем идентификации сопоставления каждой функции в конкретном входе для прогноза, сделанного DNN [39].

Считается [16], что интерпретируемость не является монолитным понятием, а отражает несколько различных идей и носит квазинаучный характер.

Рассмотрим основные концепции, описывающие интерпретируемость.

Понятность — свойство модели позволять человеку понимать ее работу без выяснения ее внутренней структуры или внутренних операций, с помощью которых модель обрабатывает данные [54]. Это свойство относится к ответу на вопрос: как работает модель DL? Естественно, что аудитория является краеуголь-

ным камнем ХАИ, когда дело доходит до понимания модели [51].

Когда дело доходит до понимания DNN, имеют дело с двумя взглядами на «понятность»:

- механистическое понимание (какой механизм сеть использует для решения проблемы или реализации функции);
- функциональное понимание (как сеть связывает входные переменные с выходными переменными).

Понятность является наиболее важным понятием в интерпретируемом DL.

Прозрачность обращается к свойству системы, чтобы объяснить, как она функционирует, даже если она ведет себя неожиданно. Очень важная проблема, которую приходится решать, заключается в том, что основная истина интерпретируемости зависит от полной прозрачности того, как модель DL приходит к своему решению [55].

Прозрачность можно дифференцировать [56] на прозрачность модели, прозрачность дизайна (построения), и алгоритмическую прозрачность.

Некоторые авторы [16] разделяют прозрачность модели по свойству понятности модели на три уровня: уровень всей модели, уровень отдельных компонентов и уровень алгоритмов.

В некоторых исследованиях интерпретируемость понимается как необходимое условие доверия [16]. Это указывает на то, что основная идея интерпретируемости состоит в том, чтобы помочь людям понять задачи прогнозирования моделей DL и доверять им.

Также интерпретируемость можно рассматривать как состоящую из трех категорий: интерпретируемость данных (какие измерения данных наиболее важны для задачи); интерпретируемость модели (как модель, принадлежащая к определенной категории, обычно выглядит в соответствии с моделью); интерпретируемость предсказания (объясняет, почему определенный шаблон x был классифицирован определенным образом $f(x)$).

Исходя из вышеизложенного, можно сосредоточиться на конкретных понятиях, описывающих интерпретируемость: понимание, прозрачность (доверие) и решение, а определение интерпретируемости можно дать следующим образом: «Интерпретируемость означает способность человека понимать и доверять результатам, полученным моделью глубокого обучения».

Инструментарий, применяемый в области объяснения решений, принимаемых ИИ

Приведем перечень программных реализаций частных методов, применяемых в области объяснения решений, принимаемых ИИ, а также ряд программных пакетов (таблица), которые могут помочь исследователям при решении задач объяснения и интерпретации наборов данных и моделей машинного обучения.

Local Interpretable Model Agnostic Explanations (LIME)

Исходный код: <https://github.com/marcotcr/lime>.

Лицензия: BSD 2-Clause «Simplified» License.

Объяснения, не зависящие от локальной интерпретируемой модели. Рассматривая модели машинного

обучения как функции ЧЯ, методы объяснения, не зависящие от модели, обычно имеют доступ только к выходным данным модели.

Anchors

Исходный код: <https://github.com/marcotcr/anchor>.

Лицензия: BSD 2-Clause «Simplified» License.

Основная идея заключается в том, что отдельные предсказания любой модели классификации ЧЯ объясняются путем нахождения решающего правила, которое в достаточной степени «закрепляет» предсказание — отсюда и название «якоря». Результирующие объяснения представляют собой правила принятия решений в форме операторов IF-THEN, которые определяют области в пространстве объектов

GraphLIME

Исходный код: <https://github.com/WilliamCCHuang/GraphLIME>.

Лицензия: MIT license.

Метод, который использует основную идею LIME, но не является линейным. Он применяется к особому типу архитектуры нейронных сетей, а именно к графовым нейронным сетям. Эти модели могут обрабатывать неевклидовы данные, поскольку они организованы в виде графовой структуры.

Layer-wise Relevance Propagation (LRP)

Исходный код: <https://github.com/chr5tphr/zennit>.

Лицензия: GPL-3.0, LGPL-3.0.

Исходный код: <https://github.com/albermax/investigate>.

Лицензия: The copyright in this software is being made available under the BSD License, included below. This software is subject to other contributor rights, including patent rights, and no such rights are granted under this license. All rights reserved.

Послойное распространение релевантности (LRP) — метод объяснения, основанный на распространении, т. е. он требует доступа к внутренним элементам модели (топологии, весам, активациям и т. д.).

Deep Taylor Decomposition (DTD)

Исходный код: <https://github.com/chr5tphr/zennit>.

Лицензия: GPL-3.0, LGPL-3.0.

Метод объяснения на основе распространения, который объясняет решения нейронной сети путем декомпозиции. Он перераспределяет значение функции (т. е. выходные данные нейронной сети) на входные переменные послойно, используя при этом математический инструмент (первого порядка) разложения Тейлора для определения пропорции или релевантности, присвоенной элементам нижнего уровня в процессе перераспределения (т. е. их соответствующие вклады). Этот подход тесно связан с методом LRP.

Prediction Difference Analysis (PDA)

Исходный код: <https://github.com/lmzintgraf/DeepVis-PredDiff>.

Лицензия: MIT license.

Метод основан на идее, когда для данного прогноза каждой входной функции присваивается значение релевантности по отношению к классу. Идея PDA заключается в том, что релевантность функции может быть оценена измерением того, как изменяется прогнозирование, когда признак неизвестен.

*Testing with Concept Activation Vectors (TCAV)*Исходный код: <https://github.com/tensorflow/tcav>.

Лицензия: Apache-2.0 license.

Целью подхода является количественная оценка того, насколько сильно концепция, такая как цвет, влияет на классификацию. Чтобы рассчитать такой TCAV, сначала необходимо собрать и объединить два набора данных: набор данных, содержащий изображения, представляющие концепцию, и набор данных, состоящий из изображений, в которых эта концепция отсутствует.

*Explainable Graph Neural Networks (XGNN)*Исходный код: <https://github.com/divelab/DIG/tree/dig-stable/benchmarks/xgraph>.

Лицензия: GPL-3.0 license.

Является методом post hoc, который работает на уровне модели, что означает, что он не стремится предоставлять объяснения на уровне отдельных примеров (метод объяснения изобретен специально для задачи классификации графов).

*Shapley Values (SHAP)*Исходный код: <https://github.com/slundberg/shap>.

Лицензия: MIT license.

Методы в этом семействе связаны с объяснениями модели функции в какой-то отдельной точке. Значения SHAP симметричны. Это означает, что, если две переменные оказывают одинаковое влияние на поведение модели, например, потому, что они принимают одинаковые значения, они получают одинаковые атрибуты.

*Asymmetric Shapley Values (ASV)*Исходный код: <https://github.com/nredell/shapFlex>.

Лицензия: MIT license.

ASV позволяют использовать дополнительные знания о причинно-следственных связях между переменными в процессе объяснения модели. Причинно-следственная связь, описанная в форме причинного графика, позволяет перераспределять атрибуцию переменных таким образом, чтобы исходные переменные имели большую атрибуцию, оказывая влияние как на другие зависимые переменные, так и на предсказания модели.

*Break-Down*Исходный код: <https://github.com/ModelOriented/DALEX>.

Лицензия: GPL-3.0 license.

Если в модели есть взаимодействия, разный порядок переменных может привести к разным вкладам. Значения SHAP усредняются по всем возможным порядкам, что приводит к аддитивным вкладам и игнорированию взаимодействий. Альтернативой является анализ различных порядков, чтобы обнаружить, когда одна переменная имеет разные вклады в зависимости от того, какие другие переменные предшествуют ей. Метод Break-Down анализирует различные порядки для идентификации и визуализации взаимодействий в модели.

*Shapley Flow*Исходный код: <https://github.com/nathanwang000/Shapley-Flow>.

Лицензия: не указано.

Shapley Flow также позволяет использовать структуру зависимостей между переменными в процессе

объяснения. Как и в ASV, взаимосвязь описывается причинно-следственным графом. Однако, в отличие от ASV и других методов, атрибуция присваивается не узлам (переменным), а ребрам (отношениям между переменными). Ребро на графике имеет значение, если его удаление изменит прогнозы модели.

*Textual Explanations of Visual Models*Исходный код: <https://github.com/LisaAnne/ECCV2016>.

Лицензия: All Rights Reserved. Permission to use, copy, modify, and distribute this software and its documentation for educational, research, and not-for-profit purposes, without fee and without a signed licensing agreement, is hereby granted, provided that the above copyright notice, this paragraph and the following two paragraphs appear in all copies, modifications, and distributions.

Генерация текстовых описаний изображений решается несколькими моделями машинного обучения, которые содержат как часть, обрабатывающую входные изображения — обычно сверточную нейронную сеть, так и ту, которая изучает адекватную текстовую последовательность, обычно рекуррентную нейронную сеть. Эти две части взаимодействуют для создания описательных предложений с изображениями, что предполагает успешное выполнение задачи классификации. Важно отметить, что простое описание содержимого изображения не эквивалентно объяснению процесса принятия решений в модели нейронной сети.

*Integrated Gradients*Исходный код: <https://github.com/ankurtaly/Integrated-Gradients>.

Лицензия: не указано.

Метод Integrated Gradients основан на двух фундаментальных аксиомах: чувствительности и инвариантности реализации. Чувствительность означает, что ненулевые атрибуты присваиваются каждому входному сигналу и базовой линии, которые отличаются по одному признаку, но имеют разные прогнозы. Инвариантность реализации означает, что если две модели ведут себя идентично/функционально эквивалентны, то атрибуции должны быть идентичными. Хотя эти две аксиомы звучат очень естественно, оказывается, что многие методы атрибуции не обладают этими свойствами. В частности, когда модель имеет сглаженные прогнозы для конкретной интересующей точки, градиент в интересующей точке обнуляется и не несет информации, полезной для объяснения.

Causal Models

Исходный код: нет.

Лицензия: нет.

Модель Causal Models можно считать расширением байесовских моделей среды RL с использованием контрфактуалов. Она учитывает события, которые могут произойти, или состояния среды, которые могут быть достигнуты при различных действиях, предпринятых агентом RL. В конечном счете цель любого агента RL — максимизировать долгосрочное вознаграждение; объяснение обеспечивает причинно-следственные связи до тех пор, пока не будет достигнуто состояние получения вознаграждения.

Таблица. Программные пакеты, которые могут применяться при решении задач объяснения/интерпретации наборов данных и моделей машинного обучения

Table. Software packages that can be used to solve problems of explaining/interpreting data sets and machine learning models

№	Инструмент	Исходный код	Лицензия
1	AI Fairness 360 (AIF360)	https://github.com/Trusted-AI/AIF360	Apache-2.0 license
2	AI Explainability 360 (AIX360)	https://github.com/Trusted-AI/AIX360	Apache-2.0 license
3	Alibi Explain	https://github.com/SeldonIO/alibi	Apache-2.0 license
4	Analysis by synthesis (ABS)	https://github.com/bethgelab/AnalysisBySynthesis	Apache-2.0 license
5	Captum	https://github.com/pytorch/captum	BSD-3-Clause license
6	DALEX	https://github.com/ModelOriented/DALEX	GPL-3.0 license
7	DeepExplain	https://github.com/marcoancona/DeepExplain	MIT license
8	Deep visualization tool	https://github.com/yosinski/deep-visualization-toolbox	MIT license
9	ELI5	https://github.com/TeamHG-Memex/eli5	MIT license
10	explainX	https://github.com/explainX/explainx	MIT license
11	FAT Forensics	https://github.com/fat-forensics/fat-forensics	BSD-3-Clause license
12	InterpretML	https://github.com/interpretml/interpret	MIT license
13	iNNvestigate	https://github.com/albermax/innvestigate	в соответствии с BSD с указанными дополнениями
14	H2O.ai	https://github.com/h2oai/mli-resources	Не указано
15	L2X	https://github.com/Jianbo-Lab/L2X	Не указано
16	Rectified gradient	https://github.com/1202kbs/Rectified-Gradient	Не указано
17	Saliency relevance propagation	https://github.com/Hey1Li/Salient-Relevance-Propagation	MIT license
18	Sensitivity analysis library (SALib)	https://github.com/SALib/SALib	MIT license
19	Skater	https://github.com/oracle/Skater	UPL-1.0 license
20	tfexplain	https://github.com/sicara/tf-explain	MIT license
21	treeinterpreter	https://pypi.org/project/treeinterpreter/	BSD License (BSD)
22	XAI	https://github.com/EthicalML/xai	MIT license

Meaningful Perturbations

Исходный код: https://github.com/ruthcfong/perturb_explanations.

Лицензия: не указано.

Метод может рассматриваться как независимый от модели метод объяснения, основанный на возмущениях. Таким образом, объяснение вычисляется исключительно на основе реакции модели на возмущенную (или закрытую) входную выборку.

EXplainable Neural-Symbolic Learning (X-NeSyL)

Исходный код: <https://github.com/JulesSanchez/X-NeSyL>.

Лицензия: не указано.

Исходный код: <https://github.com/JulesSanchez/MonumaIAutomaticStyleClassification>.

Лицензия: GPL-3.0 license.

Нейро-символические методы включают в себя предварительные человеческие знания для различных задач, таких как изучение концепций, и в то же время они дают более понятный результат, такой как математические уравнения или языки, специфичные для предметной области.

Заключение

При проектировании и создании интеллектуальных систем важно тщательно продумать, как представлять факты и знания в системе, какие механизмы обработки реализовывать в решателе, каким должен быть интеллектуальный интерфейс взаимодействия с лицами, принимающими решение, какова их квалификация, какой язык взаимодействия инженеров по знаниям с интеллектуальной системой выбрать и т. п. Ответ на эти вопросы во многом обусловит потенциальную возможность реализации подсистемы, способной объяснить конечному пользователю ход принятия решения и донести эту информацию на доступном ему языке.

Следует определять допустимую точность формируемых решений и время, отводимое на их порождение, так как в ряде случаев, учитывая данные факторы, можно отдать предпочтение менее точным и оперативным системам, но более «прозрачным», т. е. основанным на применении моделей из класса «белых ящиков», что позволит повысить степень объяснимости их работы. В случае невозможности отказаться от применения моделей типа «черного ящика», нужно понимать, какие методы и инструменты их «объяснимости» и «интер-

претируемости» существуют и какие применимы для той или иной модели.

Отсутствие возможности понять, как та или иная интеллектуальная система приходит к тому или иному решению, таит опасности, связанные как с безопасностью самой модели, так и с безопасностью реализации решений, ею порождаемых. Ввиду этого объяснимость и интерпретируемость искусственного интеллекта являются одними из важнейших качеств, которыми должен обладать современный искусственный интеллект. Чтобы объяснить, как искусственный интеллект принимает определенные решения, используются методы и инструменты, предназначенные для их интерпретации и объяснения.

Несмотря на достижения в области разработки объяснимого искусственного интеллекта, по-прежнему не понятно, как глубокие нейронные сети принимают решения, насколько они уверены в своих выводах, и когда следует корректировать их решения, чтобы было можно доверять им. Если же пользователь поймет объяснения, он будет более склонен доверять системам глубокого обучения и применять их.

Интерпретация же относится к процессу понимания того, как интеллектуальная система (и нейронная сеть — в частности) обрабатывает информацию и при-

нимает решения на основе предоставленных ей данных. Целью интерпретации является выявление взаимосвязей между входными данными и выходом, а также выявление слабых мест и возможных ошибок.

С другой стороны, объяснимость означает способность объяснить, как работает нейронная сеть и интеллектуальная система в целом, используя понятный человеку язык и концепции. Это означает, что интеллектуальная система должна быть способна объяснить свое решение на основе входных данных и параметров, используемых для обучения.

Однозначного и принятого всем научным сообществом определения «объяснимости» применительно к глубокому обучению и искусственному интеллекту в целом пока нет, но не вызывает сомнения, что объяснение должно быть точным, понятным, достаточным и не требовать большого количества ресурсов для его осуществления.

Таким образом, интерпретация и объяснимость — два разных понятия, причем объяснимость является более широким концептом, который включает в себя не только интерпретацию, но и способность объяснить работу нейронной сети и интеллектуальной системы в целом.

Литература

1. Финн В.К. Об интеллектуальном анализе данных // *Новости искусственного интеллекта*. 2004. № 3. С. 3–18.
2. Финн В.К. Искусственный интеллект: Идеальная база и основной продукт // IX Национальная конференция «Искусственный интеллект-2004». 2004. Т. 1. С. 11–20.
3. Бирюков Д.Н., Ломоко А.Г., Ростовцев Ю.Г. Облик антиципирующих систем предотвращения рисков реализации киберугроз // *Труды СПИИРАН*. 2015. № 2(39). С. 5–25.
4. Бирюков Д.Н., Ломоко А.Г. Денотационная семантика контекстов знаний при онтологическом моделировании предметных областей конфликта // *Труды СПИИРАН*. 2015. № 5(42). С. 155–179.
5. Бирюков Д.Н., Ломоко А.Г., Жолус Р.Б. Пополнение онтологических систем знаний на основе моделирования умозаключений с учетом семантики ролей // *Труды СПИИРАН*. 2016. № 4(47). С. 105–129. <https://doi.org/10.15622/sp.47.6>
6. Namatēvs I., Sudars K., Dobrājs A. Interpretability versus explainability: classification for understanding deep learning systems and models // *Computer Assisted Methods in Engineering and Science*. 2022. V. 29. N 4. P. 297–356. <http://dx.doi.org/10.24423/cames.518>
7. Gunning D. Explainable artificial intelligence (XAI). 2017. [Электронный ресурс]. URL: <https://nsarchive.gwu.edu/sites/default/files/documents/5794867/National-Security-Archive-David-Gunning-DARPA.pdf> (дата обращения: 21.10.2024).
8. Varshney K.R. Trustworthy machine learning and artificial intelligence // *XRDS: Crossroads, The ACM Magazine for Students*. 2019. V. 25. N 3. P. 26–29. <https://doi.org/10.1145/3313109>
9. Doshi-Velez F., Kim B., Towards a rigorous science of interpretable machine learning // *arXiv*. 2017. arXiv:1702.08608v2. <https://doi.org/10.48550/arXiv.1702.08608>
10. Yuan W., Liu P., Neubig G. Can we automate scientific reviewing? // *arXiv*. 2021. arXiv:2102.00176. <https://doi.org/10.48550/arXiv.2102.00176>
11. Arya V., Bellamy R.K.E., Chen P.-Yu., Dhurandhar A., Hind M., Hoffman S.C., Houde S., Liao V.Q., Luss R., Mojsilović A., et al. One explanation does not fit all: A toolkit and taxonomy of AI explainability techniques // *arXiv*. 2019. arXiv:1909.03012. <https://doi.org/10.48550/arXiv.1909.03012>
12. Samek W., Wiegand T., Müller K.-R. Explainable artificial intelligence: Understanding, visualizing and interpreting deep

References

1. Finn V.K. On intelligent data analysis. *Novosti Iskusstvennogo Intelleta*, 2004, no. 3, pp. 3–18. (in Russian)
2. Finn V.K. Artificial Intelligence: The Idea Base and the Main Product. *Proc. of the 9th National Conference on Artificial Intelligence*, 2004, vol. 1, pp. 11–20. (in Russian)
3. Biryukov D.N., Lomako A.G., Rostovtsev Yu.G. The appearance of anticipating cyber threats risk prevention systems. *SPIIRAS Proceedings*, 2015, no. 2(39), pp. 5–25. (in Russian)
4. Biryukov D.N., Lomako A.G. Denotational semantics of knowledge contexts in ontological modeling of subject domains of the conflict. *SPIIRAS Proceedings*, 2015, no. 5(42), pp. 155–179. (in Russian)
5. Biryukov D.N., Lomako A.G., Zhulus R.B. Ontological knowledge system completion based on modeling inferences taking into account role semantics. *SPIIRAS Proceedings*, 2016, no. 4(47), pp. 105–129. (in Russian). <https://doi.org/10.15622/sp.47.6>
6. Namatēvs I., Sudars K., Dobrājs A. Interpretability versus explainability: classification for understanding deep learning systems and models. *Computer Assisted Methods in Engineering and Science*, 2022, vol. 29, no. 4, pp. 297–356. <http://dx.doi.org/10.24423/cames.518>
7. Gunning D. *Explainable artificial intelligence (XAI)*, 2017. Available at: <https://nsarchive.gwu.edu/sites/default/files/documents/5794867/National-Security-Archive-David-Gunning-DARPA.pdf> (accessed: 21.10.2024).
8. Varshney K.R. Trustworthy machine learning and artificial intelligence. *XRDS: Crossroads, The ACM Magazine for Students*, 2019, vol. 25, no. 3, pp. 26–29. <https://doi.org/10.1145/3313109>
9. Doshi-Velez F., Kim B., Towards a rigorous science of interpretable machine learning. *arXiv*, 2017, arXiv:1702.08608v2. <https://doi.org/10.48550/arXiv.1702.08608>
10. Yuan W., Liu P., Neubig G. Can we automate scientific reviewing? *arXiv*, 2021, arXiv:2102.00176. <https://doi.org/10.48550/arXiv.2102.00176>
11. Arya V., Bellamy R.K.E., Chen P.-Yu., Dhurandhar A., Hind M., Hoffman S.C., Houde S., Liao V.Q., Luss R., Mojsilović A., et al. One explanation does not fit all: A toolkit and taxonomy of AI explainability techniques. *arXiv*, 2019, arXiv:1909.03012. <https://doi.org/10.48550/arXiv.1909.03012>
12. Samek W., Wiegand T., Müller K.-R. Explainable artificial intelligence: Understanding, visualizing and interpreting deep

- learning models // arXiv. 2017. arXiv.1708.08296. <https://doi.org/10.48550/arXiv.1708.08296>
13. Angelov P., Soares E. Towards explainable deep neural networks (xDNN) // *Neural Networks*. 2020. V. 130. P. 185–194. <https://doi.org/10.1016/j.neunet.2020.07.010>
 14. Oh S.J., Augustin M., Schiele B., Fritz M. Towards reverse-engineering black-box neural networks // arXiv. 2018. arXiv.1711.01768. <https://doi.org/10.48550/arXiv.1711.01768>
 15. Rai A. Explainable AI: From black box to glass box // *Journal of the Academy of Marketing Science*. 2020. V. 48. N 1. P. 137–141. <https://doi.org/10.1007/s11747-019-00710-5>
 16. Lipton Z.C. The mythos of model interpretability // arXiv. 2017. arXiv.1606.03490. <https://doi.org/10.48550/arXiv.1606.03490>
 17. Montavon G., Samek W., Müller K.-R. Methods for interpreting and understanding deep neural networks. *Digital Signal Processing*, 2018, V. 73, P. 1–15. <https://doi.org/10.1016/j.dsp.2017.10.011>
 18. Mascharka D., Tran P., Soklaski R., Majumdar A. Transparency by design: Closing the gap between performance and interpretability in visual reasoning // arXiv. 2018. arXiv:1803.05268. <https://doi.org/10.48550/arXiv.1803.05268>
 19. Beaudouin V., Bloch I., Bounie D., Cléménçon S., d'Alché-Buc F., Eagan J., Maxwell W., Mozharovskiy P., Parekh J. Flexible and context-specific AI explainability: A multidisciplinary approach // arXiv. 2020. arXiv:2003.07703v1. <https://doi.org/10.48550/arXiv.2003.07703>
 20. Sokol K., Flach P. Explainability fact sheets: A framework for systematic assessment of explainable approaches // *Proc. of the 2020 Conference on Fairness, Accountability, and Transparency (FAT*20)*. 2020. P. 56–67. <https://doi.org/10.1145/3351095.3372870>
 21. Xu F., Uszkoreit H., Du Y., Fan W., Zhao D., Zhu J. Explainable AI: A brief survey on history, research areas, approaches and challenges // *Lecture Notes in Computer Science*. 2019. V. 11839. P. 563–574. https://doi.org/10.1007/978-3-030-32236-6_51
 22. Thompson N.C., Greenwald K., Lee K., Manso G.F. The computational limits of deep learning // arXiv. 2020. arXiv:2007.05558. <https://doi.org/10.48550/arXiv.2007.05558>
 23. DuSell B., Chiang D. Learning context-free languages with nondeterministic stack RNNs // *Proc. of the 24th Conference on Computational Natural Language Learning*. 2020. P. 507–519. <https://doi.org/10.18653/v1/2020.conll-1.41>
 24. Flambeau J.K.F., Norbert T. Simplifying the explanation of deep neural networks with sufficient and necessary feature-sets: case of text classification // arXiv. 2020. arXiv:2010.03724v2. <https://doi.org/10.48550/arXiv.2010.03724>
 25. Gunning D., Stefik M., Choi J., Miller T., Stumpf S., Yang G.-Z. XAI — Explainable artificial intelligence // *Science Robotics*. 2019. V. 4. N 37. P. eaay7120. <https://doi.org/10.1126/scirobotics.aay7120>
 26. Gilpin L.H., Bau D., Yuan B.Z., Bajwa A., Specter M., Kagal L. Explaining explanations: an overview of interpretability of machine learning // arXiv. 2018. arXiv:1806.00069v3. <https://doi.org/10.48550/arXiv.1806.00069>
 27. Alber M. Software and application patterns for explanation methods // arXiv. 2019. arXiv:1904.04734v1. <https://doi.org/10.48550/arXiv.1904.04734>
 28. Zhao X., Banks A., Sharp J., Robu V., Flynn D., Fisher M., Huang X. A safety framework for critical systems utilising deep neural networks // arXiv. 2020. arXiv:2003.05311v3. https://doi.org/10.1007/978-3-030-54549-9_16
 29. Weller A. Transparency: Motivations and challenges // *Lecture Notes in Computer Science*. 2019. V. 11700. P. 23–40. https://doi.org/10.1007/978-3-030-28954-6_2
 30. Raghu M., Schmidt E. A survey of deep learning for scientific discovery // arXiv. 2020. arXiv:2003.11755v1. <https://doi.org/10.48550/arXiv.2003.11755>
 31. Hendricks L.A., Rohrbach A., Schiele B., Darrell T., Akata Z. Generating visual explanations with natural language // *Applied AI Letters*. 2021. V. 2. N 4. P. e55. <https://doi.org/10.1002/ail2.55>
 32. Kaplan J., McCandlish S., Henighan T., Brown T.B., Chess B., Child R., Gray S., Radford A., Wu J., Amodei D. Scaling laws for neural language models // arXiv. 2020. arXiv:2001.08361v1. <https://doi.org/10.48550/arXiv.2001.08361>
 33. Towell G.G., Shavlik J.W. Extracting refined rules from knowledge-based neural networks // *Machine Learning*. 1993. V. 13. N 1. P. 71–101. <https://doi.org/10.1007/bf00993103>
 34. Molnar C. Interpretable Machine Learning: A Guide for Making Black Box Models Explainable. Christoph Molnar, 2025. 392 p.
 - learning models. *arXiv*, 2017, arXiv.1708.08296. <https://doi.org/10.48550/arXiv.1708.08296>
 13. Angelov P., Soares E. Towards explainable deep neural networks (xDNN). *Neural Networks*, 2020, vol. 130, pp. 185–194. <https://doi.org/10.1016/j.neunet.2020.07.010>
 14. Oh S.J., Augustin M., Schiele B., Fritz M. Towards reverse-engineering black-box neural networks. *arXiv*, 2018, arXiv.1711.01768. <https://doi.org/10.48550/arXiv.1711.01768>
 15. Rai A. Explainable AI: From black box to glass box. *Journal of the Academy of Marketing Science*, 2020, vol. 48, no. 1, pp. 137–141. <https://doi.org/10.1007/s11747-019-00710-5>
 16. Lipton Z.C. The mythos of model interpretability. *arXiv*, 2017, arXiv.1606.03490. <https://doi.org/10.48550/arXiv.1606.03490>
 17. Montavon G., Samek W., Müller K.-R. Methods for interpreting and understanding deep neural networks. *Digital Signal Processing*, 2018, vol. 73, pp. 1–15. <https://doi.org/10.1016/j.dsp.2017.10.011>
 18. Mascharka D., Tran P., Soklaski R., Majumdar A. Transparency by design: Closing the gap between performance and interpretability in visual reasoning. *arXiv*, 2018, arXiv:1803.05268. <https://doi.org/10.48550/arXiv.1803.05268>
 19. Beaudouin V., Bloch I., Bounie D., Cléménçon S., d'Alché-Buc F., Eagan J., Maxwell W., Mozharovskiy P., Parekh J. Flexible and context-specific AI explainability: A multidisciplinary approach. *arXiv*, 2020, arXiv:2003.07703v1. <https://doi.org/10.48550/arXiv.2003.07703>
 20. Sokol K., Flach P. Explainability fact sheets: A framework for systematic assessment of explainable approaches. *Proc. of the 2020 Conference on Fairness, Accountability, and Transparency (FAT*20)*, 2020, pp. 56–67. <https://doi.org/10.1145/3351095.3372870>
 21. Xu F., Uszkoreit H., Du Y., Fan W., Zhao D., Zhu J. Explainable AI: A brief survey on history, research areas, approaches and challenges. *Lecture Notes in Computer Science*, 2019, vol. 11839, pp. 563–574. https://doi.org/10.1007/978-3-030-32236-6_51
 22. Thompson N.C., Greenwald K., Lee K., Manso G.F. The computational limits of deep learning. *arXiv*, 2020, arXiv:2007.05558. <https://doi.org/10.48550/arXiv.2007.05558>
 23. DuSell B., Chiang D. Learning context-free languages with nondeterministic stack RNNs. *Proc. of the 24th Conference on Computational Natural Language Learning*, 2020, pp. 507–519. <https://doi.org/10.18653/v1/2020.conll-1.41>
 24. Flambeau J.K.F., Norbert T. Simplifying the explanation of deep neural networks with sufficient and necessary feature-sets: case of text classification. *arXiv*, 2020, arXiv:2010.03724v2. <https://doi.org/10.48550/arXiv.2010.03724>
 25. Gunning D., Stefik M., Choi J., Miller T., Stumpf S., Yang G.-Z. XAI — Explainable artificial intelligence. *Science Robotics*, 2019, vol. 4, no. 37, pp. eaay7120. <https://doi.org/10.1126/scirobotics.aay7120>
 26. Gilpin L.H., Bau D., Yuan B.Z., Bajwa A., Specter M., Kagal L. Explaining explanations: an overview of interpretability of machine learning. *arXiv*, 2018, arXiv:1806.00069v3. <https://doi.org/10.48550/arXiv.1806.00069>
 27. Alber M. Software and application patterns for explanation methods. *arXiv*, 2019, arXiv:1904.04734v1. <https://doi.org/10.48550/arXiv.1904.04734>
 28. Zhao X., Banks A., Sharp J., Robu V., Flynn D., Fisher M., Huang X. A safety framework for critical systems utilising deep neural networks. *arXiv*, 2020, arXiv:2003.05311v3. https://doi.org/10.1007/978-3-030-54549-9_16
 29. Weller A. Transparency: Motivations and challenges. *Lecture Notes in Computer Science*, 2019, vol. 11700, pp. 23–40. https://doi.org/10.1007/978-3-030-28954-6_2
 30. Raghu M., Schmidt E. A survey of deep learning for scientific discovery. *arXiv*, 2020, arXiv:2003.11755v1. <https://doi.org/10.48550/arXiv.2003.11755>
 31. Hendricks L.A., Rohrbach A., Schiele B., Darrell T., Akata Z. Generating visual explanations with natural language. *Applied AI Letters*, 2021, vol. 2, no. 4, pp. e55. <https://doi.org/10.1002/ail2.55>
 32. Kaplan J., McCandlish S., Henighan T., Brown T.B., Chess B., Child R., Gray S., Radford A., Wu J., Amodei D. Scaling laws for neural language models. *arXiv*, 2020, arXiv:2001.08361v1. <https://doi.org/10.48550/arXiv.2001.08361>
 33. Towell G.G., Shavlik J.W. Extracting refined rules from knowledge-based neural networks. *Machine Learning*, 1993, vol. 13, no. 1, pp. 71–101. <https://doi.org/10.1007/bf00993103>

35. Kim S., Jeong M., Ko B.C. Interpretation and simplification of deep forest // *arXiv*. 2020. arXiv:2001.04721v4. <https://doi.org/10.48550/arXiv.2001.04721>
36. Nam W.-J., Gur S., Choi J., Wolf L., Lee S.-W. Relative attributing propagation: interpreting the comparative contributions of individual units in deep neural networks // *arXiv*. 2019. arXiv:1904.00605v4. <https://doi.org/10.48550/arXiv.1904.00605>
37. Oramas J.M., Wang K., Tuytelaars T. Visual explanation by interpretation: Improving visual feedback capabilities of deep neural networks // *arXiv*. 2019. arXiv:1712.06302v3. <https://doi.org/10.48550/arXiv.1712.06302>
38. Rudin C. Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead // *arXiv*. 2019. arXiv:1811.10154v3. <https://doi.org/10.48550/arXiv.1811.10154>
39. Samek W., Montavon G., Vedaldi A., Hansen L.K., Müller K.-R. Explainable AI: interpreting, explaining and visualizing deep learning // *Lecture Notes in Computer Science*. 2019. V. 11700. 439 p. <https://doi.org/10.1007/978-3-030-28954-6>
40. Hansen L.K., Rieger L. Interpretability in intelligent systems – A new concept? // *Lecture Notes in Computer Science*. 2019. V. 11700. P. 41–49. https://doi.org/10.1007/978-3-030-28954-6_3
41. Liao Q.V., Gruen D., Miller S. Questioning the AI: Informing design practices for explainable AI user experiences // *arXiv*. 2020. arXiv:2001.02478v2. <https://doi.org/10.48550/arXiv.2001.02478>
42. Holzinger A., Längs G., Denk H., Zatlouk K., Müller H. Causability and explainability of artificial intelligence in medicine // *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*. 2019. V. 9. N 4. P. e1312. <https://doi.org/10.1002/widm.1312>
43. Miller T. Explanation in artificial intelligence: insights from the social sciences // *Artificial Intelligence*. 2019. V. 267. P. 1–38. <https://doi.org/10.1016/j.artint.2018.07.007>
44. Kulesza T., Burnett M., Wong W., Stumpf S. Principles of explanatory debugging to personalize interactive machine learning // *Proc. of the 20th International Conference on Intelligent User Interfaces (IUI '15)*. 2015. P. 126–137. <https://doi.org/10.1145/2678025.2701399>
45. Tintarev N. Explaining recommendations // *Lecture Notes in Computer Science*. 2007. V. 4511. P. 470–474. https://doi.org/10.1007/978-3-540-73078-1_67
46. Chrysostomou G., Alertis N. Improving the faithfulness of attention-based explanations with task-specific information for text classification // *arXiv*. 2021. arXiv:2105.02657v2. <https://doi.org/10.48550/arXiv.2105.02657>
47. Vilone G., Longo L. Explainable artificial intelligence: A systematic review // *arXiv*. 2020. arXiv:2006.00093v3. <https://doi.org/10.48550/arXiv.2006.00093>
48. Papenmeier A., Engleblenne G., Seifert C. How model accuracy and explanation fidelity influence user trust // *arXiv*. 2019. arXiv:1907.12652v1. <https://doi.org/10.48550/arXiv.1907.12652>
49. Harutyunyan H., Achille A., Paolini G., Majumder O., Ravichandran A., Bhotika R., Soatto S. Estimating informativeness of samples with smooth unique information // *arXiv*. 2021. arXiv:2101.06640v1. <https://doi.org/10.48550/arXiv.2101.06640>
50. Liu S., Wang X., Liu M., Zhu J. Towards better analysis of machine learning models: a visual analytics perspective // *Visual Informatics*. 2017. V. 1. N 1. P. 48–56. <https://doi.org/10.1016/j.visinf.2017.01.006>
51. Arrieta A.B., Díaz-Rodríguez N., Del Ser J., Bennetot A., Tabik S., Barbado A., García S., Gil-López S., Molina D., Benjamins R., Chatila R., Herrera F. Explainable artificial intelligence (XAI): concepts, taxonomies, opportunities and challenges toward responsible AI // *Information Fusion*. 2020. V. 58. P. 82–115. <https://doi.org/10.1016/j.inffus.2019.12.012>
52. Girshick R., Donahue J., Darrell T., Malik J. Rich feature hierarchies for accurate object detection and semantic segmentation // *Proc. of the IEEE Conference on Computer Vision and Pattern Recognition*. 2014. P. 580–587. <https://doi.org/10.1109/CVPR.2014.81>
53. Ancona M., Ceolini E., Öztireli C., Gross M. Towards better understanding of gradient-based attribution methods for deep neural networks // *arXiv*. 2018. arXiv:1711.06104v4. <https://doi.org/10.48550/arXiv.1711.06104>
54. Rumelhart D.E., Hinton G.E., Williams R.J. Learning internal representations by error propagation // *Readings in Cognitive Science: A Perspective from Psychology and Artificial Intelligence*. 2013. P. 399–421.
55. Kindermans P.-J., Hooker S., Adebayo J., Alber M., Schütt K.T., Dähne S., Erhan D., Kim B. The (Un) reliability of saliency
34. Molnar C. *Interpretable Machine Learning: A Guide for Making Black Box Models Explainable*. Christoph Molnar, 2025, 392 p.
35. Kim S., Jeong M., Ko B.C. Interpretation and simplification of deep forest. *arXiv*, 2020, arXiv:2001.04721v4. <https://doi.org/10.48550/arXiv.2001.04721>
36. Nam W.-J., Gur S., Choi J., Wolf L., Lee S.-W. Relative attributing propagation: interpreting the comparative contributions of individual units in deep neural networks. *arXiv*, 2019, arXiv:1904.00605v4. <https://doi.org/10.48550/arXiv.1904.00605>
37. Oramas J.M., Wang K., Tuytelaars T. Visual explanation by interpretation: Improving visual feedback capabilities of deep neural networks. *arXiv*, 2019, arXiv:1712.06302v3. <https://doi.org/10.48550/arXiv.1712.06302>
38. Rudin C. Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. *arXiv*, 2019, arXiv:1811.10154v3. <https://doi.org/10.48550/arXiv.1811.10154>
39. Samek W., Montavon G., Vedaldi A., Hansen L.K., Müller K.-R. Explainable AI: interpreting, explaining and visualizing deep learning. *Lecture Notes in Computer Science*, 2019, vol. 11700. 439 p. <https://doi.org/10.1007/978-3-030-28954-6>
40. Hansen L.K., Rieger L. Interpretability in intelligent systems – A new concept? *Lecture Notes in Computer Science*, 2019, vol. 11700, pp. 41–49. https://doi.org/10.1007/978-3-030-28954-6_3
41. Liao Q.V., Gruen D., Miller S. Questioning the AI: Informing design practices for explainable AI user experiences. *arXiv*, 2020, arXiv:2001.02478v2. <https://doi.org/10.48550/arXiv.2001.02478>
42. Holzinger A., Längs G., Denk H., Zatlouk K., Müller H. Causability and explainability of artificial intelligence in medicine. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 2019, vol. 9, no. 4, pp. e1312. <https://doi.org/10.1002/widm.1312>
43. Miller T. Explanation in artificial intelligence: insights from the social sciences. *Artificial Intelligence*, 2019, vol. 267, pp. 1–38. <https://doi.org/10.1016/j.artint.2018.07.007>
44. Kulesza T., Burnett M., Wong W., Stumpf S. Principles of explanatory debugging to personalize interactive machine learning. *Proc. of the 20th International Conference on Intelligent User Interfaces (IUI '15)*, 2015, pp. 126–137. <https://doi.org/10.1145/2678025.2701399>
45. Tintarev N. Explaining recommendations. *Lecture Notes in Computer Science*, 2007, vol. 4511, pp. 470–474. https://doi.org/10.1007/978-3-540-73078-1_67
46. Chrysostomou G., Alertis N. Improving the faithfulness of attention-based explanations with task-specific information for text classification. *arXiv*, 2021, arXiv:2105.02657v2. <https://doi.org/10.48550/arXiv.2105.02657>
47. Vilone G., Longo L. Explainable artificial intelligence: A systematic review. *arXiv*, 2020, arXiv:2006.00093v3. <https://doi.org/10.48550/arXiv.2006.00093>
48. Papenmeier A., Engleblenne G., Seifert C. How model accuracy and explanation fidelity influence user trust. *arXiv*, 2019, arXiv:1907.12652v1. <https://doi.org/10.48550/arXiv.1907.12652>
49. Harutyunyan H., Achille A., Paolini G., Majumder O., Ravichandran A., Bhotika R., Soatto S. Estimating informativeness of samples with smooth unique information. *arXiv*, 2021, arXiv:2101.06640v1. <https://doi.org/10.48550/arXiv.2101.06640>
50. Liu S., Wang X., Liu M., Zhu J. Towards better analysis of machine learning models: a visual analytics perspective. *Visual Informatics*, 2017, vol. 1, no. 1, pp. 48–56. <https://doi.org/10.1016/j.visinf.2017.01.006>
51. Arrieta A.B., Díaz-Rodríguez N., Del Ser J., Bennetot A., Tabik S., Barbado A., García S., Gil-López S., Molina D., Benjamins R., Chatila R., Herrera F. Explainable artificial intelligence (XAI): concepts, taxonomies, opportunities and challenges toward responsible AI. *Information Fusion*, 2020, vol. 58, pp. 82–115. <https://doi.org/10.1016/j.inffus.2019.12.012>
52. Girshick R., Donahue J., Darrell T., Malik J. Rich feature hierarchies for accurate object detection and semantic segmentation. *Proc. of the IEEE Conference on Computer Vision and Pattern Recognition*, 2014, pp. 580–587. <https://doi.org/10.1109/CVPR.2014.81>
53. Ancona M., Ceolini E., Öztireli C., Gross M. Towards better understanding of gradient-based attribution methods for deep neural networks. *arXiv*, 2018, arXiv:1711.06104v4. <https://doi.org/10.48550/arXiv.1711.06104>
54. Rumelhart D.E., Hinton G.E., Williams R.J. Learning internal representations by error propagation. *Readings in Cognitive Science: A Perspective from Psychology and Artificial Intelligence*, 2013, pp. 399–421.

- methods // Lecture Notes in Computer Science. 2019. V. 11700. P. 267–280. https://doi.org/10.1007/978-3-030-28954-6_14
56. Roscher R., Bohn B., Duarte M.F., Garcke J. Explainable machine learning for scientific insights and discoveries // arXiv. 2020. arXiv:1905.08883v3. <https://doi.org/10.48550/arXiv.1905.08883>
55. Kindermans P.-J., Hooker S., Adebayo J., Alber M., Schütt K.T., Dähne S., Erhan D., Kim B. The (Un) reliability of saliency methods. *Lecture Notes in Computer Science*, 2019, vol. 11700, pp. 267–280. https://doi.org/10.1007/978-3-030-28954-6_14
56. Roscher R., Bohn B., Duarte M.F., Garcke J. Explainable machine learning for scientific insights and discoveries. *arXiv*, 2020, arXiv:1905.08883v3. <https://doi.org/10.48550/arXiv.1905.08883>

Авторы

Бирюков Денис Николаевич — доктор технических наук, профессор, начальник кафедры, Военно-космическая академия имени А.Ф. Можайского, Санкт-Петербург, 197198, Российская Федерация, [sc 57188163400](https://orcid.org/0000-0003-1300-2470), <https://orcid.org/0000-0003-1300-2470>, Biryukov.D.N@yandex.ru

Дудкин Андрей Сергеевич — кандидат технических наук, доцент, заместитель начальника кафедры, Военно-космическая академия имени А.Ф. Можайского, Санкт-Петербург, 197198, Российская Федерация, [sc 57211979130](https://orcid.org/0000-0003-0283-9048), <https://orcid.org/0000-0003-0283-9048>, andry-ll@mail.ru

Denis N. Biryukov — D.Sc., Professor, Head of Department, Mozhaisky Military Aerospace Academy, Saint Petersburg, 197198, Russian Federation, [sc 57188163400](https://orcid.org/0000-0003-1300-2470), <https://orcid.org/0000-0003-1300-2470>, Biryukov.D.N@yandex.ru

Andrey S. Dudkin — PhD, Associate Professor, Deputy Head of Department, Mozhaisky Military Aerospace Academy, Saint Petersburg, 197198, Russian Federation, [sc 57211979130](https://orcid.org/0000-0003-0283-9048), <https://orcid.org/0000-0003-0283-9048>, andry-ll@mail.ru

Статья поступила в редакцию 03.02.2025
Одобрена после рецензирования 11.04.2025
Принята к печати 26.05.2025

Received 03.02.2025
Approved after reviewing 11.04.2025
Accepted 26.05.2025



Работа доступна по лицензии
 Creative Commons
 «Attribution-NonCommercial»

ОПТИЧЕСКИЕ СИСТЕМЫ И ТЕХНОЛОГИИ OPTICAL ENGINEERING

doi: 10.17586/2226-1494-2025-25-3-387-395

УДК 618.7.03 + 628.58 + 57.083

Интенсификация золь-гель синтеза Mn-содержащих материалов системы $\text{MgO-Al}_2\text{O}_3\text{-ZrO}_2\text{-SiO}_2$

Сергей Константинович Евстропьев¹, Валентина Леонидовна Столярова²,
Дмитрий Владимирович Булыга³✉, Артем Сергеевич Саратовский⁴,
Николай Бабкенович Князян⁵, Гоарик Габриэловна Манукян⁶

^{1,3} АО «НПО ГОИ им. С.И. Вавилова», Санкт-Петербург, 192171, Российская Федерация

^{1,3,4} Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация

^{2,4} Институт химии силикатов им. И. В. Гребенщикова Российской академии наук, Санкт-Петербург, 199034, Российская Федерация

² Санкт-Петербургский государственный университет, Санкт-Петербург, 199034, Российская Федерация

^{5,6} Институт общей и неорганической химии им. М. Манвеляна НАН РА, Ереван, 0051, Армения

¹ evstropiev@bk.ru, <https://orcid.org/0000-0002-0160-8443>

² stvl08@inbox.ru, <https://orcid.org/0000-0001-7798-4405>

³ dmbulyga@yandex.ru✉, <https://orcid.org/0000-0001-8861-2527>

⁴ saratovskija@inbox.ru, <https://orcid.org/0000-0001-5171-6912>

⁵ knigo51@mail.ru, <https://orcid.org/0000-0001-9500-0588>

⁶ gmanukyan@sci.am, <https://orcid.org/0000-0001-8993-0471>

Аннотация

Введение. Стеклообразные и стеклокристаллические материалы системы $\text{MgO-Al}_2\text{O}_3\text{-SiO}_2$ имеют множество практических применений, в том числе их использование в качестве люминофоров. Актуальной задачей является понижение температуры синтеза таких материалов. **Метод.** В работе золь-гель методом синтезированы Mn-содержащие материалы системы $\text{MgO-Al}_2\text{O}_3\text{-ZrO}_2\text{-SiO}_2$. Аналитический химический состав, кристаллическая структура, морфология и спектры люминесценции исследованы методами рентгенофазового анализа, растровой электронной микроскопии, энергодисперсионного анализа и люминесцентной спектроскопии. **Основные результаты.** Установлено, что введение фторидного компонента в золи существенно ускоряет кристаллизацию Mn-содержащих гелей в процессе их термообработки и оказывает существенное влияние на морфологию ксерогелей. Фториды играют роль дополнительных центров зародышеобразования и обеспечивают формирование многочисленных мелких оксидных кристаллов. Энергодисперсионный анализ показал, что фтор полностью удаляется из структуры материалов при термообработке гелей до 900 °C. По данным рентгенофазового анализа внедрение ионов марганца в структуру формирующихся оксидных кристаллов и деформация их кристаллической решетки происходит на начальных стадиях процесса кристаллизации. В спектрах фотолюминесценции ксерогелей наблюдаются полосы эмиссии как ионов марганца, так и структурных дефектов, сформировавшихся в кристаллической решетке оксидных кристаллов. **Обсуждение.** Показано, что, помимо использования широко известного золь-гель метода, введение фтор-содержащего прекурсора значительно ускоряет кристаллизацию гелей системы $\text{MgO-Al}_2\text{O}_3\text{-ZrO}_2\text{-SiO}_2$, способствует формированию дисперсной структуры материалов, повышает интенсивность и улучшает разрешение полос эмиссии в спектрах люминесценции.

Ключевые слова

золь-гель метод, термическая обработка, структурные дефекты, люминесценция, фтор

Ссылка для цитирования: Евстропьев С.К., Столярова В.Л., Булыга Д.В., Саратовский А.С., Князян Н.Б., Манукян Г.Г. Интенсификация золь-гель синтеза Mn-содержащих материалов системы $\text{MgO-Al}_2\text{O}_3\text{-ZrO}_2\text{-SiO}_2$ // Научно-технический вестник информационных технологий, механики и оптики. 2025. Т. 25, № 3. С. 387–395.
doi: 10.17586/2226-1494-2025-25-3-387-395

Intensification of sol-gel synthesis of Mn-containing $\text{MgO-Al}_2\text{O}_3\text{-ZrO}_2\text{-SiO}_2$ system materials

Sergey K. Evstropiev¹, Valentina L. Stolyarova², Dmitry V. Bulyga³✉, Artem S. Saratovskii⁴, Nikolay B. Knyazyan⁵, Goharik G. Manukyan⁶

^{1,3} RPA "Vavilov State Optical Institute", Saint Petersburg, 192171, Russian Federation

^{1,3,4} ITMO University, Saint Petersburg, 197101, Russian Federation

^{2,4} Institute of Silicate Chemistry, Russian Academy of Sciences (ISC RAS), Saint Petersburg, 199034, Russian Federation

² St. Petersburg State University (SPbSU), Saint Petersburg, 199034, Russian Federation

^{5,6} Institute of General and Inorganic Chemistry named after Academician Manvelyan (NAS RA), Erevan, 0051, Armenia

¹ evstropiev@bk.ru, <https://orcid.org/0000-0002-0160-8443>

² stvl08@inbox.ru, <https://orcid.org/0000-0001-7798-4405>

³ dmbulyga@yandex.ru✉, <https://orcid.org/0000-0001-8861-2527>

⁴ saratovskija@inbox.ru, <https://orcid.org/0000-0001-5171-6912>

⁵ knigo51@mail.ru, <https://orcid.org/0000-0001-9500-0588>

⁶ gmanukyan@sci.am, <https://orcid.org/0000-0001-8993-0471>

Abstract

Glass and glass-crystalline $\text{MgO-Al}_2\text{O}_3\text{-SiO}_2$ system materials have many practical applications including their use as luminophores. To lower the synthesis temperature of such materials is an actual task. In this work, Mn-containing materials of $\text{MgO-Al}_2\text{O}_3\text{-ZrO}_2\text{-SiO}_2$ system were synthesized by sol-gel method. The analytical chemical composition, crystal structure, morphology and luminescence spectra were investigated by X-ray phase analysis, scanning electron microscopy, energy dispersive analysis and luminescence spectroscopy. It was found that the introduction of fluoride component into sols significantly accelerates the crystallization of Mn-containing gels during their heat treatment and has a significant effect on the morphology of xerogels. Fluorides play the role of additional nucleation centers and ensure the formation of numerous small oxide crystals. Energy dispersive analysis showed that fluoride is completely removed from the structure of materials during heat treatment of gels up to 900 °C. According to the data of X-ray phase analysis, the introduction of manganese ions into the structure of forming oxide crystals and deformation of their crystal lattice occurs at the initial stages of the crystallization process. Emission bands of both manganese ions and structural defects formed in the crystal lattice of oxide crystals are observed in the photoluminescence spectra of xerogels. It was shown that in addition to using the sol-gel method, which is a well-known approach, the addition of fluorine-containing precursor significantly accelerates crystallization of gels of $\text{MgO-Al}_2\text{O}_3\text{-ZrO}_2\text{-SiO}_2$ system, promotes formation of dispersed structure of materials, increases intensity, and improves resolution of emission bands in luminescence spectra.

Keywords

sol-gel method, thermal treatment, structural defects, luminescence, fluorine

For citation: Evstropiev S.K., Stolyarova V.L., Bulyga D.V., Saratovskii A.S., Knyazyan N.B., Manukyan G.G. Intensification of sol-gel synthesis of Mn-containing $\text{MgO-Al}_2\text{O}_3\text{-ZrO}_2\text{-SiO}_2$ system materials. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2025, vol. 25, no. 3, pp. 387–395 (in Russian). doi: 10.17586/2226-1494-2025-25-3-387-395

Введение

Разработка новых и совершенствование известных методов получения различных оксидных материалов является традиционной и актуальной проблемой. Стекла и стеклокристаллические материалы системы $\text{MgO-Al}_2\text{O}_3\text{-SiO}_2$ перспективны и широко используются в различных практических приложениях [1–9]. Традиционная методика синтеза требует применения высоких температур (не менее 1550 °C) [5, 6]. Исходя из этого, для получения материалов системы $\text{MgO-Al}_2\text{O}_3\text{-SiO}_2$ часто используются добавки фторидов в исходную шихту, понижающие температуру плавления компонентов, или низкотемпературные методы синтеза [7, 8].

Хорошо известно, что применение золь-гель метода позволяет существенно понизить температуру синтеза различных оксидных материалов [7–14]. Так, в работах [12, 13] низкотемпературным золь-гель синтезом были получены наночастицы люминофора $\text{Zn}_2\text{SiO}_4\text{:Mn}^{2+}$ и исследованы их структура и оптические свойства.

В работе [4] показано, что фторидные добавки способствуют кристаллизации стекол системы $\text{Li}_2\text{O-Al}_2\text{O}_3\text{-SiO}_2$,

SiO_2 , понижая температуру начала кристаллизации и повышая механические свойства полученных материалов. В [14] установлено, что добавки фторидов ускоряют кристаллизацию цинковоборосиликатных стекол. В работе [15] описано влияние добавки фторида лития на ускорение образования кристаллов форстерита (Mg_2SiO_4) при их твердотельном синтезе из порошков оксида магния (MgO) и диоксида кремния (SiO_2).

Кроме воздействия на структуру материалов, добавки фторидов используются для повышения люминесцентных характеристик различных оксидных люминофоров [16–18]. Применение фторидов металлов в золь-гель композиционных покрытиях на основе SiO_2 рассмотрено в [19].

Mn-содержащие люминесцентные материалы системы $\text{MgO-Al}_2\text{O}_3\text{-SiO}_2$ перспективны для различных практических приложений [20–24]. Вместе с тем в настоящее время дальнейшая разработка и оптимизация методов их синтеза является актуальной проблемой. Однако исследования по влиянию добавок фторидов на формирование системы $\text{MgO-Al}_2\text{O}_3\text{-SiO}_2$ золь-гель методом до настоящего времени не проводились.

Целью настоящей работы стало исследование влияния добавки фтора на процессы формирования при золь-гель синтезе кристаллической структуры, морфологии и люминесцентных свойств Mn-содержащих материалов системы $\text{MgO-Al}_2\text{O}_3\text{-ZrO}_2\text{-SiO}_2$.

Материалы и методы

Композиционные золи были получены смешением в заданных объемах водных растворов нитрата алюминия ($\text{Al}(\text{NO}_3)_3$), нитрата магния ($\text{Mg}(\text{NO}_3)_2$), оксид-дихлорида циркония (ZrOCl_2), сульфата марганца и поливинилпирролидона (ПВП; молекулярная масса 25 000–35 000 г/моль) с тетраэтоксисилоаном (ТЕОС) и пропанолом-2. В качестве источника фторид-анионов использован водный раствор $\text{NH}_4\text{F}\cdot\text{HF}$ (концентрация 5 масс.%). Химический состав исходных растворов представлен в табл. 1.

Процесс гелеобразования протекал при комнатной температуре в течение двух суток. После завершения гелеобразования образцы подвергались сушке при 70 °С с последующей термообработкой при температурах 900 °С или 1150 °С в воздушной атмосфере. Полученные образцы представляли собой белые однородные высокодисперсные порошки. Аналитический состав материалов, прокаленных при температуре 1150 °С, приведен в табл. 2.

Из приведенных в табл. 2 данных видно, что в составе всех синтезированных материалов отсутствуют какие-либо соединения фтора. Это свидетельствует об улетучивании фтора в процессе термообработки гелей и объясняется высокой летучестью фторидных соединений.

Кристаллическая структура материалов была исследована методом рентгенофазового анализа на приборе Rigaku SmartLab 3 (излучение $\text{CuK}\alpha$ ($\lambda = 0,015418$ нм, 40 кВ, 44 мА)). Для оценки среднего размера кристаллитов в материалах применялась формула Шеррера. При этих расчетах использовались данные наиболее интенсивных пиков кристаллов.

Изучение морфологии и аналитического химического состава полученных материалов выполнялось методами электронно-микроскопического и энергодисперсионного анализов, используя прибор VEGA3 TESCAN с приставкой Advanced Aztec Energy (Oxford Instruments).

Измерения спектров фотолуминесценции образцов проводилось на спектрофлуориметре Perkin Elmer LS 50B в спектральном диапазоне.

Результаты и обсуждения

Рис. 1 представляет рентгенограммы гелей, термообработанных при различных температурах. На рентгенограммах гелей, прокаленных при 900 °С, наблюдаются пики, характерные для кубических кристаллов шпинели (MgAl_2O_4), ZrO_2 , твердых растворов со структурой β -кварца. В гелях, прокаленных при 1150 °С, также присутствуют фазы ZrSiO_4 , Mg_2SiO_4 , энстатита (MgSiO_3), кордиерита и SiO_2 . Видно также, что значительная часть материалов сохраняется в аморфном состоянии.

Известно [9], что кристаллы ZrO_2 формируются на начальных стадиях кристаллизации стекол системы $\text{MgO-Al}_2\text{O}_3\text{-ZrO}_2\text{-SiO}_2$ и играют роль нуклеаторов для образования других кристаллов. Кристаллы MgAl_2O_4 и твердые растворы со структурой β -кварца также формируются на начальных стадиях кристаллизации стекол [2, 6, 9].

Следует отметить схожесть характера структурной эволюции материалов на начальных этапах кристаллизации гелей и стекол системы $\text{MgO-Al}_2\text{O}_3\text{-ZrO}_2\text{-SiO}_2$. Несмотря на близость химического состава гелей и стекол это явление не является очевидным. Начальные стадии процессов кристаллизации стекол протекают в объеме плотной аморфной среды, в то время как кристаллизация в гелях протекает в условиях высокой дисперсности материала с относительно большим пространственным разделением компонентов.

Образование в гелях кристаллов силикатов магния — MgSiO_3 и Mg_2SiO_4 наблюдается при повыше-

Таблица 1. Химический состав исходных растворов, масс.%

Table 1. Chemical compositions of the initial solutions, mol.%

Номер раствора	$\text{Al}(\text{NO}_3)_3$	$\text{Mg}(\text{NO}_3)_2$	ZrOCl_2	H_2O	ПВП	Пропанол-2	$\text{Mn}(\text{NO}_3)_2$	$\text{NH}_4\text{F}\cdot\text{HF}$	ТЕОС
1	7,59	5,27	1,74	47,89	0,07	5,34	0,04	—	32,06
2	7,12	4,95	1,63	44,95	0,07	5,02	0,04	6,13	30,09
3	7,59	5,27	1,74	47,91	0,07	5,35	—	—	32,07

Таблица 2. Аналитический химический состав гелей, мол.%

Table 2. Analytical chemical composition of the gels, mol.%

Образец	Компоненты				
	SiO_2	Оксид алюминия (Al_2O_3)	Оксид циркония (ZrO_2)	MgO	Оксид марганца
1	58,3	11,5	6,0	23,2	1,0
2	68,6	16,4	2,4	12,0	0,6
3	58,9	11,6	6,1	23,4	—

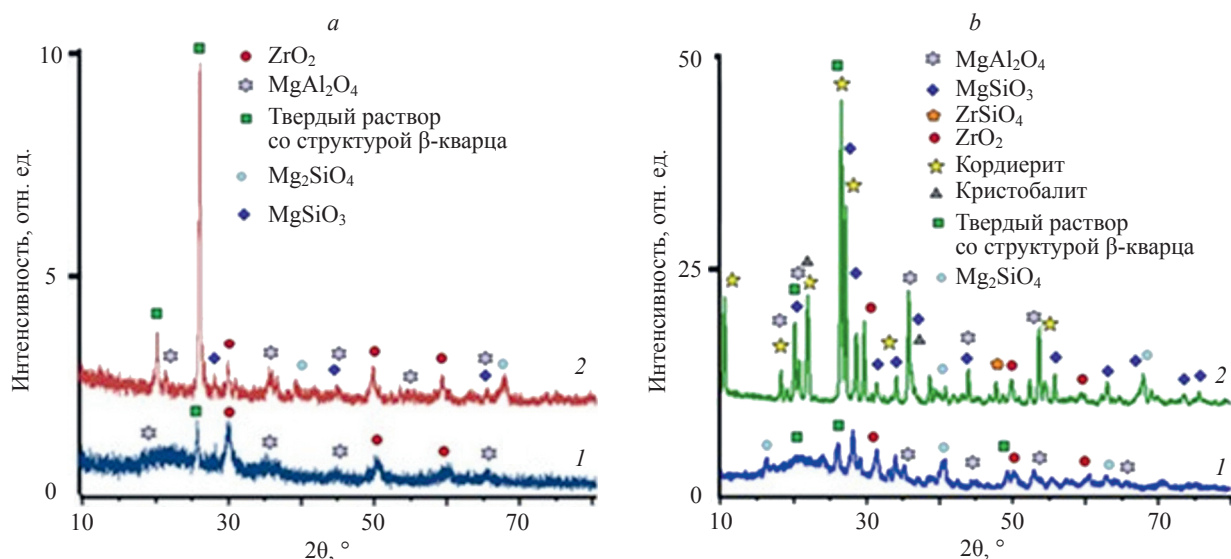


Рис. 1. Рентгенограммы гелей 1 (a) и 2 (b), термообработанных при температурах 900 °С (кривые 1) и 1150 °С (кривые 2)
 Fig. 1. XRD patterns of gels 1 (a) and 2 (b), calcined at 900 °С (curves 1) and 1150 °С (curves 2)

нии температуры термообработки до 1150 °С. Пики кордиеритовых кристаллов проявляется на рентгенограмме геля 2 после его термообработки при 1150 °С. Аналогичное явление получено ранее в [8] при кристаллизации кордиеритовых гелей при 1100 °С.

Сопоставление дифрактограмм, приведенных на рис. 1, показывает существенное различие между кристаллическими структурами гелей 1 и 2. Количество пиков различных кристаллов и их интенсивность значительно выше на дифрактограмме геля 2, полученного с применением фторидной добавки.

Средние размеры некоторых кристаллов, сформировавшихся в гелях, приведены в табл. 3. Размеры кристаллов MgAl_2O_4 и твердых растворов со структурой β -кварца, сформировавшихся при термообработке гелей 1 и 2 при 1150 °С составляют 30–34 нм. Учитывая, что интенсивность пиков на дифрактограмме геля 2 (рис. 1, b) значительно выше, чем у геля 1 (рис. 1, a) можно предположить, что фтор играет роль нуклеатора на начальных стадиях кристаллизации гелей. Это определяет формирование на начальных стадиях кристаллизации значительного большего количества различных мелких кристаллов в геле 2 по сравнению с гелем 1.

Отметим отсутствие пиков каких-либо соединений марганца на дифрактограммах (рис. 1). Учитывая низкое содержание марганца в гелях, можно предположить, что это явление связано со способностью ионов марганца встраиваться в структуру кристаллов при кристаллизации материалов системы $\text{MgO-Al}_2\text{O}_3\text{-SiO}_2$ [20–23].

Результаты сопоставления позиций наиболее интенсивных пиков кубических кристаллов MgAl_2O_4 на дифрактограммах приведены в табл. 4. Из полученных результатов видно, что наблюдается заметный сдвиг пиков в сторону меньших значений 2θ по сравнению со стандартными величинами (JCPDS 89-1627).

Это явление объясняется некоторым увеличением объема элементарной ячейки кристаллов MgAl_2O_4 при замещении небольших ионов Mg^{2+} (ионный радиус 0,0057 нм [25]) на большие по размеру ионы Mn^{2+} (ионный радиус 0,066 нм [25]). Следует также отметить, что этот сдвиг пиков наблюдается и на дифрактограммах гелей, термообработанных при 900 °С. Этот факт позволяет сделать вывод о том, что внедрение ионов марганца в структуру кристаллов MgAl_2O_4 протекает уже на начальных этапах кристаллизации.

Таблица 3. Средние размеры кристаллов, сформировавшихся в гелях при их термообработке, нм

Table 3. Average sizes of crystals formed in the gels during the calcination, nm

Материал	Кристаллическая фаза	При температуре термообработки	
		900 °С	1150 °С
Гель 1	ZrO_2	22	36
	MgAl_2O_4	26	30
	Твердый раствор со структурой β -кварца	30	33
Гель 2	ZrO_2	15	26
	MgAl_2O_4	10	34
	Твердый раствор со структурой β -кварца	16	31
	Кордиерит	—	30

Таблица 4. Результаты сопоставления положения на дифрактограммах гелей наиболее интенсивных пиков кубических кристаллов MgAl_2O_4 , 2θ , град.

Table 4. Results of the comparison of the most intense peaks position of MgAl_2O_4 in the XRD patterns of gels, 2θ , degree

(hkl)	Положение пика согласно стандарту JCPDS 89-1627	Гель 1		Гель 2	
		при температуре термообработки			
		900 °C	1150 °C	900 °C	1150 °C
(111)	19,30	19,14	19,07	19,10	19,14
(311)	36,96	36,76	36,79	36,85	36,68
(440)	65,70	65,54	65,39	65,50	65,53

Примечание. hkl — обозначения кристаллографических плоскостей.

Электронно-микроскопические снимки полученных материалов приведены на рис. 2. Сравнение снимков гелей, подвергнутых термообработке при различных температурах, демонстрирует существенное влияние фтора на морфологию материалов. Гель 1, термообработанный при 900 °C, состоит из частиц размером около 0,5 мкм (рис. 2, *a, b*). Повышение температуры термообработки до 1150 °C приводит к росту частиц размерами до 5 мкм (рис. 2, *c*) и их частичному спеканию (рис. 2, *c, d*).

Рис. 2, *e, f* показывает, что гель 2, полученный с использованием фторидной добавки имеет рыхлую структуру, состоящую из маленьких частиц размером около 0,1 мкм. Увеличение температуры термообработки приводит к формированию агрегатов частиц (рис. 2, *g, h*).

Полученные ксерогели 1, 2 и 3 являются высокодисперсными материалами и состоят из различных нанокристаллов и остаточной аморфной фазы. Известно, что такие дисперсные материалы содержат множество собственных поверхностных и объемных структур-

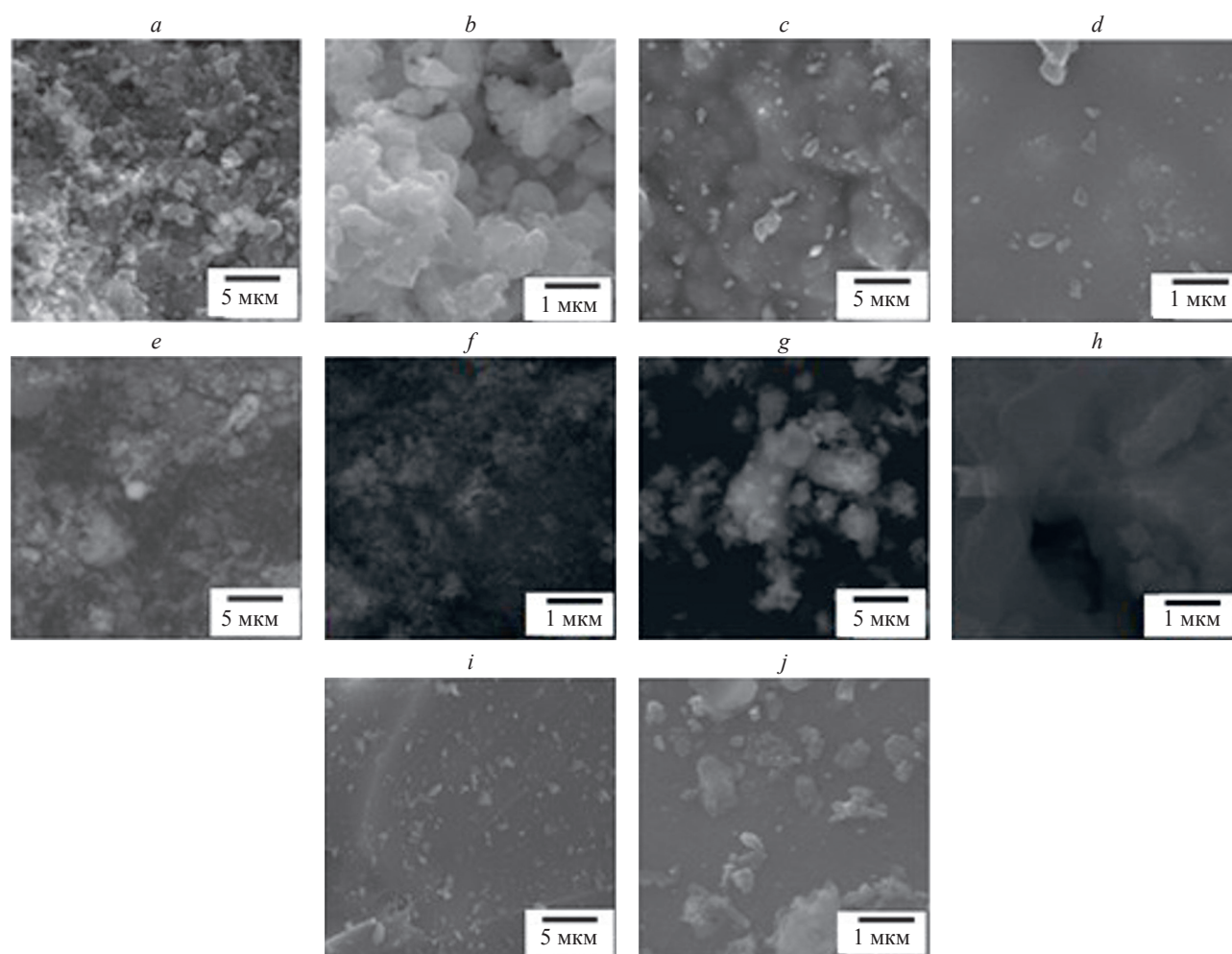


Рис. 2. Электронно-микроскопические снимки образцов гелей 1 (*a–d*), 2 (*e–h*) и 3 (*i, j*) при разном увеличении, термообработанных при 900 °C (*a, b, e, f*) и 1150 °C (*c, d, g, h, i, j*)

Fig. 2. SEM images of gel samples at different magnifications: 1 (*a–d*), 2 (*e–h*) and 3 (*i, j*), heat-treated at 900 °C (*a, b, e, f*) and 1150 °C (*c, d, g, h, i, j*)

ных дефектов [26]. Присутствие таких структурных дефектов часто определяет интенсивную люминесценцию дисперсных материалов в видимой части спектра [20, 27–29]. Кроме того, некоторая часть нанокристаллов в Mn-содержащих гелях деформирована из-за внедрения в их кристаллическую решетку ионов марганца.

Рис. 3 демонстрирует спектры фотолюминесценции гелей, термообработанных при 900 °C в видимом спектральном диапазоне. В спектрах наблюдаются многочисленные пики эмиссии ионов марганца и собственных дефектов в структуре нанокристаллов. Невысокая интенсивность пиков и их взаимное наложение затрудняют их точную идентификацию. Тем не менее, сравнение спектров фотолюминесценции Mn-содержащих гелей 1 (кривые 1) и 2 (кривые 2) со спектрами геля 3, не содержащего марганца (кривые 3), позволяет идентифицировать полосы эмиссии, связанные с ионами марганца.

Полосы эмиссии, расположенные в видимой области спектра (455–460 нм), наблюдаются в спектрах всех

гелей и могут относиться к структурным дефектам в кристаллах MgAl_2O_4 [26, 29]. Так, полоса люминесценции с $\lambda_{\text{max}} = 460$ нм наблюдалась в спектре эмиссии MgAl_2O_4 кристаллов под действием ультрафиолетового излучения [30]. В [26] наличие этой полосы предположительно связано с процессами возбуждения и рекомбинации положительно заряженных кислородных вакансий в кристалле MgAl_2O_4 .

Зеленая люминесценция в Mn-содержащих материалах обычно приписывается эмиссии ионов Mn^{2+} , находящихся в тетраэдрической координации [11, 13, 20, 21, 23, 24]. Сопоставление спектров эмиссии гелей 1 и 2, содержащих ионы марганца (кривые 1, 2 на рис. 3, а, б) и геля 3, не содержащего марганца, позволяет предположить, что полосы с $\lambda_{\text{max}} = 490$ нм и $\lambda_{\text{max}} = 525$ нм определяются эмиссией ионов марганца. Полоса с $\lambda_{\text{max}} = 525$ нм, связанная с люминесценцией ионов Mn^{2+} в тетраэдрической координации, наблюдалась в [11, 13] при возбуждении синим светом ($\lambda_{\text{ex}} = 450$ нм) Mn-содержащих кристаллов

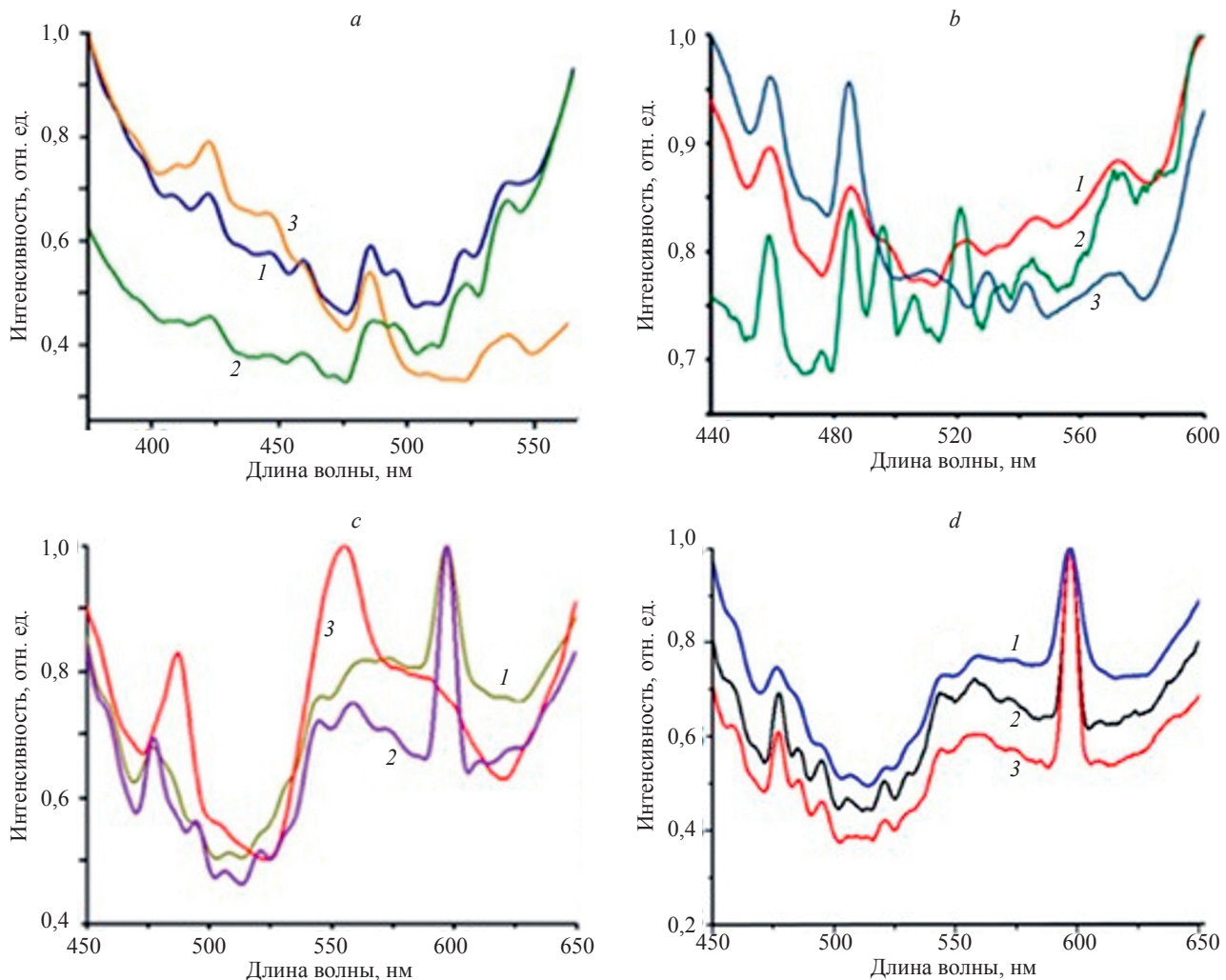


Рис. 3. Спектры фотолюминесценции, полученные после термообработки при 900 °C для гелей: 1 (кривая 1), 2 (кривая 2) и 3 (кривая 3) при длинах волн возбуждения: 300 нм (а), 350 нм (б) и 400 нм (с), и геля 1 после термообработки при 1150 °C (кривая 1), геля 2 — при 900 °C (кривая 2) и геля 2 — при 1150 °C (кривая 3) при длине волны возбуждения 400 нм (д)

Fig. 3. Photoluminescence spectra obtained after heat treatment at 900 °C for gels: 1 (curve 1), 2 (curve 2) and 3 (curve 3) at excitation wavelengths: 300 nm (a), 350 nm (b) and 400 nm (c), and gel 1 after heat treatment at 1150 °C (curve 1), gel 2 at 900 °C (curve 2) and gel 2 at 1150 °C (curve 3) at an excitation wavelength of 400 nm (d)

MgAl₂O₄. Аналогичная полоса эмиссии проявлялась также в спектрах люминесценции нанокompозита SiO₂/Zn₂SiO₄:Mn [31].

Эмиссия Mn-содержащих материалов в оранжево-красной области спектра определяется как эмиссией ионов марганца (ионы Mn²⁺ в октаэдрической координации и ионы Mn⁴⁺ [23]), так и собственных структурных дефектов кристаллов [20]. Полоса эмиссии, расположенная в спектральном диапазоне 560–608 нм наблюдалась ранее в нанокompозитах SiO₂/Zn₂SiO₄:Mn [13] и нанокристаллах Zn₂SiO₄:Mn²⁺ и приписывалась электронному переходу ⁴T₁(⁴G) → ⁶A₁(⁶S) ионов Mn²⁺ [32]. Однако полоса эмиссии в этом спектральном диапазоне наблюдается и в спектре люминесценции геля 3, не содержащего марганца (рис. 3, а). Полученные результаты позволяют предположить, что данная полоса может иметь сложный характер и являться результатом наложения пиков эмиссии ионов Mn²⁺ и структурных дефектов кристаллов.

Сопоставление кривых 1 (рис. 3, а, б) и 2 (рис. 3, б) показывает, что добавки фторида в золь влияют на форму полос эмиссии, делая эти полосы более разрешенными. Такие различия в спектрах люминесценции могут объясняться существенно большей кристаллическостью ксерогеля 2 по сравнению с ксерогелем 1 (рис. 1).

Таким образом, полученные результаты показывают, что добавки фторида в исходную смесь могут существенно увеличить кристалличность полученных

золь-гель методом Mn-содержащих материалов системы MgO-Al₂O₃-ZrO₂-SiO₂ и оказывают заметное влияние на их люминесцентные свойства, повышая разрешение полос эмиссии в спектрах люминесценции.

Заключение

Mn-содержащие материалы системы MgO-Al₂O₃-ZrO₂-SiO₂ были синтезированы золь-гель методом. В работе изучено влияние фторидной добавки в золи на кристаллическую структуру, морфологию и люминесцентные свойства материалов. При термообработке гелей до температуры от 900 °C наблюдается формирование различных оксидных нанокристаллов (MgAl₂O₄; ZrO₂ и др.). Экспериментальные результаты показали, что добавка фторидов оказывает сильное влияние на кристаллическую структуру, морфологию и люминесцентные свойства Mn-содержащих материалов системы MgO-Al₂O₃-ZrO₂-SiO₂. Фторидная добавка значительно ускоряет кристаллизацию гелей при их термообработке при температурах от 900 °C. Данные рентгенофазового анализа показали, что при термообработке происходит внедрение ионов Mn²⁺ в кристаллическую решетку MgAl₂O₄ уже на начальных этапах кристаллизации гелей. Добавка фторида в золи оказывает заметное влияние на люминесцентные свойства полученных при термообработке ксерогелей, повышая разрешение полос эмиссии в спектрах люминесценции.

Литература

1. Chen G.-H., Liu X.-Y. Sintering, crystallization and properties of MgO-Al₂O₃-SiO₂ system glass-ceramics containing ZnO // *Journal of Alloys and Compounds*. 2007. V. 431. N 1-2. P. 282–286. <https://doi.org/10.1016/j.jallcom.2006.05.060>
2. Bortkevich A.V., Dymshits O.S., Zhilin A.A., Polushkin A.Yu., Tsenter M.Ya., Shashkin A.V., Golubkov V.V., B'en V.-B., Li K.-K., Pak E.-B., Pak K.H. Study of phase transformations in titanium-containing magnesium-aluminum silicate glasses and glass-ceramics for diffuse reflectors // *Journal of Optical Technology*. 2002. V. 69. N 8. P. 558–594. <https://doi.org/10.1364/JOT.69.000588>
3. Evstropiev S.K., Yurchenko D.A., Stolyarova V.L., Knyazyan N.B., Manukyan G.G., Shashkin A.V. Some features of the surface modification of MgO-Al₂O₃-TiO₂-SiO₂ glass and glass ceramics by Ag diffusion // *Ceramics International*. 2022. V. 48. N 17. P. 24517–24522. <https://doi.org/10.1016/j.ceramint.2022.05.090>
4. Guo X., Yang H. Effects of fluorine on crystallization, structure and performances of lithium aluminosilicate glass ceramic // *Materials Research Bulletin*. 2006. V. 41. N 2. P. 396–405. <https://doi.org/10.1016/j.materresbull.2005.08.002>
5. Guo H., Liu X.Y., Li F., Wei R.F., Wei Y.L., Ma C. Enhanced white luminescence in mixed-valence Eu-doped BaAl₂Si₂O₈ glass ceramics for W-LEDs // *Journal of the Electrochemical Society*. 2012. V. 159. N 6. P. J223–J226. <https://doi.org/10.1149/2.jes113286>
6. Evstropiev S.K., Shashkin A.V., Knyazyan N.B., Manukyan G.G., Bagramyan V.V., Timchuk A.V., Stolyarova V.L. Eu-doped BaO-Al₂O₃-SiO₂-MgF₂ glass and glass ceramics // *Journal of Non-Crystalline Solids*. 2022. V. 580. P. 121386. <https://doi.org/10.1016/j.jnoncrystsol.2021.121386>
7. Evstrop'ev S.K., Volynkin V.M., Saratovskii A.S., Danilovich D.P., Demidov V.V., Dukel'skii K.V., Bulyga D.V., Sysolyatin S.O. Modification of quartz ceramics by applying a sol-gel composition of MgO-Al₂O₃-ZrO₂-SiO₂ system // *Russian Journal of Applied Chemistry*. 2023. V. 96. N 2. P. 190–197. <https://doi.org/10.1134/S1070427223020090>
8. Petrović R., Janačković D., Zec S., Drmanić S. Ž., Kostić-Gvozdenović L.J. Crystallization behavior of alkoxy-derived cordierite gels // *Journal*

References

1. Chen G.-H., Liu X.-Y. Sintering, crystallization and properties of MgO-Al₂O₃-SiO₂ system glass-ceramics containing ZnO. *Journal of Alloys and Compounds*. 2007, vol. 431, no. 1-2, pp. 282–286. <https://doi.org/10.1016/j.jallcom.2006.05.060>
2. Bortkevich A.V., Dymshits O.S., Zhilin A.A., Polushkin A.Yu., Tsenter M.Ya., Shashkin A.V., Golubkov V.V., B'en V.-B., Li K.-K., Pak E.-B., Pak K.H. Study of phase transformations in titanium-containing magnesium-aluminum silicate glasses and glass-ceramics for diffuse reflectors. *Journal of Optical Technology*. 2002, vol. 69, no. 8, pp. 558–594. <https://doi.org/10.1364/JOT.69.000588>
3. Evstropiev S.K., Yurchenko D.A., Stolyarova V.L., Knyazyan N.B., Manukyan G.G., Shashkin A.V. Some features of the surface modification of MgO-Al₂O₃-TiO₂-SiO₂ glass and glass ceramics by Ag diffusion. *Ceramics International*, 2022, vol. 48, no. 17, pp. 24517–24522. <https://doi.org/10.1016/j.ceramint.2022.05.090>
4. Guo X., Yang H. Effects of fluorine on crystallization, structure and performances of lithium aluminosilicate glass ceramic. *Materials Research Bulletin*, 2006, vol. 41, no. 2, pp. 396–405. <https://doi.org/10.1016/j.materresbull.2005.08.002>
5. Guo H., Liu X.Y., Li F., Wei R.F., Wei Y.L., Ma C. Enhanced white luminescence in mixed-valence Eu-doped BaAl₂Si₂O₈ glass ceramics for W-LEDs. *Journal of the Electrochemical Society*, 2012, vol. 159, no. 6, pp. J223–J226. <https://doi.org/10.1149/2.jes113286>
6. Evstropiev S.K., Shashkin A.V., Knyazyan N.B., Manukyan G.G., Bagramyan V.V., Timchuk A.V., Stolyarova V.L. Eu-doped BaO-Al₂O₃-SiO₂-MgF₂ glass and glass ceramics. *Journal of Non-Crystalline Solids*, 2022, vol. 580, pp. 121386. <https://doi.org/10.1016/j.jnoncrystsol.2021.121386>
7. Evstrop'ev S.K., Volynkin V.M., Saratovskii A.S., Danilovich D.P., Demidov V.V., Dukel'skii K.V., Bulyga D.V., Sysolyatin S.O. Modification of quartz ceramics by applying a sol-gel composition of MgO-Al₂O₃-ZrO₂-SiO₂ system. *Russian Journal of Applied Chemistry*, 2023, vol. 96, no. 2, pp. 190–197. <https://doi.org/10.1134/S1070427223020090>
8. Petrović R., Janačković D., Zec S., Drmanić S. Ž., Kostić-Gvozdenović L.J. Crystallization behavior of alkoxy-derived cordierite

- of Sol-Gel Science and Technology, 2003. V. 28. N. 1. P. 111–118. <https://doi.org/10.1023/A:1025649406466>
9. Dittmer M., Rüssel C. Colorless and high strength $\text{MgO}/\text{Al}_2\text{O}_3/\text{SiO}_2$ glass-ceramic dental material using zirconia as nucleating agent // *Journal of Biomedical Materials Research Part B Applied Biomaterials*. 2012. V. 100. N 2. P. 463–470. <https://doi.org/10.1002/jbm.b.31972>
 10. Ni S., Chou L., Chang J. Preparation and characterization of forsterite (Mg_2SiO_4) bioceramics // *Ceramics International*. 2007. V. 33. N 1. P. 83–88. <https://doi.org/10.1016/j.ceramint.2005.07.021>
 11. Khaidukov N.M., Brekhovskikh M.N., Kirikova N.Yu., Kondratyuk V.A., Makhov V.N. Luminescence of MgAl_2O_4 and ZnAl_2O_4 spinel ceramics containing some 3d ions // *Ceramics International*. 2020. V. 46. N 13. P. 21351–21359. <https://doi.org/10.1016/j.ceramint.2020.05.231>
 12. Lukić S.R., Petrović D.M., Dramićanin M.D., Mitrić M., Dačanin Lj., Optical and structural properties of $\text{Zn}_2\text{SiO}_4\text{:Mn}^{2+}$ green phosphor nanoparticles obtained by a polymer-assisted sol-gel method // *Scripta Materialia*. 2008. V. 58. N 8. P. 655–658. <https://doi.org/10.1016/j.scriptamat.2007.11.045>
 13. Omri K., El Ghoul J., Alyamani A., Barthou C., El Mir L. Luminescence properties of green emission of $\text{SiO}_2/\text{Zn}_2\text{SiO}_4\text{:Mn}$ nanocomposite prepared by sol-gel method // *Physica E: Low-dimensional Systems and Nanostructures*. 2013. V. 53. P. 48–54. <https://doi.org/10.1016/j.physe.2013.04.020>
 14. Kullberg A.T.G., Lopes A.A.S., Monteiro R.C.C. Effect of ZnF_2 addition on the crystallization behaviour of a zinc borosilicate glass // *Journal of Non-Crystalline Solids*. 2017. V. 468. P. 100–104. <https://doi.org/10.1016/j.jnoncrysol.2017.04.030>
 15. Kan A., Hirabayashi R., Takahashi S., Ogawa H. Low-temperature crystallization and microwave dielectric properties of forsterite generated in MgO-SiO_2 system following LiF addition // *Ceramics International*. 2023. V. 49. N 6. P. 9883–9892. <https://doi.org/10.1016/j.ceramint.2022.11.163>
 16. Yoon S., Otal E.H., Maegli A.E., Karvonen L., Matam S.K., Ebbinghaus S.G., Walfort B., Hagemann H., Pokrant S., Weidenkaff A. Improved persistent luminescence of $\text{CaTiO}_3\text{:Pr}$ by fluorine substitution and thermochemical treatment // *Journal of Alloys and Compounds*. 2014. V. 613. P. 338–343. <https://doi.org/10.1016/j.jallcom.2014.06.041>
 17. Yu Y., Wang H., Li L., Chen Y., Zeng R. Effects of various fluxes on the morphology and optical properties of $\text{Lu}_{3-x}\text{Al}_5\text{O}_{12}\text{:xCe}^{3+}$ green phosphors // *Ceramics International*. 2014. V. 40. N 9. Part A. P. 14171–14175. <https://doi.org/10.1016/j.ceramint.2014.06.004>
 18. Wang H., Mao F., Liu Y., Jiang X., Ma B., Wei L., Wu F., Li L. Effect of fluxes on luminescence properties of color-tunable $\text{Ba}_{1.3}\text{Ca}_{0.7}\text{SiO}_4\text{:Eu}^{2+},\text{Mn}^{2+}$ phosphor for Near-Ultraviolet white-LEDs // *Materials Research Bulletin*. 2020. V. 125. P. 110808. <https://doi.org/10.1016/j.materresbull.2020.110808>
 19. Fujihara S., Koji S., Kimura T. Structure and optical properties of $(\text{Gd,Eu})\text{F}_3$ -nanocrystallized sol-gel silica films // *Journal of Materials Chemistry*. 2004. V. 14. N 8. P. 1331–1335. <https://doi.org/10.1039/b313784h>
 20. Евстропьев С.К., Столярова В.Л., Саратовский А.С., Бульга Д.В., Дукельский К.В., Князян Н.Б., Юрченко Д.В. Люминесцентные Mn^{2+} -содержащие золь-гель материалы системы $\text{MgO-Al}_2\text{O}_3\text{-ZrO}_2\text{-SiO}_2$ // *Журнал неорганической химии*. 2024. Т. 69. № 3. С. 394–401. <https://doi.org/10.31857/S0044457X24030134>
 21. Khaidukov N.M., Brekhovskikh M.N., Kirikova N.Yu., Kondratyuk V.A., Makhov V.N., Luminescence properties of spinels doped with manganese ions // *Russian Journal of Inorganic Chemistry*. 2020. V. 65. N 8. P. 1135–1141. <https://doi.org/10.1134/S0036023620080069>
 22. Song E., Zhou Y., Wei Y., Han X., Tao Z., Qiu R., Xia Z., Zhang Q. A thermally stable narrow-band green-emitting phosphor $\text{MgAl}_2\text{O}_4\text{:Mn}^{2+}$ for wide color gamut backlight display application // *Journal of Materials Chemistry C*. 2019. V. 7. N 27. P. 8192–8198. <https://doi.org/10.1039/c9tc02107h>
 23. El Ghoul J., Omri K., Alyamani A., Barthou C., El Mir L. Synthesis and luminescence of $\text{SiO}_2/\text{Zn}_2\text{SiO}_4$ and $\text{SiO}_2/\text{Zn}_2\text{SiO}_4\text{:Mn}$ composite with sol-gel methods // *Journal of Luminescence*. 2013. V. 138. P. 218–222. <https://doi.org/10.1016/j.jlumin.2013.02.009>
 24. Shannon R.D. Revised effective ionic radii and systematic studies of interatomic distances in halides and chalcogenides // *Acta Crystallographica Section A*. 1976. V. 32. N 5. P. 751–767. <https://doi.org/10.1107/s0567739476001551>
 - gels. *Journal of Sol-Gel Science and Technology*, 2003, vol. 28, no. 1, pp. 111–118. <https://doi.org/10.1023/A:1025649406466>
 9. Dittmer M., Rüssel C. Colorless and high strength $\text{MgO}/\text{Al}_2\text{O}_3/\text{SiO}_2$ glass-ceramic dental material using zirconia as nucleating agent. *Journal of Biomedical Materials Research Part B Applied Biomaterials*, 2012, vol. 100, no. 2, pp. 463–470. <https://doi.org/10.1002/jbm.b.31972>
 10. Ni S., Chou L., Chang J. Preparation and characterization of forsterite (Mg_2SiO_4) bioceramics. *Ceramics International*, 2007, vol. 33, no. 1, pp. 83–88. <https://doi.org/10.1016/j.ceramint.2005.07.021>
 11. Khaidukov N.M., Brekhovskikh M.N., Kirikova N.Yu., Kondratyuk V.A., Makhov V.N. Luminescence of MgAl_2O_4 and ZnAl_2O_4 spinel ceramics containing some 3d ions. *Ceramics International*, 2020, vol. 46, no. 13, pp. 21351–21359. <https://doi.org/10.1016/j.ceramint.2020.05.231>
 12. Lukić S.R., Petrović D.M., Dramićanin M.D., Mitrić M., Dačanin Lj., Optical and structural properties of $\text{Zn}_2\text{SiO}_4\text{:Mn}^{2+}$ green phosphor nanoparticles obtained by a polymer-assisted sol-gel method. *Scripta Materialia*. 2008, vol. 58, no. 8, pp. 655–658. <https://doi.org/10.1016/j.scriptamat.2007.11.045>
 13. Omri K., El Ghoul J., Alyamani A., Barthou C., El Mir L. Luminescence properties of green emission of $\text{SiO}_2/\text{Zn}_2\text{SiO}_4\text{:Mn}$ nanocomposite prepared by sol-gel method. *Physica E: Low-dimensional Systems and Nanostructures*, 2013, vol. 53, pp. 48–54. <https://doi.org/10.1016/j.physe.2013.04.020>
 14. Kullberg A.T.G., Lopes A.A.S., Monteiro R.C.C. Effect of ZnF_2 addition on the crystallization behaviour of a zinc borosilicate glass. *Journal of Non-Crystalline Solids*, 2017, vol. 468, pp. 100–104. <https://doi.org/10.1016/j.jnoncrysol.2017.04.030>
 15. Kan A., Hirabayashi R., Takahashi S., Ogawa H. Low-temperature crystallization and microwave dielectric properties of forsterite generated in MgO-SiO_2 system following LiF addition. *Ceramics International*, 2023, vol. 49, no. 6, pp. 9883–9892. <https://doi.org/10.1016/j.ceramint.2022.11.163>
 16. Yoon S., Otal E.H., Maegli A.E., Karvonen L., Matam S.K., Ebbinghaus S.G., Walfort B., Hagemann H., Pokrant S., Weidenkaff A. Improved persistent luminescence of $\text{CaTiO}_3\text{:Pr}$ by fluorine substitution and thermochemical treatment. *Journal of Alloys and Compounds*, 2014, vol. 613, pp. 338–343. <https://doi.org/10.1016/j.jallcom.2014.06.041>
 17. Yu Y., Wang H., Li L., Chen Y., Zeng R. Effects of various fluxes on the morphology and optical properties of $\text{Lu}_{3-x}\text{Al}_5\text{O}_{12}\text{:xCe}^{3+}$ green phosphors. *Ceramics International*, 2014, vol. 40, no. 9, part A, pp. 14171–14175. <https://doi.org/10.1016/j.ceramint.2014.06.004>
 18. Wang H., Mao F., Liu Y., Jiang X., Ma B., Wei L., Wu F., Li L. Effect of fluxes on luminescence properties of color-tunable $\text{Ba}_{1.3}\text{Ca}_{0.7}\text{SiO}_4\text{:Eu}^{2+},\text{Mn}^{2+}$ phosphor for Near-Ultraviolet white-LEDs. *Materials Research Bulletin*, 2020, vol. 125, pp. 110808. <https://doi.org/10.1016/j.materresbull.2020.110808>
 19. Fujihara S., Koji S., Kimura T. Structure and optical properties of $(\text{Gd,Eu})\text{F}_3$ -nanocrystallized sol-gel silica films. *Journal of Materials Chemistry*, 2004, vol. 14, no. 8, pp. 1331–1335. <https://doi.org/10.1039/b313784h>
 20. Evstropiev S.K., Stolyarova V.L., Saratovskii A.S., Bulyga D.V., Dukelskii K.V., Knyazyan N.B., Yurchenko D.A. Luminescent Mn^{2+} -Doped $\text{MgO-Al}_2\text{O}_3\text{-ZrO}_2\text{-SiO}_2$ Sol-Gel Materials. *Russian Journal of Inorganic Chemistry*, 2024, vol. 69, no. 3, pp. 385–391. <https://doi.org/10.1134/S0036023623603446>
 21. Khaidukov N.M., Brekhovskikh M.N., Kirikova N.Yu., Kondratyuk V.A., Makhov V.N., Luminescence properties of spinels doped with manganese ions. *Russian Journal of Inorganic Chemistry*, 2020, vol. 65, no. 8, pp. 1135–1141. <https://doi.org/10.1134/S0036023620080069>
 22. Song E., Zhou Y., Wei Y., Han X., Tao Z., Qiu R., Xia Z., Zhang Q. A thermally stable narrow-band green-emitting phosphor $\text{MgAl}_2\text{O}_4\text{:Mn}^{2+}$ for wide color gamut backlight display application. *Journal of Materials Chemistry C*, 2019, vol. 7, no. 27, pp. 8192–8198. <https://doi.org/10.1039/c9tc02107h>
 23. El Ghoul J., Omri K., Alyamani A., Barthou C., El Mir L. Synthesis and luminescence of $\text{SiO}_2/\text{Zn}_2\text{SiO}_4$ and $\text{SiO}_2/\text{Zn}_2\text{SiO}_4\text{:Mn}$ composite with sol-gel methods. *Journal of Luminescence*, 2013, vol. 138, pp. 218–222. <https://doi.org/10.1016/j.jlumin.2013.02.009>
 24. Shannon R.D. Revised effective ionic radii and systematic studies of interatomic distances in halides and chalcogenides. *Acta Crystallographica Section A*, 1976, vol. 32, no. 5, pp. 751–767. <https://doi.org/10.1107/s0567739476001551>

25. Fidalgo A., Ilharco L.M., The defect structure of sol-gel-derived silica/polytetrahydrofuran hybrid films by FTIR // *Journal of Non-Crystalline Solids*. 2001. V. 283. N 1-3. P. 144–154. [https://doi.org/10.1016/S0022-3093\(01\)00418-5](https://doi.org/10.1016/S0022-3093(01)00418-5)
26. Sawai S., Uchino T. Visible photoluminescence from MgAl₂O₄ spinel with cation disorder and oxygen vacancy // *Journal of Applied Physics*. 2012. V. 112. N 10. P. 103523. <https://doi.org/10.1063/1.4767228>
27. Dlamini C., Mhlomo M.R., Koao L.F., Motaung T.E., Hlatshwayo T.T., Motloung S.V. The effects of varying the annealing period on the structure, morphology and optical properties of MgAl₂O₄:0.1% Mn²⁺ nanophosphors // *Applied Physics A*. 2020. V. 126. N 1. P. 75. <https://doi.org/10.1007/s00339-019-3248-7>
28. Sandeep K.M., Bhat S., Dharmaprakash S.M. Structural defects and photoluminescence studies of sol-gel prepared ZnO and Al-doped ZnO films // *Applied Physics A*. 2016. V. 122. N 11. P. 975. <https://doi.org/10.1007/s00339-016-0512-y>
29. Wang Y.K., Xie X., Zhu C.G. Self-propagating high-temperature synthesis of magnesium aluminate spinel using Mg-Al alloy // *ACS Omega*. 2022. V. 7. N 15. P. 12617–12623. <https://doi.org/10.1021/acsomega.1c06583>
30. Lin J., Huang Y., Zhang J., Shi F., Wei S., Gao J., Ding X., Tang C. Synthesis and photoluminescence properties of MgAl₂O₄:Mn²⁺ hexagonal nanoplates // *Materials Research Bulletin*. 2009. V. 44. N 1. P. 106–109. <https://doi.org/10.1016/j.materresbull.2008.03.031>
31. Taghavia N., Lerondel G., Makino H., Yamanoto A., Yao T., Kawazoe Y., Golo T., Growth of luminescent Zn₂SiO₄:Mn²⁺ particles inside oxidized porous silicon: emergence of yellow luminescence // *Journal of Crystal Growth*. 2002. V. 237–238. Part 1. P. 869–873. [https://doi.org/10.1016/S0022-0248\(01\)02041-3](https://doi.org/10.1016/S0022-0248(01)02041-3)
32. Jiang Y., Chen J., Xie Z., Zheng L. Syntheses and optical properties of α- and β-Zn₂SiO₄:Mn nanoparticles by solvothermal method in ethylene glycol-water system // *Materials Chemistry and Physics*. 2010. V. 120. N 2–3. P. 313–318. <https://doi.org/10.1016/j.matchemphys.2009.11.002>
25. Fidalgo A., Ilharco L.M., The defect structure of sol-gel-derived silica/polytetrahydrofuran hybrid films by FTIR. *Journal of Non-Crystalline Solids*, 2001, vol. 283, no. 1-3, pp. 144–154. [https://doi.org/10.1016/S0022-3093\(01\)00418-5](https://doi.org/10.1016/S0022-3093(01)00418-5)
26. Sawai S., Uchino T. Visible photoluminescence from MgAl₂O₄ spinel with cation disorder and oxygen vacancy. *Journal of Applied Physics*, 2012, vol. 112, no. 10, pp. 103523. <https://doi.org/10.1063/1.4767228>
27. Dlamini C., Mhlomo M.R., Koao L.F., Motaung T.E., Hlatshwayo T.T., Motloung S.V. The effects of varying the annealing period on the structure, morphology and optical properties of MgAl₂O₄:0.1% Mn²⁺ nanophosphors. *Applied Physics A*, 2020, vol. 126, no. 1, pp. 75. <https://doi.org/10.1007/s00339-019-3248-7>
28. Sandeep K.M., Bhat S., Dharmaprakash S.M. Structural defects and photoluminescence studies of sol-gel prepared ZnO and Al-doped ZnO films. *Applied Physics A*, 2016, vol. 122, no. 11, pp. 975. <https://doi.org/10.1007/s00339-016-0512-y>
29. Wang Y.K., Xie X., Zhu C.G. Self-propagating high-temperature synthesis of magnesium aluminate spinel using Mg-Al alloy. *ACS Omega*, 2022, vol. 7, no. 15, pp. 12617–12623. <https://doi.org/10.1021/acsomega.1c06583>
30. Lin J., Huang Y., Zhang J., Shi F., Wei S., Gao J., Ding X., Tang C. Synthesis and photoluminescence properties of MgAl₂O₄:Mn²⁺ hexagonal nanoplates. *Materials Research Bulletin*, 2009, vol. 44, no. 1, pp. 106–109. <https://doi.org/10.1016/j.materresbull.2008.03.031>
31. Taghavia N., Lerondel G., Makino H., Yamanoto A., Yao T., Kawazoe Y., Golo T., Growth of luminescent Zn₂SiO₄:Mn²⁺ particles inside oxidized porous silicon: emergence of yellow luminescence. *Journal of Crystal Growth*, 2002, vol. 237–238, part 1, pp. 869–873. [https://doi.org/10.1016/S0022-0248\(01\)02041-3](https://doi.org/10.1016/S0022-0248(01)02041-3)
32. Jiang Y., Chen J., Xie Z., Zheng L. Syntheses and optical properties of α- and β-Zn₂SiO₄:Mn nanoparticles by solvothermal method in ethylene glycol-water system. *Materials Chemistry and Physics*, 2010, vol. 120, no. 2–3, pp. 313–318. <https://doi.org/10.1016/j.matchemphys.2009.11.002>

Авторы

Евстропьев Сергей Константинович — доктор химических наук, начальник отдела, АО «НПО ГОИ им. С.И. Вавилова», Санкт-Петербург, 192171, Российская Федерация; ведущий научный сотрудник, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, [sc 6507317768](https://orcid.org/0000-0002-0160-8443), <https://orcid.org/0000-0002-0160-8443>, evstropiev@bk.ru

Столярова Валентина Леонидовна — доктор химических наук, профессор, академик, начальник отдела, Институт химии силикатов им. И. В. Гребенщикова Российской академии наук, Санкт-Петербург, 199034, Российская Федерация; профессор, Санкт-Петербургский государственный университет, Санкт-Петербург, 199034, Российская Федерация, [sc 7005479531](https://orcid.org/0000-0001-7798-4405), <https://orcid.org/0000-0001-7798-4405>, stvl08@inbox.ru

Булыга Дмитрий Владимирович — младший научный сотрудник, АО «НПО ГОИ им. С.И. Вавилова», Санкт-Петербург, 192171, Российская Федерация; инженер-исследователь, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, [sc 57217158694](https://orcid.org/0000-0001-8861-2527), <https://orcid.org/0000-0001-8861-2527>, dmbulyga@yandex.ru

Саратовский Артем Сергеевич — инженер-исследователь, Институт химии силикатов им. И. В. Гребенщикова Российской академии наук, Санкт-Петербург, 199034, Российская Федерация; инженер, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, [sc 57217158694](https://orcid.org/0000-0001-5171-6912), <https://orcid.org/0000-0001-5171-6912>, saratovskija@inbox.ru

Князян Николай Бабкенович — доктор технических наук, профессор, заместитель директора, Институт общей и неорганической химии им. М. Манвеляна НАН РА, Ереван, 0051, Армения, [sc 57217158694](https://orcid.org/0000-0001-9500-0588), <https://orcid.org/0000-0001-9500-0588>, knigo51@mail.ru

Манукян Гоарик Габриеловна — кандидат технических наук, научный сотрудник, Институт общей и неорганической химии им. М. Манвеляна НАН РА, Ереван, 0051, Армения, [sc 57217158694](https://orcid.org/0000-0001-8993-0471), <https://orcid.org/0000-0001-8993-0471>, gmanukyan@sci.am

Authors

Sergey K. Evstropiev — D.Sc. (Chemistry), Head of Department, RPA “Vavilov State Optical Institute”, Saint Petersburg, 192171, Russian Federation; Leading Researcher, ITMO University, Saint Petersburg, 197101, Russian Federation, [sc 6507317768](https://orcid.org/0000-0002-0160-8443), <https://orcid.org/0000-0002-0160-8443>, evstropiev@bk.ru

Valentina L. Stolyarova — D.Sc. (Chemistry), Professor, Academician, Head of Department, Institute of Silicate Chemistry, Russian Academy of Sciences (ISC RAS), Saint Petersburg, 199034, Russian Federation; Professor, St. Petersburg State University (SPbSU), Saint Petersburg, 199034, Russian Federation, [sc 7005479531](https://orcid.org/0000-0001-7798-4405), <https://orcid.org/0000-0001-7798-4405>, stvl08@inbox.ru

Dmitry V. Bulyga — Junior Researcher, RPA “Vavilov State Optical Institute”, Saint Petersburg, 192171, Russian Federation; Engineer-Researcher, ITMO University, Saint Petersburg, 197101, Russian Federation, [sc 57217158694](https://orcid.org/0000-0001-8861-2527), <https://orcid.org/0000-0001-8861-2527>, dmbulyga@yandex.ru

Artem S. Saratovskii — Engineer-Researcher, Institute of Silicate Chemistry, Russian Academy of Sciences (ISC RAS), Saint Petersburg, 199034, Russian Federation; Engineer, ITMO University, Saint Petersburg, 197101, Russian Federation, [sc 57217158694](https://orcid.org/0000-0001-5171-6912), <https://orcid.org/0000-0001-5171-6912>, saratovskija@inbox.ru

Nikolay B. Knyazyan — D.Sc., Professor, Deputy Director, Institute of General and Inorganic Chemistry named after Academician Manvelyan (NAS RA), Erevan, 0051, Armenia, [sc 57217158694](https://orcid.org/0000-0001-9500-0588), <https://orcid.org/0000-0001-9500-0588>, knigo51@mail.ru

Goharik G. Manukyan — PhD, Scientific Researcher, Institute of General and Inorganic Chemistry named after Academician Manvelyan (NAS RA), Erevan, 0051, Armenia, [sc 57217158694](https://orcid.org/0000-0001-8993-0471), <https://orcid.org/0000-0001-8993-0471>, gmanukyan@sci.am

Статья поступила в редакцию 31.10.2024
Одобрена после рецензирования 04.03.2025
Принята к печати 27.05.2025

Received 31.10.2024
Approved after reviewing 04.03.2025
Accepted 27.05.2025



Работа доступна по лицензии
Creative Commons
«Attribution-NonCommercial»

НОВЫЕ МАТЕРИАЛЫ И НАНОТЕХНОЛОГИИ MATERIAL SCIENCE AND NANOTECHNOLOGIES

doi: 10.17586/2226-1494-2025-25-3-396-405

УДК 539.2:541.68

Конформационные свойства полимерных щеток из гребнеобразных макромолекул в условиях сильного растяжения на кубической решетке

Иван Васильевич Лукиев^{1✉}, Иван Викторович Михайлов², Олег Владимирович Борисов³

^{1,2,3} Филиал НИЦ «Курчатовский институт» - ПИЯФ - ИВС, Санкт-Петербург, 199004, Российская Федерация

^{1,3} Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация

¹ ivan.lukiev@mail.ru✉, <https://orcid.org/0000-0003-3022-0496>

² georg_om@mail.ru, <https://orcid.org/0000-0001-5591-1002>

³ oleg.borisov@univ-pau.fr, <https://orcid.org/0000-0002-9281-9093>

Аннотация

Введение. Гребнеобразные полимеры, благодаря своей разветвленной структуре и ряду уникальных физико-химических свойств, используются для модификации различных поверхностей. При достаточно густой прививке макромолекулы образуют однородную полимерную щетку, полностью покрывающую модифицируемую поверхность. Полимерные щетки из гребнеобразных полимеров находят применение в качестве биомедицинских покрытий, смазочных материалов, сенсоров, систем адресной доставки лекарств и многих других. Ввиду широкой востребованности гребнеобразных полимерных покрытий, практически значимым является предсказание их конформационных свойств в зависимости от архитектуры прививаемых полимеров. Щетки из гребнеобразных полимеров достаточно хорошо изучены как теоретически, так и экспериментально при малых плотностях прививки. Однако отсутствуют аналитические модели количественно описывающие свойства этих щеток в условиях больших плотностей прививки при растяжении остовов макромолекул, близких к предельным. **Метод.** Для исследования конформационных свойств плоских полимерных щеток из гребнеобразных полимеров были применены два взаимодополняющих подхода: аналитический и численный методы самосогласованного поля. Первый метод применялся для аналитического описания профиля объемной доли мономерных звеньев привитых макромолекул при их растяжении на объемно-центрированной кубической решетке, второй — для проверки предложенной аналитической модели путем сравнения ее результатов с результатами численных расчетов на простой кубической решетке. В обоих подходах использовалась огрубленная крупнозернистая модель полимерных щеток, в которой минимальными структурными единицами являлись мономерные звенья привитого полимера и молекулы растворителя. **Основные результаты.** Впервые получена универсальная аналитическая формула, описывающая профиль объемной доли мономерных звеньев привитых гребнеобразных макромолекул в широком интервале значений плотности прививки в условиях атермического низкомолекулярного растворителя. Проведена количественная оценка средней толщины полимерных щеток и средней плотности мономерных звеньев при разной эффективной плотности прививки гребнеобразных полимеров, соответствующей отношению действительной плотности прививки к максимально возможной плотности прививки макромолекул с заданной архитектурой, а также при разной разветвленности этих макромолекул. **Обсуждение.** Показано, что в условиях атермического растворителя с ростом степени разветвленности привитых макромолекул при фиксированных плотности прививки и контурной длине основной цепи макромолекул увеличивается средняя толщина полимерной щетки и уменьшается средняя плотность мономерных звеньев. При этом при высоких значениях разветвленности привитых цепей зависимость средней плотности стремится к линейной зависимости от эффективной плотности прививки. Предложенная аналитическая модель растяжения на объемно-центрированной кубической решетке показала высокое согласие с данными, полученными численным моделированием на простой кубической решетке. Результаты исследования позволяют прогнозировать конформационные свойства полимерных щеток в условиях высокой плотности прививки и степени разветвленности привитых гребнеобразных макромолекул.

Ключевые слова

объемно-центрированная кубическая решетка, гребнеобразные макромолекулы, метод самосогласованного поля

© Лукиев И.В., Михайлов И.В., Борисов О.В., 2025

Ссылка для цитирования: Лукиев И.В., Михайлов И.В., Борисов О.В. Конформационные свойства полимерных щеток из гребнеобразных макромолекул в условиях сильного растяжения на кубической решетке // Научно-технический вестник информационных технологий, механики и оптики. 2025. Т. 25, № 3. С. 396–405. doi: 10.17586/2226-1494-2025-25-3-396-405

Conformational properties of polymer brushes with aggrecan-like macromolecules under strong stretching conditions on a cubic lattice

Ivan V. Lukiev^{1✉}, Ivan V. Mikhailov², Oleg V. Borisov³

^{1,2,3} NRC “Kurchatov Institute” - PNPI - IMC, Saint Petersburg, 199004, Russian Federation

^{1,3} ITMO University, Saint Petersburg, 197101, Russian Federation

¹ ivan.lukiev@mail.ru✉, <https://orcid.org/0000-0003-3022-0496>

² georg_om@mail.ru, <https://orcid.org/0000-0001-5591-1002>

³ oleg.borisov@univ-pau.fr, <https://orcid.org/0000-0002-9281-9093>

Abstract

The comb-like polymers are used to modify various surfaces due to their branched structure and a number of unique physical and chemical properties. With a sufficiently dense grafting, the macromolecules form a homogeneous polymer brush that completely covers the surface to be modified. Comb-like polymer brushes find applications as biomedical coatings, lubricants, sensors, targeted drug delivery systems, and many others. Given the wide demand for comb-like polymer coatings, it is of practical importance to predict their conformational properties as a function of the architecture of the grafted polymers. Comb-like polymer brushes have been reasonably well studied both theoretically and experimentally at low grafting densities. However, there are no analytical models that quantitatively describe the properties of these brushes under conditions of high grafting densities and near-limit stretching of the macromolecular backbones. To study the conformational properties of planar polymer brushes made of comb-like polymers, two complementary approaches have been applied: analytical and numerical methods of the self-consistent field. The former was used for analytical description of the volume fraction profile of monomeric units of grafted macromolecules under their stretching on a body-centered cubic lattice, and the latter was used for validation of the proposed analytical model by comparing its results with the numerical calculation data on a simple cubic lattice. A universal analytical formula has been obtained that describes the profile of the volume fraction of monomeric units of grafted comb-like macromolecules in a wide range of grafting density values under conditions of athermal low-molecular-weight solvent. The study proceeded with the quantitative estimation of the average thickness of polymer brushes and the average density of monomeric units at different effective grafting densities of comb-like polymers. This was achieved by determining the ratio of the actual grafting density to the maximum possible grafting density of macromolecules with a given architecture as well as at different branching of these macromolecules. It has been demonstrated that, under conditions of athermal solvent, there is an increase in the average thickness of the polymer brush and a decrease in the average density of monomer units, as the branching degree of grafted macromolecules increases at a fixed grafting density and contour length of the main chain of macromolecules. Furthermore, at elevated levels of branching in grafted chains, the observed dependence of the average density on the effective grafting density approaches a linear relationship. The proposed analytical stretching model on a body-centered cubic lattice showed high agreement with the data obtained by numerical simulation on a simple cubic lattice. The findings of this study provide a foundation for predicting the conformational properties of polymer brushes under conditions of high grafting density and the degree of branching of grafted comb-like macromolecules.

Keywords

body-centered cubic lattice, aggrecan-like macromolecules, self-consistent field method

For citation: Lukiev I.V., Mikhailov I.V., Borisov O.V. Conformational properties of polymer brushes with aggrecan-like macromolecules under strong stretching conditions on a cubic lattice. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2025, vol. 25, no. 3, pp. 396–405 (in Russian). doi: 10.17586/2226-1494-2025-25-3-396-405

Введение

Полимерные щетки состоят из макромолекул, плотно привитых одним концом к непроницаемой поверхности [1, 2]. Модификация поверхностей полимерными щетками используется при производстве сенсоров [3], обеспечении коллоидной стабильности наночастиц [4], улучшении лубрикатных свойств поверхностей [5, 6], а также в биомедицинских приложениях [7]. Плотность прививки полимерных цепей и их молекулярная архитектура — два основных структурных параметра полимерных щеток, регулируя которые можно тонко настраивать лубрикатные, трибологические и защитные свойства модифицированных поверхностей. В качестве

макромолекул-модификаторов могут выступать как линейные полимерные цепи, так и разветвленные.

Частным примером разветвленных цепей являются гребнеобразные полимеры (ГП), состоящие из основной цепи и привитых к ней боковых цепей (рис. 1).

Наиболее распространены регулярные симметричные ГП, в которых боковые цепи располагаются равномерно вдоль остова макромолекулы и имеют одинаковую молекулярную массу и химическую природу. Достижения в синтезе ГП позволили создать функциональные материалы с новыми свойствами, что делает разветвленные полимеры универсальной платформой для молекулярной инженерии в разных областях применения. Примерами таких систем являются: случайные

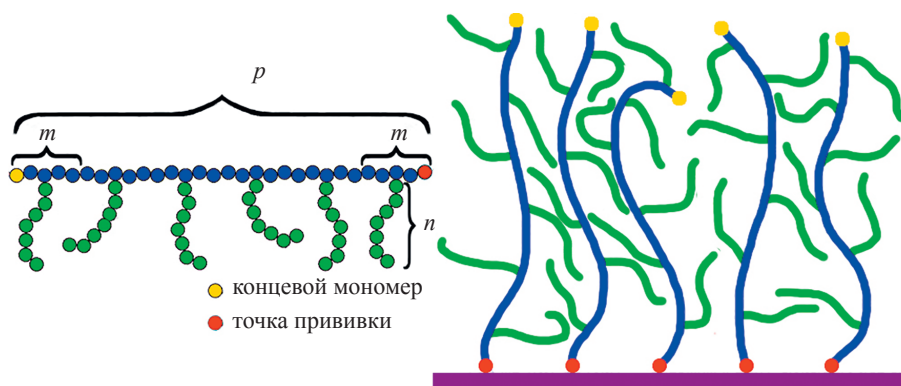


Рис. 1. Схематическое представление плоской полимерной щетки из гребнеобразных макромолекул.

n — длина боковой цепи; m — длина спейсера; p — число повторяющихся разветвленных блоков

Fig. 1. Schematic representation of a planar polymer brush composed of comb-like macromolecules.

n is the side chain length; m is the spacer length; p is the number of repeating branched blocks

гребнеобразные полимерные щетки из поли(этиленгликоля) (ПЭГ)–поли(диметилсилоксана), привитые к кремниевым подложкам с разной морфологией поверхности [8], ПЭГ-метакрилат и 2-гидроксиэтилметакрилат гребнеобразные полимерные щетки на модифицированной 4-винилбензилхлоридом кремниевой подложке [9] и многие другие.

Макромолекулы ГП также встречается в природе в составе различных биологических объектов. Их разнообразие простирается от версиканов, влияющих на пролиферацию и адгезию клеток [10, 11], до агреганов [12–14], обеспечивающих стабилизацию матрикса суставного хряща и его лубрикацию.

Исследованию полимерных щеток посвящено большое число как экспериментальных, так и теоретических работ [15–28]. Исходным объектом внимания научного сообщества являлись плоские щетки из цепей линейного строения. Экспериментально было показано [15–18], что свободные концы привитых цепей распределяются по всему объему, а плотность мономерных звеньев монотонно спадает от поверхности прививки к периферии щетки. В дальнейшем распределения плотности звеньев и свободных концов были количественно описаны в рамках аналитической теории самосогласованного поля (ССП) [18–21], а также показано, что в условиях гауссового растяжения привитых цепей при достаточно низких плотностях прививки профиль химического потенциала щетки может быть описан простой параболической зависимостью. Эта модель была в последующем распространена на щетки из разветвленных цепей [22–26].

Основной недостаток гауссовой модели — возможность бесконечного растяжения полимерной цепи. В ряде работ [27, 28] была сделана попытка построения теории среднего поля для щеток, состоящих из полимерных цепей линейного строения с конечной растяжимостью. В частности, были найдены точные аналитические решения для ряда решеточных моделей [28, 29]. Данные решения позволили устранить разногласия между результатами численных расчетов [30, 31] и аналитическими теориями для гауссовой щетки [19, 21, 32] для простейшей модели в среднем

диапазоне значений плотностей прививки. Несмотря на многочисленные работы по моделированию щеток из ГП (хороший обзор работ представлен в [33]) и технологическую востребованность в предсказании свойств таких щеток, строгое аналитическое описание их конформационных характеристик в научных работах отсутствует. В настоящей работе, опираясь на существующие аналитические модели, показано, как можно модифицировать теорию [28], распространив ее на случай гребнеобразных цепей.

Цель работы — исследование конформационных свойств плоских полимерных щеток, состоящих из привитых ГП, в условиях сильного и даже предельного растяжения макромолекул на объемно-центрированной решетке, а именно, определение равновесной толщины и плотности щетки, как функции ее структурных параметров. Для реализации поставленной цели в рамках метода ССП была построена аналитическая модель, количественно описывающая профиль распределения объемной доли и химического потенциала мономерных звеньев привитых разветвленных полимеров. Для подтверждения разработанной теории методом Схойтенса–Флира на простой кубической решетке проведено независимое численное моделирование полимерных щеток при вариации разветвленности привитых цепей.

Описание модели

Рассматривалась крупнозернистая модель полимерной щетки из гребнеобразных макромолекул, погруженная в низкомолекулярный растворитель на объемно-центрированной и кубической решетках (рис. 1). Предполагалось, что мономерные звенья остова, боковых цепей и молекулы растворителя идентичны и обладают единичным объемом a^3 . Мономерные звенья соединены между собой абсолютно жесткими единичными связями длины a . Полимер и растворитель полагаются несжимаемыми.

Привитые ГП характеризуются следующими структурными параметрами: контурные длины боковых цепей n и спейсеров m , соединяющих соседние по основной цепи точки ветвления; число точек ветвле-

ния в основной цепи p [34]; степень разветвленности $\eta = \sqrt{1 + n/m}$ (топологический коэффициент, который зависит только от топологии привитой молекулы и инвариантен к прочим параметрам. Значение $\eta = 1$ соответствует линейным цепям, $\eta > 1$ — разветвленным); степень полимеризации основной цепи $N_b = pm$; общая степень полимеризации $N = N_b(1 + n/m) = p(m + n) = N_b\eta^2$ [34].

Плотность прививки ГП в щетке характеризуется безразмерной величиной $\sigma = a^2/s$, где s — площадь поверхности прививки, приходящаяся на одну привитую макромолекулу. Максимально возможная плотность прививки [34] ограничивается разветвленностью ГП:

$$\sigma_{\max} = \frac{N_b}{N} = \frac{1}{\eta^2}.$$

Для численного моделирования был выбран следующий набор параметров: $p = 100$, $m = 5$, $n = 0, 15, 40, 75$ (что соответствует $\eta = 1, 2, 3, 4$). Для ГП структурные параметры подбирались так, чтобы длина остова была много больше длины боковых цепей и соседние боковые цепи достаточно сильно перекрывались. Плотность прививки варьировалась в широком диапазоне значений. Растворитель во всех случаях принимался атермическим. В качестве единицы длины использовался линейный размер мономерного звена a , в качестве единицы энергии — $k_B T$, где k_B — константа Больцмана; T — абсолютная температура [35].

Аналитический метод ССП

Аналитический подход ССП предполагает однородность распределения плотности мономерных звеньев полимерных щеток в пределах любой плоскости параллельной плоскости прививки. В любой из точек, равноудаленных от поверхности прививки на расстояние z , объемная доля мономерных звеньев $\varphi(z)$ и химический потенциал $u(z)$ считаются постоянными в пределах одного слоя и являются функцией только одной координаты z , нормальной к поверхности прививки [34]. Данное допущение справедливо для случая достаточно густой прививки цепей, когда соседние цепи сильно перекрываются [34]. Взаимодействие между привитыми цепями приводит к их вытяжке вдоль оси z . Для случая гауссового растяжения цепей (когда растягивающая сила линейно пропорциональна растяжению для каждого участка цепи [32]) было показано [18–26], что профиль химического потенциала полимерной щетки, как для линейных, так и разветвленных привитых цепей имеет параболическую форму:

$$\frac{u(z)}{k_B T} = \frac{3}{2a^2} \left(\frac{\pi\eta}{2N} \right)^2 (H^2 - z^2), \quad (1)$$

где N — степень полимеризации привитой цепи; H — толщина щетки ($u(z \geq H) = 0$); $\eta = \sqrt{1 + n/m}$ — топологический коэффициент. Уравнение (1) применимо к ГП при условии, что число точек ветвления p достаточно велико ($p \rightarrow \infty$) и контурная длина боковых цепей n сопоставима с гауссовым размером спейсеров, соединя-

ющих точки ветвления ($n > \sqrt{m}$). И как было отмечено в разделе «Введение», параболический потенциал не применим при больших плотностях прививки, когда привитые цепи сильно растянуты.

В работе [29] было впервые получено уравнение для потенциала щетки из линейных цепей с конечной растяжимостью в максимально удобной для анализа форме:

$$\int_{z=0}^{z=L_{\max}N} \exp(-\omega V(z)) dz \int_{A=0}^{A=\infty} Z^{-\omega}(f) dA = \frac{1}{\omega}, \quad (2)$$

где L_{\max} — максимальное растяжение привитой цепи ($L_{\max} \leq 1$ для реальных цепей и $L_{\max} = \infty$ для гауссовой цепи); $V(z) = [u(z) + V(H)]/(k_B T)$ — безразмерный потенциал; $Z(f)$ — статистическая сумма свободно-сочлененной цепи под действием растягивающей силы f , индуцированной взаимодействием цепи с внешним эффективным полем, создаваемым соседними цепями; $A = af/(k_B T)$ — эффективная работа, потраченная на ориентацию какого-либо сегмента цепи вдоль вектора растягивающей силы; $\omega \neq 0$ — некоторая вспомогательная величина.

Можно показать, что существует класс точных аналитических решений уравнения (2) для моделей полимерных цепей, в которых зависимость статистической суммы от растягивающей силы имеет следующий вид:

$$Z(A) = \cosh^{\beta}(cA),$$

где c — некоторая константа.

К таким моделям относятся цепи на линейной ($Z(A) = \cosh(A)$), плоской квадратной ($Z(A) = \cosh^2(A/2)$), алмазной ($Z(A) = \cosh^2(A/\sqrt{6})$) и объемно-центрированной кубической ($Z(A) = \cosh^3(A/3)$) (ОЦК) решетках. Потенциал $V(z)$ для этих моделей описывается уравнением:

$$V(z) = \beta \ln \cos \frac{\pi z}{2aNL_{\max}}. \quad (3)$$

Если g — координационное число решетки (число ближайших соседей), то $\beta = \log_2 g(z)$, а $L_{\max} = \sqrt{2/3}$ для алмазной решетки и $L_{\max} = 1$ для всех остальных.

К сожалению, найти столь же простое решение (3) для других моделей, в частности для простой кубической решетки ($Z(A) = (\cosh(A) + 2)/3$), не представляется возможным. Однако имитационное моделирование, как правило, осуществляется именно на простой кубической решетке ввиду простоты реализации алгоритма. Наиболее близким типом решетки к простой кубической является ОЦК решетка.

Профиль химического потенциала $u(z) = V(z) - V$ является основополагающей характеристикой полимерной щетки, зная его можно рассчитать все остальные конформационные характеристики. В частности, профиль объемной доли полимера в условиях атермического растворителя связан с потенциалом следующим образом:

$$\varphi(z) = 1 - \exp \frac{u(z)}{k_B T}. \quad (4)$$

Численный метод Схойтенса–Флира

Численный метод Схойтенса–Флира [36], так же как и аналитический подход, основан на приближении среднего поля, но он свободен от ограничений, налагаемых приближением сильного растяжения. Этот подход использует дискретное решеточное представление пространства. Пространство принимается состоящим из узлов простой кубической решетки с шагом равным линейному размеру мономерного звена a [34]. Узлы решетки организованы в слои, которые расположены параллельно поверхности прививки и пронумерованы как $z = 0, \dots, z_{\max}$. Слой $z = 0$ соответствует поверхности прививки, в слое $z = 1$ фиксируется первый сегмент цепи. Объемная доля ϕ мономерных звеньев и химический потенциал u считаются постоянными в пределах одного слоя и изменяются только в направлении z .

Основная и боковые цепи привитых макромолекул рассматриваются как свободно-сочлененные цепи во внешнем эффективном поле $u(z)$, индуцированном меж- и внутримолекулярными взаимодействиями. Корреляции большего радиуса действия игнорируются. Перекрывание двух сегментов на узле решетки разрешено, но сильно нивелируется условием несжимаемости:

$$\phi(z) + \phi_s(z) = 1, \forall z,$$

где $\phi(z)$ и $\phi_s(z)$ — объемные доли полимера и растворителя.

Свободная энергия F , приходящаяся на одну макромолекулу, в условиях атермического растворителя определяется как отрицательный логарифм статистической суммы Q привитых макромолекул в поле эффективного химического потенциала [34] ($-\ln Q[u]$) за вычетом работы этого поля $\sum_z u(z)\phi(z)$:

$$F = -\ln Q[u(z)] - \frac{1}{\sigma} \sum_{z=0}^{z_{\max}} u(z)\phi(z)\Delta z.$$

Расчет профиля объемной доли $\phi(z)$ и профиля потенциала $u(z)$, соответствующего минимуму свободной энергии при условии несжимаемости, достигается путем оптимизации функционала:

$$F = F\{\phi(z), u(z)\} + \sum_{z=0}^{z_{\max}} \alpha(z)[\phi(z) + \phi_s(z) - 1]\Delta z,$$

где $\alpha(z)$ — поле Лагранжа (набор множителей Лагранжа).

Минимизация функционала F по $\phi(z)$ позволяет вычислить потенциал $u(z)$:

$$\frac{\delta F}{\delta \phi(z)} = 0 \Rightarrow u(z) \leftarrow \alpha(z).$$

Минимизация F по $u(z)$ дает способ вычисления распределения объемной доли $\phi(z)$:

$$\frac{\delta F}{\delta u(z)} = 0 \Rightarrow \phi(z) \leftarrow \frac{\delta(-\ln Q)}{\delta u(z)}.$$

Поле Лагранжа $\alpha(z)$ вычисляется в ходе итеративной процедуры градиентного спуска:

$$\alpha(z) \leftarrow \alpha + h \frac{\delta F}{\delta \alpha(z)} = \alpha(z) + h[\phi_{br}(z) + \phi_{lp}(z) - 1],$$

где $h < 1/2$ — величина шага сходимости.

Общий алгоритм вычислительной процедуры состоит из следующих циклически повторяющихся шагов.

Шаг 0. Инициализируем произвольное поле Лагранжа (например, нулевое):

$$\alpha(z) := 0.$$

Шаг 1. Вычисляем статистический вес Больцмана (вероятность нахождения любого мономерного звена или молекулы растворителя в z -м слое):

$$W_B(z) := \exp\{\alpha(z)\}.$$

Шаг 2. Задаем начальные условия для матриц-пропагаторов (функций Грина), которые определяют вероятности распределения сегментов с номером s по слоям с номером z .

Начальное условие обратного пропагатора для боковых цепей:

$$G_{arm,b}(z, s = n + 1) = W_B(z).$$

Прямой и обратный пропагаторы для основной цепи:

$$G_f(z = 1, s = 1) = W_B(1)G_{arm,b}(z = D, s = N_b) = W_B(z).$$

Шаг 3. «Заполняем» матрицы-пропагаторы по следующим рекуррентным формулам:

$$G_f(s, z) := W_B(z) \left[\frac{1}{6} G_f(s-1, z-1) + \frac{4}{6} G_f(s-1, z) + \frac{1}{6} G_f(s-1, z+1) \right],$$

$$G_b(s, z) := W_B(z) \left[\frac{1}{6} G_f(s+1, z-1) + \frac{4}{6} G_f(s+1, z) + \frac{1}{6} G_f(s+1, z+1) \right].$$

Если номер сегмента s соответствует точке ветвления, то вероятность его нахождения в z -м слое вычисляется как вероятность встречи двух цепей Маркова при «блуждании» по основной цепи и по боковой цепи от свободного конца до точки ветвления:

$$G_f(s, z) := G_f(s, z) \cdot G_{arm,b}(z, s = 1) / W_B(z),$$

$$G_b(s, z) := G_b(s, z) \cdot G_{arm,b}(z, s = 1) / W_B(z).$$

Шаг 4. «Заполняем» пропагаторы для прямого «блуждания» по боковым цепям от точек ветвления к свободным концам:

$$G_{arm,f}(s = 1, z, k) := G_f(s = km + 1, z) \cdot G_b(s = km + 1, z) / W_B(z),$$

$$G_{arm,f}(s \neq 1, z, k) := \frac{W_B(z)}{6} [G_{arm,f}(s-1, z-1, k) + 4G_{arm,f}(s-1, z, k) + G_{arm,f}(s-1, z+1, k)],$$

где k — порядковый номер боковой цепи (считая от привитого конца основной цепи).

Шаг 5. Рассчитываем распределение по слоям парциальной статистической суммы $q(z)$:

$$q(z) = \frac{1}{W_B(z)} = \left[\sum_{s=1}^{s=N_b} G_f(z, s) G_b(z, s) + \sum_{k=1}^{k=N_b/(m-1)} \sum_{s=1}^{s=n} G_{arm,f}(z, s, k) G_{arm,b}(z, s) \right].$$

Шаг 6. Рассчитываем распределение объемной доли:

$$\varphi(z) = \frac{q(z)}{\sum_{z=1}^{z=z_{\max}} q(z)} N\sigma.$$

Шаг 7. Повторно вычисляем поле Лагранжа:

$$\alpha(z) := \alpha(z) + h[\varphi(z) + \varphi_s(z) - 1].$$

Шаг 8. Находим отклонение от условия несжимаемости:

$$\epsilon = \sqrt{\sum_{z=1}^{z=z_{\max}} [\varphi(z) + \varphi_s(z) - 1]^2}.$$

Если отклонение больше указанного значения ($\epsilon > 10^{-8}$), возвращаемся к шагу 1. В противном случае выходим из цикла [34].

Результаты и обсуждение

Для того чтобы построить модель, описывающую профиль объемной доли мономерных звеньев в щетках из гребнеобразных макромолекул, отметим следующий факт: в условиях гауссового растяжения основных цепей ГП ($H/N_b \ll 1$) профиль потенциала $v(z) = u(z) + v$ стремится к форме функции (1), которую также можно переписать в виде:

$$\frac{u(z)}{k_B T} = \frac{-3}{2a^2 l_k} \left(\frac{\pi \eta}{2N} \right)^2 z^2 = \frac{-3}{2a^2 l_k} \left(\frac{\pi}{N_b \eta} \right)^2 z^2 = \frac{-3}{2a^2 l_k} \left(\frac{\pi}{2N_b} \right)^2 z^2.$$

Другими словами, при малых растяжениях цепей, что реализуется при малых плотностях прививки, профиль потенциала $v(z)$ для ГП аналогичен профилю потенциала для щетки из жестких линейных цепей со степенью полимеризации N_b и длиной сегмента Куна равной $l_K = \eta^2$. При этом толщина щетки H , может быть найдена из условия нормировки распределения объемной доли полимера $\varphi(z) = 1 - \exp[-u(z)/(k_B T)]$:

$$\int_{z=0}^{z=H} \varphi(z) dz = N\sigma = N_b \eta^2 \sigma = N_b \sigma_{eff}. \quad (5)$$

Таким образом, исследование щетки из ГП можно заменить изучением щетки из линейных цепей с той

же максимально возможной вытяжкой, но с большей термодинамической жесткостью $l_K = \eta^2$ и большей эффективной плотностью прививки $\sigma_{eff} = \sigma \eta^2 = \sigma / \sigma_{\max}$, которая по сути является отношением действительной плотности прививки к максимально возможной для заданной архитектуры ГП.

Выдвинем гипотезу, что описанное модельное отображение из множества ГП щеток в множество щеток из линейных цепей справедливо не только при малых плотностях прививки, но и для всего интервала значений $\sigma_{eff} \in (0, 1)$. Тогда потенциал ГП щеток также должен описываться уравнением (3) при моделировании цепей на ОЦК решетке и близкой к ней простой кубической решетках.

Разложим уравнение (3) в ряд Маклорена по малому параметру $z/(aN_b)$:

$$V(z) = \beta \ln \cos \frac{\pi z}{2aN_b} = -\beta \left(\frac{\pi^2}{8} \right) \left(\frac{z}{aN_b} \right)^2 - \dots$$

Пренебрегая членами разложения выше второго порядка и приравнявая потенциалы $V(z)$ и $v(z)$, найдем, что значение коэффициента β равно $3/\eta^2$.

С учетом уравнения (4) получим искомое универсальное аналитическое выражение, описывающее профиль объемной доли мономерных звеньев ГП щетки во всей области плотностей прививки (и, соответственно, во всей области растяжения остовов макромолекул):

$$\varphi(z) = 1 - \left(\frac{\cos \frac{\pi H}{2aN_b}}{\cos \frac{\pi z}{2aN_b}} \right)^{\frac{3}{\eta^2}}, \quad (6)$$

где H — толщина щетки, полученная из условия нормировки (5).

Для проверки выдвинутой гипотезы и, соответственно, полученной для $\varphi(z)$ формулы (6), был проведен ряд прямых численных моделирований щеток из ГП на кубической решетке методом Схойтенса–Флира. На рис. 2 представлены профили объемной доли для полимерных щеток из ГП различной разветвленности η при разных эффективных плотностях прививки.

Из графиков (рис. 2) видно, что данные, полученные численным моделированием с большой точностью, согласуются с предложенной аналитической моделью (6). Таким образом, можно утверждать, что гипотеза о применимости представленной модели подтверждается на уровне численного моделирования.

При увеличении эффективной плотности прививки профили объемной доли $\varphi(z)$ естественным образом меняют свою форму с параболической на форму ступеньки, так как в пределе $\sigma_{eff} = 1$ привитый полимерный слой представляет собой «сухую» щетку, полностью заполненную мономерными звеньями и не содержащую молекул растворителя ($\varphi(z) = 1$ при $z < N_b$). Однако при увеличении разветвленности привитых ГП плато $\varphi(z) = \text{const}$ отчетливо наблюдается при все более меньших эффективных плотностях прививки, что может объясняться увеличением наведенной жесткости

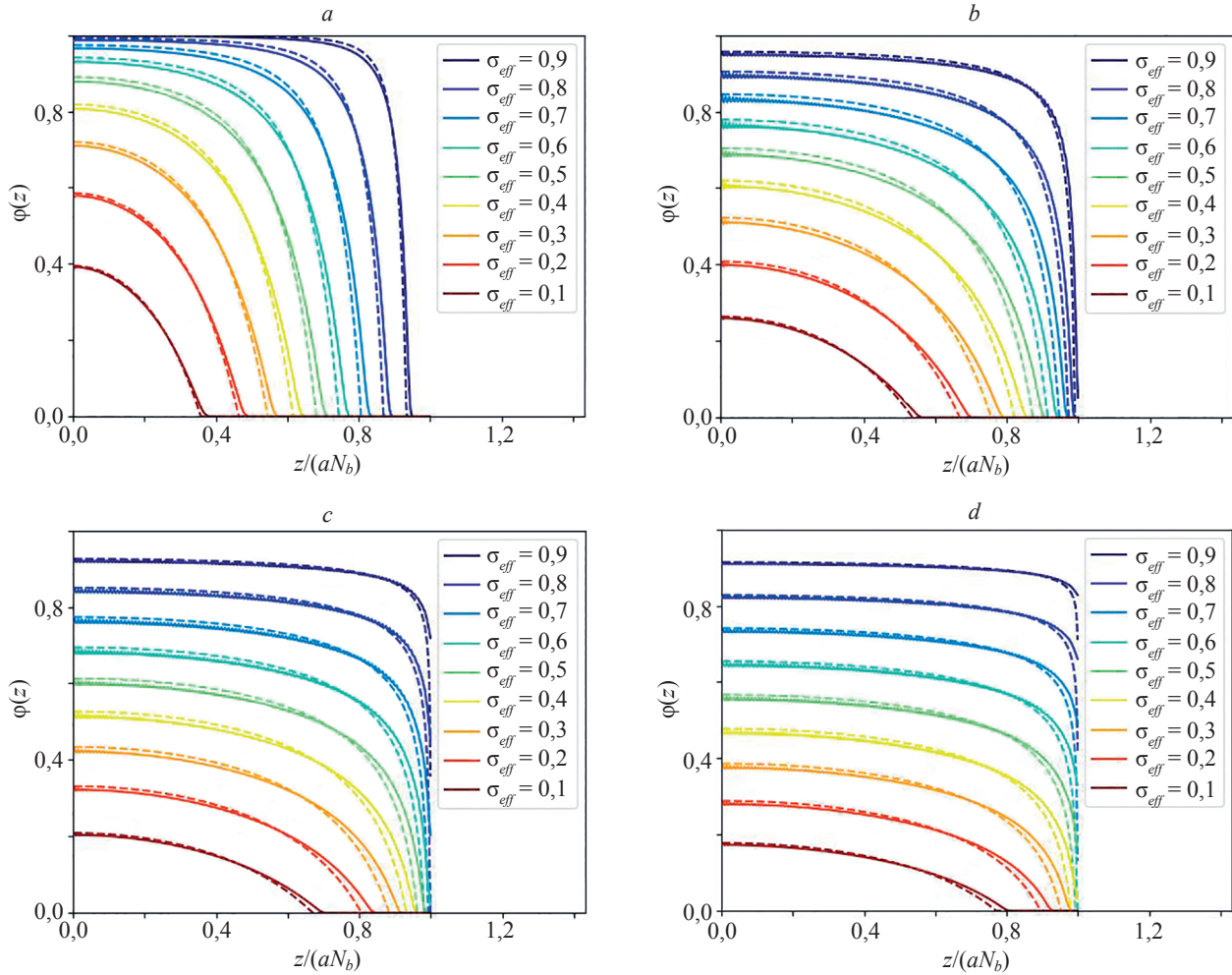


Рис. 2. Профили объемной доли мономерных звеньев щеток из гребнеобразных полимеров для разной степени разветвленности привитых цепей: $\eta = 1$ (a); $\eta = 2$ (b); $\eta = 3$ (c); $\eta = 4$ (d) в приведенных координатах. Сплошными линиями обозначены данные моделирования методом Схойтенса–Флира, пунктирными – соответствующие аналитические зависимости (6)

Fig. 2. Profiles of the volume fraction of monomeric units of brushes with comb-like chains for different branching degree of grafted chains: $\eta = 1$ (a); $\eta = 2$ (b); $\eta = 3$ (c); $\eta = 4$ (d) in reduced coordinates. Solid lines represent simulation data by the Scheutjens–Fleer method, dashed lines correspond to the analytical dependences (6)

основной цепи с возрастанием степени полимеризации боковых цепей.

С практической точки зрения большой интерес представляет возможности теоретического прогнозирования средней толщины щетки:

$$\langle H \rangle = \frac{\int_{z=0}^{z=H} \phi(z) z dz}{\int_{z=0}^{z=H} \phi(z) dz}$$

и средней плотности мономерных звеньев:

$$\langle \phi \rangle = \frac{\int_{z=0}^{z=H} \phi^2(z) dz}{\int_{z=0}^{z=H} \phi(z) dz}.$$

Зависимости данных характеристик от σ_{eff} в сопоставлении с теоретическими кривыми приведены на

рис. 3. Толщина щеток ожидаемо увеличивается, а средняя плотность уменьшается с ростом разветвленности привитых ГП η при одинаковых величинах эффективной плотности прививки и контурной длины основных цепей макромолекул, что особенно выражено в средней области значений σ_{eff} . Отметим, что средняя плотность стремится к линейной зависимости от эффективной плотности прививки ГП с ростом разветвленности и практически ее воспроизводит при $\eta = 4$. Из этого можно заключить, что σ_{eff} является удобным управляющим параметром для определения средней плотности.

Также следует отметить, что представленные в работе зависимости показаны в приведенных координатах ($z/(aN_b)$) на рис. 2 и $\langle H \rangle/(aN_b)$ на рис. 3) и поэтому являются универсальными для разных степеней полимеризации основных цепей N_b (при $N_b \gg n$ и растворитель атермический).

Для того чтобы любой желающий имел возможность самостоятельно воспроизвести данные, представ-

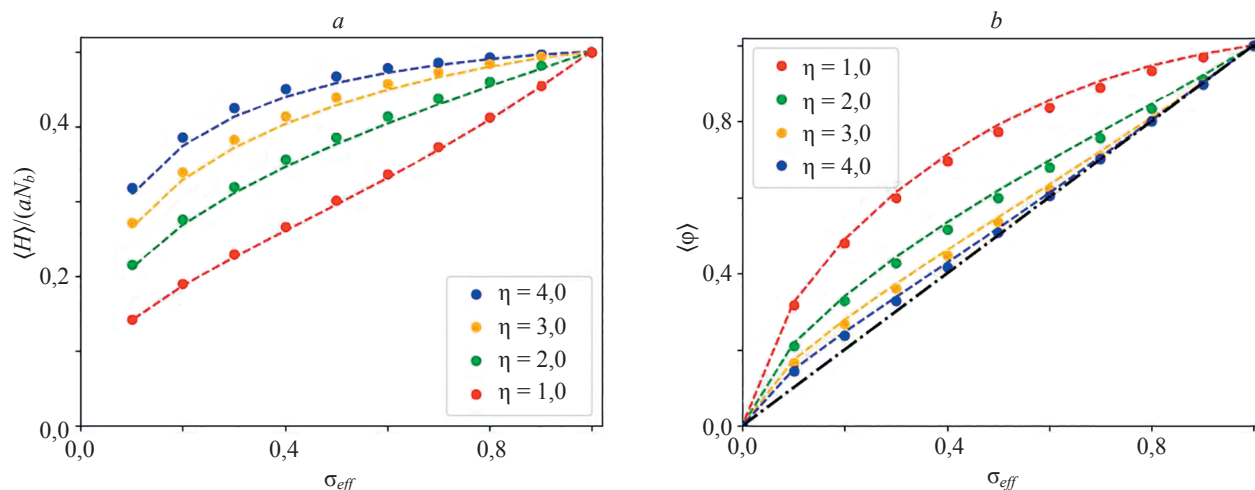


Рис. 3. Зависимости конформационных характеристик полимерных щеток из гребнеобразных полимеров от эффективной плотности прививки σ_{eff} макромолекул: средней толщины щетки в приведенных координатах (а) и средней плотности мономерных звеньев в щетке (б).

Точками обозначены данные моделирования методом Схойтенса-Флира, пунктирными линиями — соответствующие аналитические кривые

Fig. 3. Dependences of conformational characteristics of polymer brushes with comb-like chains on the effective grafting density σ_{eff} of macromolecules: average brush thickness in reduced coordinates (a); average density of monomer units in the brush (b).

Dots represent simulation data by the Scheutjens–Fleer method, dashed lines correspond to analytical curves

ленные в настоящей работе, размещен программный код для расчетов в открытом доступе на языке Python¹.

Заключение

В рамках модели растяжения на объемно-центрированной кубической решетке аналитически выведена универсальная формула профиля объемной доли мономерных звеньев полимерных щеток в широком диапазоне плотности прививки. Проведен анализ зависимости средней толщины щетки и средней плотности мономерных звеньев от степени разветвленности боковых цепей η в указанном диапазоне значений эффективной плотности прививки σ_{eff} .

Показано, что увеличение η при фиксированных σ_{eff} и длине основной цепи макромолекул приводит к увеличению средней толщины щетки. В то же время средняя плотность мономерных звеньев уменьшается. Установлено, что при высоких значениях η средняя плотность вырождается в линейную зависимость от σ_{eff} .

Сопоставление результатов аналитической модели в рамках объемно-центрированной кубической решетки с данными численного моделирования на простой кубической решетке показало их высокую степень согласия. Данный факт подтверждает возможность применения предложенной модели для дальнейшего изучения конформационных свойств полимерных щеток с разветвленными макромолекулами. Стоит также отметить, что все полученные результаты позволяют прогнозировать конформационные свойства полимерных щеток с гребнеобразными макромолекулами в зависимости от густоты прививки и архитектуры привитых цепей, а также управлять ими.

¹ [Электронный ресурс]. Режим доступа: <https://github.com/IvanMikhailovIMCRAS/ascf.git> (дата обращения: 30.12.2024).

Литература

1. Alexander S. Adsorption of chain molecules with a polar head a scaling description // *Journal de Physique*. 1977. V. 38. N 8. P. 983–987. <https://doi.org/10.1051/jphys:01977003808098300>
2. de Gennes P.G. Conformations of polymers attached to an interface // *Macromolecules*. 1980. V. 13. N 5. P. 1069–1075. <https://doi.org/10.1021/ma60077a009>
3. Schüwer N., Klok H.A. A potassium-selective quartz crystal microbalance sensor based on crown-ether functionalized polymer brushes // *Advanced Materials*. 2010. V. 22. N 30. P. 3251–3255. <https://doi.org/10.1002/adma.201000377>
4. Sato T., Ruch R. *Stabilization of Colloidal Dispersions by Polymer Adsorption*. Marcel Dekker Inc., 1980. 176 p.
5. Klein J., Perahia D., Warburg S. Forces between polymer-bearing surfaces undergoing shear // *Nature*. 1991. V. 352. N 6331. P. 143–145. <https://doi.org/10.1038/352143a0>

References

1. Alexander S. Adsorption of chain molecules with a polar head a scaling description. *Journal de Physique*, 1977, vol. 38, no. 8, pp. 983–987. <https://doi.org/10.1051/jphys:01977003808098300>
2. de Gennes P.G. Conformations of polymers attached to an interface. *Macromolecules*, 1980, vol. 13, no. 5, pp. 1069–1075. <https://doi.org/10.1021/ma60077a009>
3. Schüwer N., Klok H.A. A potassium-selective quartz crystal microbalance sensor based on crown-ether functionalized polymer brushes. *Advanced Materials*, 2010, vol. 22, no. 30, pp. 3251–3255. <https://doi.org/10.1002/adma.201000377>
4. Sato T., Ruch R. *Stabilization of Colloidal Dispersions by Polymer Adsorption*. Marcel Dekker Inc., 1980, 176 p.
5. Klein J., Perahia D., Warburg S. Forces between polymer-bearing surfaces undergoing shear. *Nature*, 1991, vol. 352, no. 6331, pp. 143–145. <https://doi.org/10.1038/352143a0>

6. Kreer T. Polymer-brush lubrication: a review of recent theoretical advances // *Soft Matter*. 2016. V. 12. N 15. P. 3479–3501. <https://doi.org/10.1039/C5SM02919H>
7. Ohm C., Welch M.E., Ober C.K. Materials for biosurfaces // *Journal of Materials Chemistry*. 2012. V. 22. N 37. P. 19343–19347. <https://doi.org/10.1039/C2JM90126A>
8. Synytska A., Svetushkina E., Martina D., Bellmann C., Simon F., Ionov L., Stamm M., Creton C. Intelligent materials with adaptive adhesion properties based on comb-like polymer brushes // *Langmuir*. 2012. V. 28. N 47. P. 16444–16454. <https://doi.org/10.1021/la303773b>
9. Zhai G., Cao Y., Gao J. Covalently tethered comb-like polymer brushes on hydrogen-terminated Si (100) surface via consecutive aqueous atom transfer radical polymerization of methacrylates // *Journal of Applied Polymer Science*. 2006. V. 102. N 3. P. 2590–2599. <https://doi.org/10.1002/app.24698>
10. Naso M.F., Zimmermann D.R., Iozzo R.V. Characterization of the complete genomic structure of the human versican gene and functional analysis of its promoter // *Journal of Biological Chemistry*. 1994. V. 269. N 52. P. 32999–33008. [https://doi.org/10.1016/S0021-9258\(20\)30090-9](https://doi.org/10.1016/S0021-9258(20)30090-9)
11. Wu Y.J., La Pierre D.P., Wu J., Yee A.J., Yang B.B. The interaction of versican with its binding partners // *Cell Research*. 2005. V. 15. N 7. P. 483–494. <https://doi.org/10.1038/sj.cr.7290318>
12. Klein J. Molecular mechanisms of synovial joint lubrication // *Proc. of the Institution of Mechanical Engineers, Part J: Journal of Engineering Tribology*. 2006. V. 220. N 8. P. 691–710. <https://doi.org/10.1243/13506501JET143>
13. Chen M., Briscoe W.H., Armes S.P., Klein J. Lubrication at physiological pressures by polyzwitterionic brushes // *Science*. 2009. V. 323. N 5922. P. 1698–1701. <https://doi.org/10.1126/science.1169399>
14. Seror J., Merkher Y., Kampf N., Collinson L., Day A.J., Maroudas A., Klein J. Articular cartilage proteoglycans as boundary lubricants: structure and frictional interaction of surface-attached hyaluronan and hyaluronan–aggrecan complexes // *Biomacromolecules*. 2011. V. 12. N 10. P. 3432–3443. <https://doi.org/10.1021/bm2004912>
15. Бирштейн Т.М., Караев А.В. Конформация макромолекул во взаимодействующих плоских слоях привитых цепей // *Высокомолекулярные соединения. Серия А*. 1987. Т. 29. № 9. С. 1882–1887.
16. Cosgrove T., Heath T., Vanlent B., Leermakers F., Scheutjens J. Configuration of terminally attached chains at the solid/solvent interface: self-consistent field theory and a Monte Carlo model // *Macromolecules*. 1987. V. 20. N 7. P. 1692–1696. <https://doi.org/10.1021/ma00173a041>
17. Сковцов А.М., Павлушков И.В., Горбунов А.А. О структуре монослоя привитых полимерных цепей // *Высокомолекулярные соединения. Серия А*. 1988. Т. 30. № 3. С. 503–508.
18. Сковцов А.М., Павлушков И.В., Горбунов А.А., Жулина Е.Б., Борисов О.В., Прямицын В.А. Структура густо привитых полимерных монослоев // *Высокомолекулярные соединения. Серия А*. 1988. Т. 30. № 8. С. 1615–1622.
19. Жулина Е.Б., Прямицын В.А., Борисов О.В. Структура и конформационные переходы в слоях привитых полимерных цепей. Новая теория // *Высокомолекулярные соединения. Серия А*. 1989. Т. 31. № 1. С. 185–194.
20. Milner S.T., Witten T.A., Cates M.E. A parabolic density profile for grafted polymers // *Europhysics Letters*. 1988. V. 5. N 5. P. 413–418. <https://doi.org/10.1209/0295-5075/5/5/006>
21. Milner S.T., Witten T.A., Cates M.E. Theory of the grafted polymer brush // *Macromolecules*. 1988. V. 21. N 8. P. 2610–2619. <https://doi.org/10.1021/ma00186a051>
22. Pickett G.T. Classical path analysis of end-grafted dendrimers: dendrimer forest // *Macromolecules*. 2001. V. 34. N 25. P. 8784–8791. <https://doi.org/10.1021/ma010917y>
23. Zhulina E.B., Mikhailov I.V., Borisov O.V. Theory of mesophases of triblock comb-shaped copolymers: effects of dead zones and bridging // *Macromolecules*. 2022. V. 55. N 14. P. 6040–6055. <https://doi.org/10.1021/acs.macromol.2c00418>
24. Polotsky A.A., Leermakers F.A.M., Zhulina E.B., Birshtein T.M. On the two-population structure of brushes made of arm-grafted polymer stars // *Macromolecules*. 2012. V. 45. N 17. P. 7260–7273. <https://doi.org/10.1021/ma300691b>
25. Zhulina E.B., Leermakers F.A.M., Borisov O.V. Theory of Brushes Formed by Ψ -Shaped Macromolecules at Solid–Liquid Interfaces // *Langmuir*. 2015. V. 31. N 23. P. 6514–6522. <https://doi.org/10.1021/acs.langmuir.5b00947>
6. Kreer T. Polymer-brush lubrication: a review of recent theoretical advances. *Soft Matter*, 2016, vol. 12, no. 15, pp. 3479–3501. <https://doi.org/10.1039/C5SM02919H>
7. Ohm C., Welch M.E., Ober C.K. Materials for biosurfaces. *Journal of Materials Chemistry*, 2012, vol. 22, no. 37, pp. 19343–19347. <https://doi.org/10.1039/C2JM90126A>
8. Synytska A., Svetushkina E., Martina D., Bellmann C., Simon F., Ionov L., Stamm M., Creton C. Intelligent materials with adaptive adhesion properties based on comb-like polymer brushes. *Langmuir*, 2012, vol. 28, no. 47, pp. 16444–16454. <https://doi.org/10.1021/la303773b>
9. Zhai G., Cao Y., Gao J. Covalently tethered comb-like polymer brushes on hydrogen-terminated Si (100) surface via consecutive aqueous atom transfer radical polymerization of methacrylates. *Journal of Applied Polymer Science*, 2006, vol. 102, no. 3, pp. 2590–2599. <https://doi.org/10.1002/app.24698>
10. Naso M.F., Zimmermann D.R., Iozzo R.V. Characterization of the complete genomic structure of the human versican gene and functional analysis of its promoter. *Journal of Biological Chemistry*, 1994, vol. 269, no. 52, pp. 32999–33008. [https://doi.org/10.1016/S0021-9258\(20\)30090-9](https://doi.org/10.1016/S0021-9258(20)30090-9)
11. Wu Y.J., La Pierre D.P., Wu J., Yee A.J., Yang B.B. The interaction of versican with its binding partners. *Cell Research*, 2005, vol. 15, no. 7, pp. 483–494. <https://doi.org/10.1038/sj.cr.7290318>
12. Klein J. Molecular mechanisms of synovial joint lubrication. *Proc. of the Institution of Mechanical Engineers, Part J: Journal of Engineering Tribology*, 2006, vol. 220, no. 8, pp. 691–710. <https://doi.org/10.1243/13506501JET143>
13. Chen M., Briscoe W.H., Armes S.P., Klein J. Lubrication at physiological pressures by polyzwitterionic brushes. *Science*, 2009, vol. 323, no. 5922, pp. 1698–1701. <https://doi.org/10.1126/science.1169399>
14. Seror J., Merkher Y., Kampf N., Collinson L., Day A.J., Maroudas A., Klein J. Articular cartilage proteoglycans as boundary lubricants: structure and frictional interaction of surface-attached hyaluronan and hyaluronan–aggrecan complexes. *Biomacromolecules*, 2011, vol. 12, no. 10, pp. 3432–3443. <https://doi.org/10.1021/bm2004912>
15. Birshtein T.M., Karaev A.K. Conformation of macromolecules in interacting plane layers of grafted chains. *Vysokomolekulyarnye Soedineniya. Seriya A*, 1987, vol. 29, no. 9, pp. 1882–1887. (in Russian)
16. Cosgrove T., Heath T., Vanlent B., Leermakers F., Scheutjens J. Configuration of terminally attached chains at the solid/solvent interface: self-consistent field theory and a Monte Carlo model. *Macromolecules*, 1987, vol. 20, no. 7, pp. 1692–1696. <https://doi.org/10.1021/ma00173a041>
17. Skvortcov A.M., Pavlushkov I.V., Gorbunov A.A. On the structure of a monolayer of grafted polymer chains. *Vysokomolekulyarnye Soedineniya. Seriya A*, 1988, vol. 30, no. 3, pp. 503–508. (in Russian)
18. Skvortcov A.M., Pavlushkov I.V., Gorbunov A.A., Zhulina E.B., Borisov O.V., Priamitsyn V.A. Structure of dense-grafted polymer monolayers. *Vysokomolekulyarnye Soedineniya. Seriya A*, 1988, vol. 30, no. 8, pp. 1615–1622. (in Russian)
19. Zhulina E.B., Pryamitsyn V.A., Borisov O.V. Structure and conformational transitions in grafted polymer chains layers: new theory. *Vysokomolekulyarnye Soedineniya. Seriya A*, 1989, vol. 31, no. 1, pp. 185–194. (in Russian)
20. Milner S.T., Witten T.A., Cates M.E. A parabolic density profile for grafted polymers. *Europhysics Letters*, 1988, vol. 5, no. 5, pp. 413–418. <https://doi.org/10.1209/0295-5075/5/5/006>
21. Milner S.T., Witten T.A., Cates M.E. Theory of the grafted polymer brush. *Macromolecules*, 1988, vol. 21, no. 8, pp. 2610–2619. <https://doi.org/10.1021/ma00186a051>
22. Pickett G.T. Classical path analysis of end-grafted dendrimers: dendrimer forest. *Macromolecules*, 2001, vol. 34, no. 25, pp. 8784–8791. <https://doi.org/10.1021/ma010917y>
23. Zhulina E.B., Mikhailov I.V., Borisov O.V. Theory of mesophases of triblock comb-shaped copolymers: effects of dead zones and bridging. *Macromolecules*, 2022, vol. 55, no. 14, pp. 6040–6055. <https://doi.org/10.1021/acs.macromol.2c00418>
24. Polotsky A.A., Leermakers F.A.M., Zhulina E.B., Birshtein T.M. On the two-population structure of brushes made of arm-grafted polymer stars. *Macromolecules*, 2012, vol. 45, no. 17, pp. 7260–7273. <https://doi.org/10.1021/ma300691b>
25. Zhulina E.B., Leermakers F.A.M., Borisov O.V. Theory of Brushes Formed by Ψ -Shaped Macromolecules at Solid–Liquid Interfaces.

26. Zhulina E.B., Leermakers F.A.M., Borisov O.V. Ideal mixing in multicomponent brushes of branched polymers // *Macromolecules*. 2015. V. 48. N 21. P. 8025–8035. <https://doi.org/10.1021/acs.macromol.5b01722>
27. Shim D.F.K., Cates M.E. Finite extensibility and density saturation effects in the polymer brush // *Journal de Physique*. 1989. V. 50. N 24. P. 3535–3551. <https://doi.org/10.1051/jphys:0198900500240353500>
28. Amoskov V.M., Pryamitsyn V.A. Theory of monolayers of non-Gaussian polymer chains grafted onto a surface. Part 1. — General theory // *Journal of the Chemical Society, Faraday Transactions*. 1994. V. 90. N 6. P. 889–893. <https://doi.org/10.1039/FT9949000889>
29. Амосков В.М., Прямыцын В.А. Теория привитых полимерных монослоев. Модели цепей с конечной растяжимостью // *Высокомолекулярные соединения. Серия А*. 1995. Т. 37. № 7. С. 1198–1205.
30. Scheutjens J., Fleer G.J. Statistical theory of the adsorption of interacting chain molecules. 1. Partition function, segment density distribution, and adsorption isotherms // *Journal of Physical Chemistry*. 1979. V. 83. N 12. P. 1619–1635. <https://doi.org/10.1021/j100475a012>
31. Scheutjens J., Fleer G.J. Statistical theory of the adsorption of interacting chain molecules. 2. Train, loop, and tail size distribution // *The Journal of Physical Chemistry*. 1980. V. 84. N 2. P. 178–190. <https://doi.org/10.1021/j100439a011>
32. Semenov A.N. Contribution to the theory of microphase layering in block-copolymer melts // *JETP*. 1985. V. 61. N 4. P. 733–742.
33. Leuty G.M., Tsige M., Grest G.S., Rubinstein M. Tension amplification in tethered layers of bottle-brush polymers // *Macromolecules*. 2016. V. 49. N 5. P. 1950–1960. <https://doi.org/10.1021/acs.macromol.5b02305>
34. Викторovich И. Теория изгибной жесткости полимерных щеток из привитых дендронов. Диссертация на соискание ученой степени кандидата физико-математических наук // Федеральное государственное бюджетное учреждение науки институт высокомолекулярных соединений российской академии наук, 2018, [Электронный ресурс]. URL: https://macro.ru/OLD_DOC/council/dis/MikhailovIV/MikhailovIV_dis.pdf (дата обращения: 10.01.2025).
35. Лукиев И.В., Михайлов И.В., Борисов О.В. Влияние качества растворителя на трибологические свойства полимерных щеток // *Научно-технический вестник информационных технологий, механики и оптики*. 2024. Т. 24. № 5. С. 751–757. <https://doi.org/10.17586/2226-1494-2024-24-5-751-757>
36. Fleer G.J., Cohen Stuart M.A., Scheutjens J.M.H.M., Cosgrove T., Vincent B. *Polymers at Interfaces*. Springer Science & Business Media, 1993. 496 p.
- Langmuir*, 2015, vol. 31, no. 23, pp. 6514–6522. <https://doi.org/10.1021/acs.langmuir.5b00947>
26. Zhulina E.B., Leermakers F.A.M., Borisov O.V. Ideal mixing in multicomponent brushes of branched polymers. *Macromolecules*, 2015, vol. 48, no. 21, pp. 8025–8035. <https://doi.org/10.1021/acs.macromol.5b01722>
27. Shim D.F.K., Cates M.E. Finite extensibility and density saturation effects in the polymer brush. *Journal de Physique*, 1989, vol. 50, no. 24, pp. 3535–3551. <https://doi.org/10.1051/jphys:0198900500240353500>
28. Amoskov V.M., Pryamitsyn V.A. Theory of monolayers of non-Gaussian polymer chains grafted onto a surface. Part 1. — General theory. *Journal of the Chemical Society, Faraday Transactions*, 1994, vol. 90, no. 6, pp. 889–893. <https://doi.org/10.1039/FT9949000889>
29. Amoskov V.M., Priamitsyn V.A. Theory of grafted polymer monolayers. Chains models with finite extensibility. *Vysokomolekulyarnye Soedineniya. Seriya A*, 1995, vol. 37, no. 7, pp. 1198–1205. (in Russian)
30. Scheutjens J., Fleer G.J. Statistical theory of the adsorption of interacting chain molecules. 1. Partition function, segment density distribution, and adsorption isotherms. *Journal of Physical Chemistry*, 1979, vol. 83, no. 12, pp. 1619–1635. <https://doi.org/10.1021/j100475a012>
31. Scheutjens J., Fleer G.J. Statistical theory of the adsorption of interacting chain molecules. 2. Train, loop, and tail size distribution. *The Journal of Physical Chemistry*, 1980, vol. 84, no. 2, pp. 178–190. <https://doi.org/10.1021/j100439a011>
32. Semenov A.N. Contribution to the theory of microphase layering in block-copolymer melts. *JETP*, 1985, vol. 61, no. 4, pp. 733–742.
33. Leuty G.M., Tsige M., Grest G.S., Rubinstein M. Tension amplification in tethered layers of bottle-brush polymers. *Macromolecules*, 2016, vol. 49, no. 5, pp. 1950–1960. <https://doi.org/10.1021/acs.macromol.5b02305>
34. Viktorovich I. *Theory of bending stiffness of polymer brushes from graft dendrons*. Dissertation for the degree of candidate of physical and mathematical sciences. 2018. Available at: https://macro.ru/OLD_DOC/council/dis/MikhailovIV/MikhailovIV_dis.pdf (accessed: 10.01.2025). (in Russian)
35. Lukiev I.V., Mikhailov I.V., Borisov O.V. Impact of solvent quality on tribological properties of polymer brushes. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2024, vol. 24, no. 5, pp. 751–757. (in Russian). <https://doi.org/10.17586/2226-1494-2024-24-5-751-757>
36. Fleer G.J., Cohen Stuart M.A., Scheutjens J.M.H.M., Cosgrove T., Vincent B. *Polymers at Interfaces*. Springer Science & Business Media, 1993, 496 p.

Авторы

Лукиев Иван Васильевич — инженер, филиал НИЦ «Курчатовский институт» - ПИЯФ - ИВС, Санкт-Петербург, 199004, Российская Федерация; аспирант, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация; [sc 57386570500](https://orcid.org/0000-0003-3022-0496), <https://orcid.org/0000-0003-3022-0496>, ivan.lukiev@mail.ru

Михайлов Иван Викторович — кандидат физико-математических наук, старший научный сотрудник, филиал НИЦ «Курчатовский институт» - ПИЯФ - ИВС, Санкт-Петербург, 199004, Российская Федерация, [sc 56288088800](https://orcid.org/0000-0001-5591-1002), <https://orcid.org/0000-0001-5591-1002>, georg_om@mail.ru

Борисов Олег Владимирович — доктор физико-математических наук, профессор, заведующий лабораторией, филиал НИЦ «Курчатовский институт» - ПИЯФ - ИВС, Санкт-Петербург, 199004, Российская Федерация; профессор, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, [sc 7005214753](https://orcid.org/0000-0002-9281-9093), <https://orcid.org/0000-0002-9281-9093>, oleg.borisov@univ-pau.fr

Authors

Ivan V. Lukiev — Engineer, NRC “Kurchatov Institute” - PNPI - IMC, Saint Petersburg, 199004, Russian Federation; PhD Student, ITMO University, Saint Petersburg, 197101, Russian Federation, [sc 57386570500](https://orcid.org/0000-0003-3022-0496), <https://orcid.org/0000-0003-3022-0496>, ivan.lukiev@mail.ru

Ivan V. Mikhailov — PhD (Physics & Mathematics), Senior Researcher, NRC “Kurchatov Institute” - PNPI - IMC, Saint Petersburg, 199004, Russian Federation, [sc 56288088800](https://orcid.org/0000-0001-5591-1002), <https://orcid.org/0000-0001-5591-1002>, georg_om@mail.ru

Oleg V. Borisov — D.Sc. (Physics & Mathematics), Professor, Head of Laboratory, NRC “Kurchatov Institute” - PNPI - IMC, Saint Petersburg, 199004, Russian Federation; Professor, ITMO University, Saint Petersburg, 197101, Russian Federation, [sc 7005214753](https://orcid.org/0000-0002-9281-9093), <https://orcid.org/0000-0002-9281-9093>, oleg.borisov@univ-pau.fr

Статья поступила в редакцию 30.01.2025
Одобрена после рецензирования 27.03.2025
Принята к печати 28.05.2025

Received 30.01.2025
Approved after reviewing 27.03.2025
Accepted 28.05.2025



Работа доступна по лицензии
Creative Commons
«Attribution-NonCommercial»

doi: 10.17586/2226-1494-2025-25-3-406-416

УДК 546.07

Генератор синглетного кислорода в газовой фазе атмосферного воздуха для практического многофункционального применения

Лариса Львовна Хомутинникова¹✉, Егор Павлович Быков², Семён Алексеевич Плясцов³,
Сергей Константинович Евстропьев⁴, Игорь Касьянович Мешковский⁵,
Сергей Григорьевич Журавский⁶, Владимир Николаевич Баушев⁷

^{1,2,3,4,5,6,7} Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация

⁶ Национальный медицинский исследовательский центр имени В.А. Алмазова, Санкт-Петербург, 197341, Российская Федерация

¹ larahlesnyh@yandex.ru✉, <https://orcid.org/0000-0003-0918-6350>

² egor1999b@gmail.com, <https://orcid.org/0000-0002-5829-2590>

³ s.plyastsov@gmail.com, <https://orcid.org/0000-0002-5764-0960>

⁴ evstropiev@bk.ru, <https://orcid.org/0000-0002-0160-8443>

⁵ <https://orcid.org/0000-0003-3470-1000>

⁶ s.jour@mail.ru, <https://orcid.org/0000-0002-5960-068X>

⁷ vb@npp17.ru, <https://orcid.org/0009-0004-3116-3563>

Аннотация

Введение. Синглетный кислород относится к метастабильным активным формам кислорода. Синглетный кислород участвует в ряде биохимических реакций и физиологических процессов. Это предполагает возможность его применения для решения прикладных задач медицины и в сферах безопасности жизнедеятельности человека. Благодаря своим окислительным свойствам эффективно уничтожает патогенные организмы, включая бактерии, грибки и вирусы, а также используется в фотодинамической терапии для лечения различных заболеваний, включая онкологические и дерматологические патологии. Для его генерации традиционно применяются фотосенсибилизаторы, которые имеют ряд значительных недостатков, таких как токсичность, низкая селективность по отношению к пораженным клеткам и необходимость использования мощного оптического излучения. Одним из решений этих проблем является применение фотокаталитических материалов, которые генерируют синглетный кислород как в жидкой, так и в газовой средах. В газовой фазе время жизни молекул синглетного кислорода значительно выше, чем в жидкостях. Исследование методов генерации синглетного кислорода в газовой фазе сегодня является актуальной научной задачей. В научной литературе отсутствуют публикации, описывающие качественные и количественные характеристики генерируемых воздушных смесей, обогащенных активными формами кислорода. Разработка генераторов синглетного кислорода в газовой фазе атмосферного воздуха представляет собой актуальную задачу, имеющую много функциональных приложений в областях медицины и технологии безопасности. **Метод.** В настоящей работе представлен и исследован экспериментальный образец прибора для генерации синглетного кислорода в газовой фазе атмосферного воздуха. При его проектировании использованы научные исследования авторов по созданию оригинального фотокаталитического нанокристаллического покрытия на основе $\text{ZnO-SnO}_2\text{-Fe}_2\text{O}_3$, способного в условиях воздействия оптическим излучением, близким к видимому спектру (длина волны 405 нм), к генерации синглетного кислорода. **Основные результаты.** Разработана оригинальная модель прибора. Особенностью устройства является применение фотокатализатора многократного использования. Применяемые материалы исследованы методами рентгенофазового анализа и атомно-силовой микроскопии. Активность генерации синглетного кислорода оценена методом электронного парамагнитного резонанса. Достигнутая скорость фотогенерации синглетного кислорода составила 100 (мкмоль/л)/мин. Расчетная концентрация синглетного кислорода в воздухе на выходе устройства при нормальных условиях в зоне работающего генератора, по величине скорости фотодегградации красителя родамина 6Ж в пористом стекле, составила величину 10 (мкмоль/л)/мин. Представленный прототип прибора отличается низким классом энергопотребления, экологической безопасностью, доступностью материалов, использованием излучения, близкого к видимому спектру, а

© Хомутинникова Л.Л., Быков Е.П., Плясцов С.А., Евстропьев С.К., Мешковский И.К., Журавский С.Г., Баушев В.Н., 2025

также возможностью генерации синглетного кислорода без токсичных окисляющих примесей. **Обсуждение.** Разработанный прототип генератора предполагает возможность создания ряда его модификаций для создания линейки multifunctional приборов: индивидуального, группового применения с лечебной целью, для решения конструкторских задач обеспечения безопасной среды жизнедеятельности и других. Селективная генерация синглетного кислорода позволяет применять приборы для решения медицинских задач как при непосредственном контакте с тканями человека, так и при создании воздушной среды для дыхания и жизнедеятельности человека.

Ключевые слова

генератор синглетного кислорода, оксид цинка, фотокатализ, фотокаталитические наноструктуры

Ссылка для цитирования: Хомутинникова Л.Л., Быков Е.П., Плясцов С.А., Евстропьев С.К., Мешковский И.К., Журавский С.Г., Баушев В.Н. Генератор синглетного кислорода в газовой фазе атмосферного воздуха для практического multifunctional применения // Научно-технический вестник информационных технологий, механики и оптики. 2025. Т. 25, № 3. С. 406–416. doi: 10.17586/2226-1494-2025-25-3-406-416

Atmospheric air-phase singlet oxygen generator for practical multifunctional applications

Larisa L. Khomutinnikova^{1✉}, Egor P. Bykov², Semyon A. Plyastsov³, Sergey K. Evstropiev⁴, Igor K. Meshkovskii⁵, Sergei G. Zhuravskii⁶, Vladimir N. Baushev⁷

^{1,2,3,4,5,6,7} ITMO University, Saint Petersburg, 197101, Russian Federation

⁶ Almazov National Medical Research Centre, Saint Petersburg, 197341, Russian Federation

¹ larahlesnyh@yandex.ru✉, <https://orcid.org/0000-0003-0918-6350>

² egor1999b@gmail.com, <https://orcid.org/0000-0002-5829-2590>

³ s.plyastsov@gmail.com, <https://orcid.org/0000-0002-5764-0960>

⁴ evstropiev@bk.ru, <https://orcid.org/0000-0002-0160-8443>

⁵ <https://orcid.org/0000-0003-3470-1000>

⁶ s.jour@mail.ru, <https://orcid.org/0000-0002-5960-068X>

⁷ vb@npp17.ru, <https://orcid.org/0009-0004-3116-3563>

Abstract

Singlet oxygen is a metastable reactive oxygen species involved in numerous biochemical reactions and physiological processes. This suggests its potential applicability in addressing practical challenges in medicine and human safety. Due to its oxidative properties, singlet oxygen effectively eliminates pathogenic organisms, including bacteria, fungi, and viruses, and is utilized in photodynamic therapy for the treatment of various diseases, including oncological and dermatological pathologies. Traditionally, photosensitizers are employed for its generation; however, they exhibit significant drawbacks, such as toxicity, low selectivity toward affected cells, and the requirement for high-intensity optical radiation. One promising solution involves the use of photocatalytic materials capable of generating singlet oxygen in both liquid and gaseous phases. The lifetime of singlet oxygen molecules in the gas phase is substantially longer than in liquids. Investigating methods for generating singlet oxygen in the gas phase represents a pressing scientific challenge. Currently, there is a lack of publications in scientific literature describing the qualitative and quantitative characteristics of air mixtures enriched with reactive oxygen species. The development of singlet oxygen generators in the gas phase of atmospheric air is an urgent task with multiple functional applications in medicine and safety technologies. This study presents and examines an experimental prototype of a device designed for generating singlet oxygen in the gas phase of atmospheric air. The design incorporates the authors' research on the development of an original photocatalytic nanocrystalline coating based on ZnO-SnO₂-Fe₂O₃, capable of producing singlet oxygen under irradiation with optical radiation near the visible spectrum (405 nm). A novel device model has been developed, featuring a reusable photocatalyst. The materials were characterized using X-ray diffraction analysis and atomic force microscopy. Singlet oxygen generation activity was assessed via electron paramagnetic resonance spectroscopy. The achieved photogeneration rate of singlet oxygen was 100 (μmol/L)/min. The calculated concentration of singlet oxygen in the air at the device outlet under normal conditions, determined based on the photodegradation rate of rhodamine 6G dye in porous glass, reached 10 (μmol/L)/min. The presented prototype exhibits low energy consumption, environmental safety, cost-effective materials, utilization of near-visible spectrum radiation, and the ability to generate singlet oxygen without toxic oxidizing byproducts. The developed prototype allows for the creation of multiple modifications, enabling a range of multifunctional devices for individual or group therapeutic use as well as for engineering solutions aimed at ensuring a safe living environment. Selective singlet oxygen generation permits the application of these devices in medical settings, both for direct tissue contact and for establishing breathable air environments conducive to human life.

Keywords

singlet oxygen generator, zinc oxide, photocatalysis, photocatalytic nanostructures

For citation: Khomutinnikova L.L., Bykov E.P., Plyastsov S.A., Evstropiev S.K., Meshkovskii I.K., Zhuravskii S.G., Baushev V.N. Atmospheric air-phase singlet oxygen generator for practical multifunctional applications. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2025, vol. 25, no. 3, pp. 406–416 (in Russian). doi: 10.17586/2226-1494-2025-25-3-406-416

Введение

Выделяют два практических направления применения синглетного кислорода ($^1\text{O}_2$) в медицинской практике: фотодинамическая терапия (ФДТ) [1] и воздействие газовыми смесями [2–4]. В реферируемых изданиях за последние три десятка лет теме ФДТ посвящено более 2000 публикаций. Несмотря на свои достоинства, метод имеет ряд ограничений, связанных с малым временем жизни активных форм кислорода (АФК), недостаточной селективностью по отношению к клеткам с признаками патологии, необходимостью инъекционного введения фотокатализатора и инвазивного применения оптического излучателя в полости организма [5]. При этом из-за короткого времени жизни молекул АФК [6] терапевтическая интенсивность фотохимических процессов способна развиваться только *in situ* — на поверхности непосредственного контакта светового излучателя с тканью, насыщенной фотосенсибилизатором. Но наиболее важным недостатком ФДТ является токсичность фотосенсибилизаторов, поскольку в их основе лежат органические красители [5]. Решить эти задачи возможно с помощью принципиально иных подходов к генерации $^1\text{O}_2$, предлагающих вместо инвазивной фотогенерации физиологичный способ инкорпорации АФК путем ингаляционного введения.

Примечательно, что в газовой фазе время жизни молекул $^1\text{O}_2$ на несколько порядков превосходит их время жизни в жидкостях [6]. В связи с этим показано, что газообразный $^1\text{O}_2$ способен проявлять свою реагентную способность уже не *in situ*, а на расстоянии нескольких сантиметров от места (поверхности) фотогенерации АФК [7, 8]. Это свидетельствует о значительном потенциале данного направления в медицине. Однако данный способ гораздо менее известен. По теме воздействия газовыми смесями с $^1\text{O}_2$ в рамках настоящей работы было найдено не более 10 публикаций, большая часть которых собрана в [2].

Известно [8], что существенное влияние на эффективность фотогенерации материалами АФК оказывает их химический состав, структура и морфология. По этой причине выбор материала-фотогенератора АФК является важным этапом создания фотокаталитического прибора. При выборе материала-фотогенератора необходимо учитывать спектральные и энергетические параметры источника возбуждающего излучения.

Существует ряд зарубежных приборов-генераторов для ингаляционного способа введения АФК (Airnergy, Vital Air), применяемых в неврологии, кардиологии, реабилитации, спортивной медицине [4, 9]. Для них представлены результаты клинического применения [2]. Однако в представленных работах не рассматриваются аспекты генерации АФК, а также отсутствуют данные доклинических исследований, которые могли бы способствовать более полному пониманию механизма действия и определению клинически значимых показаний.

Целью работы является создание опытного образца прибора, генерирующего из кислорода воздуха $^1\text{O}_2$ с терапевтически-эффективной концентрацией для массового применения в медицине, а также оценка его эффективности.

Выбор фотокаталитического материала

Хорошо известно, что в ФДТ широко используются различные органические красители, обладающие высокой способностью к фотогенерации АФК и вводимые в организм пациента [10]. Низкие стабильность и химическая стойкость органических красителей во многих случаях не ограничивают возможность их эффективного практического применения. Однако для использования в приборах, предназначенных для длительного и многократного использования, не предусматривающего прямого контакта пациента с фотокатализатором, целесообразно применение стабильных и высокоэффективных оксидных фотокаталитических материалов. Целесообразность разработки новых наноструктур для повышения эффективности генерации $^1\text{O}_2$ рассмотрены в [11].

Многочисленными исследованиями было показано, что многие оксидные полупроводники (оксид титана (TiO_2), оксид цинка (ZnO), оксид олова (SnO_2) и другие) обладают высокими фотокаталитическими свойствами и способностью к фотогенерации химически активного кислорода. Введение в широко используемые однокомпонентные фотокаталитические оксиды (TiO_2 , ZnO) модифицирующих добавок других оксидов (SnO_2 , оксид железа (Fe_2O_3) и другие) позволяет сформировать полупроводниковые гетероструктуры [11–13]. Это обеспечивает подавление процессов рекомбинации в полупроводниках фотогенерированных электронно-дырочных пар и существенное усиление их фотокаталитической и антибактериальной активности.

Химический состав, кристаллическая структура и морфология оксидных материалов оказывают сильное влияние на их фотокаталитическую активность и способность к фотогенерации АФК. Процессы фотогенерации АФК протекают на поверхности материалов и повышение их дисперсности сопровождается значительным возрастанием фотокаталитической активности. В связи с этим, целесообразность применения наноразмерных материалов, обладающих большой удельной поверхностью и высокой фотокаталитической активностью, была отмечена в работах [14, 15].

Материалы и методы

В основе предлагаемого экспериментального образца лежит научный задел авторов настоящей работы по разработке оригинального фотокаталитического покрытия, представляющего собой систему $\text{ZnO-SnO}_2\text{-Fe}_2\text{O}_3$ [8, 11]. Кристаллическая структура образцов была исследована рентгенофазовым методом на дифрактометре Rigaku Ultima IV (Rigaku Corporation, Япония). Средний размер кристаллитов в структуре фотокатализатора был рассчитан по формуле Шеррера. Морфология и химический состав материалов изучены методами сканирующей электронной микроскопии с энергодисперсионным анализом на микроскопе TESCAN VEGA3, снабженном приставкой Advanced Aztec Energy (Oxford Instruments, Англия), и атомно-силовой микроскопией (Solver PRO-M (NT-MDT, Россия), полуконтактный режим, зонд HA_NC Etalon (радиус кривизны острия

менее 10 нм)). Для анализа пропускания покрытий использовался спектрофотометр «Lambda 850» с областью длин волн 175–900 нм (сканирующий двухлучевой с двойным монохроматором), источники излучения: в ультрафиолетовом (УФ) диапазоне — дейтериевая лампа, в видимом и ближнем инфракрасном диапазоне — галогенная лампа накаливания.

Для оценки способности прибора генерировать из свободного кислорода атмосферного воздуха $^1\text{O}_2$ был применен метод фотолуминесценции. Измерение спектров флуоресценции в ближней инфракрасной области спектра осуществлялось на спектрометре SDH-IV (SOLAR Laser Systems, Республика Беларусь) при возбуждении люминесценции светодионом HPR40E, излучающим синий свет (максимум длины волны излучения $\lambda_{\text{max}} = 405$ нм) с плотностью излучения 19,5–30,6 мВт/см².

Для количественной оценки генерации и определения эффективности фотокаталитических материалов $^1\text{O}_2$ в работе использовались спиновые ловушки в сочетании с электронным парамагнитным резонансом (ЭПР). Селективность генерации $^1\text{O}_2$ подтверждалась методом, основанным на обесцвечивании водного раствора N,N-dimethyl-p-nitrosoaniline (RNO) с добавлением имидазола (Imd), который подробно описан в [16]. Концентрации RNO (Sigma Aldrich, США) и Imd (Sigma Aldrich, США) в приготовленном растворе составляли $2,5 \cdot 10^{-6}$ моль и $50 \cdot 10^{-3}$ моль соответственно. Кинетика обесцвечивания RNO регистрировалась спектрофотометрическим методом путем измерения оптической плотности раствора на длине волны максимума поглощения 440 нм с течением времени в два этапа: темновой (без УФ-облучения) и под УФ-воздействием. Общая площадь поверхности фотокаталитического покрытия, используемого для обесцвечивания раствора объемом 3 мл, составила 6,5 см². В качестве УФ-источника применялся диод максимумом длины волны излучения $\lambda_{\text{max}} = 365$ нм, спектральные измерения были проведены на спектрофлуориметре «Флюорат-02-Панорама» (ООО «Люмекс», Санкт-Петербург, Россия).

Для количественной оценки фотокаталитической активности предлагаемого прибора использовался метод на основе фотодеградации красителя родамина 6Ж. В качестве матрицы-носителя применены пластины пористого стекла марки ДШ-1М со средним размером пор 7 нм. Последние вымачивали в спиртовом растворе органического красителя в течение 2 ч. Концентрация раствора составила 10^{-3} моль/л. Затем матрицу с красителем высушивали в течение 12 ч. Пористое стекло устанавливалось на выходе и на расстоянии 40 мм от генератора СК. Для оценки концентрации $^1\text{O}_2$ измерялась интенсивность флуоресценции родамина 6Ж на поверхности пористого стекла на длине волны 560 нм при помощи спектрофлуориметра «Флюорат-02-Панорама» (ООО «Люмекс»).

Исследование структурных и оптических свойств покрытия

На рис. 1 представлена рентгенограмма разработанного фотокаталитического покрытия. На рентгенограм-

мах у образцов покрытий наблюдаются интенсивно выраженные пики, соответствующие отражению от плоскостей (100), (002) и (101) в структуре кристаллов ZnO, соответствующие углам Брэгга 31,8°, 34,5° и 36,4°, свидетельствующие о гексагональной структуре вюрцита ZnO. Пик (002) имеет самую высокую интенсивность в покрытиях ZnO, что указывает на преимущественную ориентацию ZnO в направлении (002), перпендикулярном к поверхности подложки. Наличие текстур в покрытиях на основе ZnO уже наблюдалось в работе [17]. Отсутствие каких-либо пиков, связанных с кристаллами SnO₂ и Fe₂O₃, может быть сопряжено с небольшой концентрацией этих компонентов в материале и с тем, что ионы Fe/Sn занимают позиции ионов Zn в решетке ZnO [18].

Расчеты, проведенные на основании данных рентгенофазового анализа, показали, что размеры нанокристаллов ZnO в разработанном покрытии не превышают 30 нм. Введение в состав материалов модифицирующих добавок SnO₂ и Fe₂O₃ способствует уменьшению размеров формирующихся нанокристаллов [13].

На снимках, полученных методом атомно-силовой микроскопии (рис. 2), наблюдается морфология поверхности покрытия, где высота кристаллов ZnO не превышает 35 нм. Это значение согласуется с данными рентгенофазового анализа и подтверждает однородность и мелкодисперсность полученного покрытия. Наличие выраженной кристалличности ZnO указывает на отсутствие значительных агрегаций в процессе синтеза, что положительно сказывается на фотокаталитических свойствах материала за счет создания дополнительных активных центров, что способствует повышению генерации $^1\text{O}_2$ и улучшению взаимодействия с органическими соединениями.

На рис. 3 представлены снимки композитного покрытия, полученные с помощью сканирующей электронной микроскопии. Покрытие состоит из плотно упакованных наночастиц ZnO, размеры и толщина которого не превышают 100 нм [8].

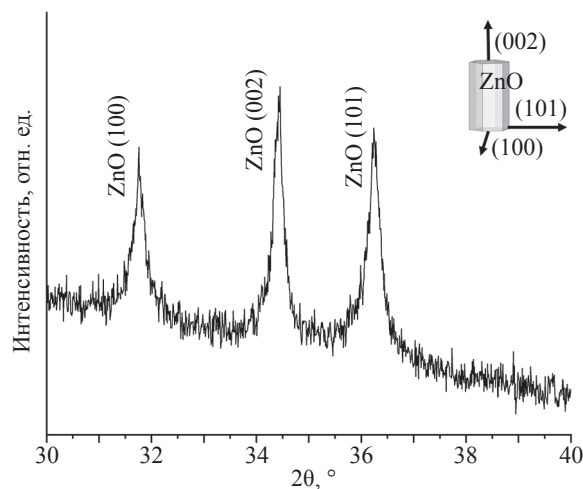


Рис. 1. Рентгеновская дифракционная картина покрытия на основе ZnO на стеклянной подложке

Fig. 1. X-ray diffraction pattern of ZnO-based coating on glass substrate

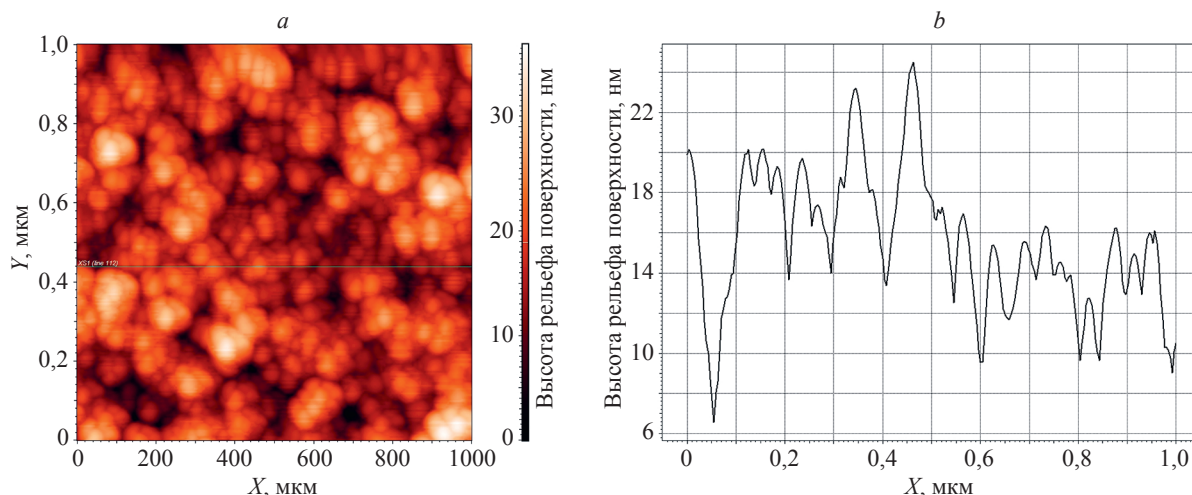


Рис. 2. Атомно-силовое изображение поверхности покрытия на основе ZnO на стеклянной подложке (a) и профилограмма по высоте (b)

Fig. 2. Atomic force microscopy image of the surface of ZnO-based coating on glass substrate (a) and height profile (b)

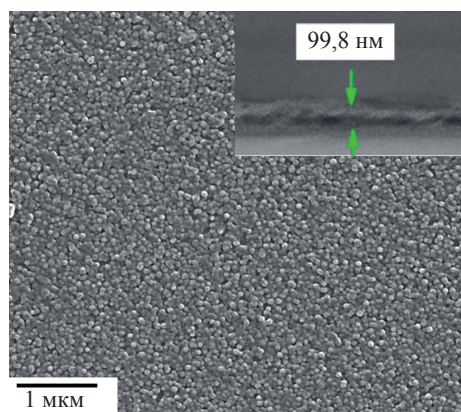


Рис. 3. Электронно-микроскопические снимки покрытия на основе ZnO на стеклянной подложке. На вставке показан торец покрытия на подложке

Fig. 3. Electron microscopic images of ZnO-based coating on glass substrate. The inset shows the cross-sectional view of the coating on the substrate

На рис. 4 продемонстрированы спектры пропускания покрытий. Полученные оксидные покрытия демонстрируют коэффициент пропускания для односто-

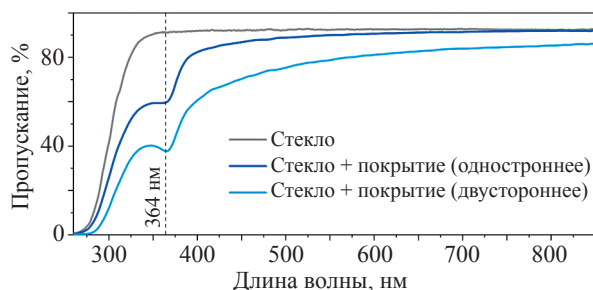


Рис. 4. Спектры поглощения образцов покрытия на основе ZnO на стеклянной подложке

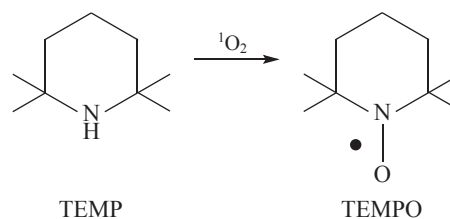
Fig. 4. Absorption spectra of ZnO-based coating samples on glass substrate

ронного (82 %) и двустороннего (60 %) покрытий в видимых областях спектра. Видно, что на длинах волн менее 400 нм наблюдается экситонная полоса поглощения ZnO $\lambda_{\text{max}} = 364$ нм, соответствующая межзонным переходам в кристалле [19]. Дополнительный слой на стекле увеличивает поглощения в УФ-спектральном диапазоне.

Исследование процессов генерации $^1\text{O}_2$ в сочетании с ЭПР

Для количественной оценки генерации и определения эффективности фотокаталитических материалов $^1\text{O}_2$ в работе использовались спиновые ловушки в сочетании с ЭПР [20]. В качестве спиновой ловушки для обнаружения $^1\text{O}_2$ применено органическое соединение 2,2,6,6-тетраметилпиперидин (ТЕМР), избирательно и количественно реагирующий с $^1\text{O}_2$ по реакции с образованием радикала 2,2,6,6-тетраметилпиперидин-1-ил оксил (ТЕМРО), с характерным трехлинейчатым спектром ЭПР (рис. 5).

В качестве исследуемого образца использовалось двустороннее покрытие. На рис. 6 приведены ЭПР-спектры свободного радикала ТЕМРО, измеренные при облучении источником излучения с длиной волны 405 нм ($P = 18,4$ мВт/см²) различной продолжительности в присутствии стеклянной подложки с покрытием. Съемка спектров ЭПР проводилась при частоте



ТЕМР

ТЕМРО

Рис. 5. Реакция образования ТЕМРО

Fig. 5. Reaction scheme for TEMPO formation

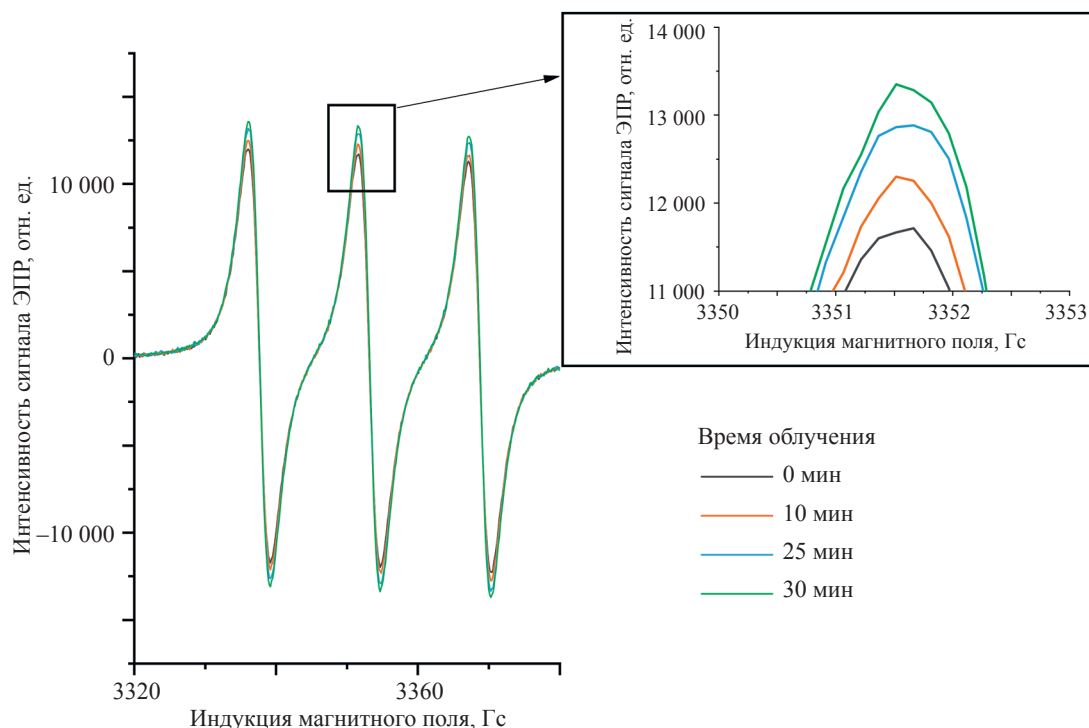


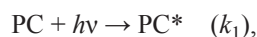
Рис. 6. Спектры с электронным парамагнитным резонансом (ЭПР) свободного радикала TEMPO в присутствии покрытия на основе ZnO в зависимости от продолжительности облучения источником излучения с длиной волны 405 нм ($P = 18,4 \text{ мВт/см}^2$)

Fig. 6. Electron Paramagnetic Resonance spectra of the free TEMPO radical in the presence of ZnO-based coating as a function of irradiation duration by a radiation source with a wavelength of 405 nm ($P = 18.4 \text{ mW/cm}^2$)

100 кГц. Наблюдается, что при УФ-облучении происходит увеличение концентрации нитроксильного радикала, что говорит об образовании $^1\text{O}_2$.

Для количественной оценки образовавшегося кислорода необходимо рассмотреть схему реакции фотокатализа с последующей реакцией TEMPO с $^1\text{O}_2$:

1. Фотовозбуждение фотокатализатора (PC)



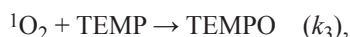
где PC^* — возбужденное состояние фотокатализатора; $h\nu$ — энергия фотона; k_1 — константа скорости фотовозбуждения PC.

2. Перенос энергии на триплетный кислород ($^3\text{O}_2$):



где k_2 — константа скорости переноса энергии.

3. Захват $^1\text{O}_2$ молекулой TEMPO:



где k_3 — константа скорости реакции TEMPO с $^1\text{O}_2$.

Поскольку в данном случае достигается квазистационарный режим по $[\text{PC}^*]$ и $[^1\text{O}_2]$, скорости по этим компонентам равны нулю и их можно записать следующим образом:

$$\frac{d[\text{PC}^*]}{dt} = k_1[\text{PC}] - k_2[\text{PC}^*][^3\text{O}_2] = 0,$$

$$\frac{d[\text{PC}^*]}{dt} = k_2[\text{PC}^*][^3\text{O}_2] - k_3[\text{TEMPO}][^1\text{O}_2] = 0,$$

где квадратные скобки $[\]$ — обозначение концентрации. Откуда для $[^1\text{O}_2]$ получаем:

$$[^1\text{O}_2] = \frac{k_1[\text{PC}]}{k_3[\text{TEMPO}]}.$$

Скорость суммарной реакции равна скорости образования TEMPO:

$$\frac{d[\text{TEMPO}]}{dt} = k_3[\text{TEMPO}][^1\text{O}_2].$$

После подстановки $[^1\text{O}_2]$:

$$\frac{d[\text{TEMPO}]}{dt} = k_1[\text{PC}].$$

Таким образом, для оценки образовавшегося $^1\text{O}_2$ необходима скорость образования TEMPO. Поскольку TEMPO взят в избытке, то скорость образования TEMPO зависит только от скорости образования $^1\text{O}_2$. Наблюдается, что при добавлении фотокаталитического образца происходит накопление нитроксильного радикала. Приведенные зависимости (рис. 7) линейны (коэффициент детерминации $R^2 > 0,9$), что позволяет сделать вывод о том, что кинетика фотогенерации $^1\text{O}_2$ покрытием описывается кинетическим уравнением псевдонулевого порядка. Расчеты показыва-

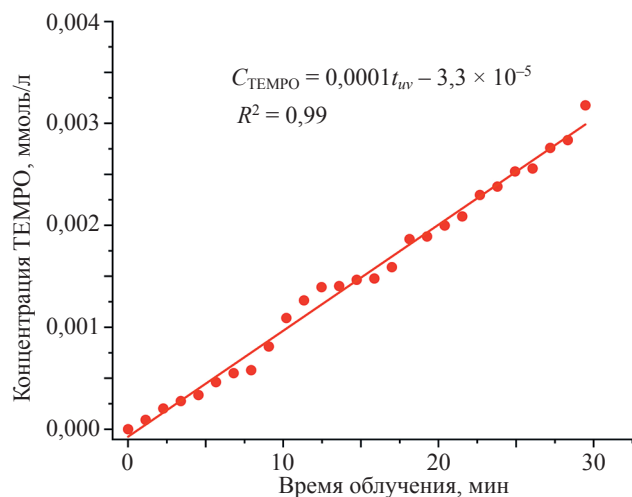


Рис. 7. Кинетика накопления нитроксильного радикала TEMPO в зависимости от времени освещения источника излучения в присутствии покрытия на основе ZnO в зависимости от продолжительности облучения источником излучения с длиной волны 405 нм ($P = 18,4 \text{ мВт/см}^2$).

C_{TEMPO} — концентрация радикала TEMPO; t_{uv} — время облучения

Fig. 7. Kinetics of TEMPO nitroxide radical accumulation as a function of irradiation time by a radiation source in the presence of ZnO-based coating vs. the duration of exposure to a radiation source with a wavelength of 405 nm ($P = 18.4 \text{ mW/cm}^2$).

C_{TEMPO} — concentration of the TEMPO radical; t_{uv} — irradiation time

ли, что константа скорости фотогенерации составляет $k = 100 \text{ (мкмоль/л)/мин}$.

Исследование генерации $^1\text{O}_2$ (фотокаталитических свойств)

Для селективного детектирования $^1\text{O}_2$ применялся метод с использованием RNO и Imd. В процессе облучения фотокатализатора генерируется $^1\text{O}_2$, который взаимодействует с имидазолом, что приводит к образованию промежуточного продукта — трансанулярного пероксида. Окисление имидазола до пероксо-формы вызывает обесцвечивание RNO. При этом такие АФК, как супероксидный и гидроксильный радикалы и перекись водорода, не оказывают значительного влияния на фотодеградацию RNO [16, 21]. По изменению оптической плотности полосы поглощения красителя в области длины волны 440 нм была выполнена оценка скорости продуцирования $^1\text{O}_2$. В качестве растворителя использовалась дистиллированная вода для создания условий, близким к среде в разрабатываемом генераторе.

Для исследования влияния фотокаталитических (ФК) покрытий на кинетику фоторазложения RNO были рассчитаны константы скорости нулевого порядка обесцвечивания красителя (k) по линейной зависимости концентрации (C) от времени (t) на длине волны максимума поглощения из уравнения [21]:

$$C = C_0 - kt,$$

где C_0 — начальная концентрация красителя в растворе.

На рис. 8 представлена зависимость концентрации RNO от времени взаимодействия с УФ-излучением без фотокатализатора (фотолиз) и с ФК-покрытиями без облучения (темновой период) и под УФ-воздействием. В темновой период концентрация оставалась постоянной, что свидетельствует об отсутствии существенного вклада абсорбции красителя на ФК-структурах. Под УФ-воздействием в процессе генерации $^1\text{O}_2$ фотокаталитическими покрытиями $\text{ZnO-SnO}_2\text{-Fe}_2\text{O}_3$ наблюдается значительное увеличение скорости обесцвечивания ($k_{\text{ФК}} = 0,01 \text{ (мкмоль/л)/мин}$) по сравнению с процессом фотолиза ($k_{\text{ФЛ}} = 0,002 \text{ (мкмоль/л)/мин}$), что подтверждает наличие $^1\text{O}_2$ в группе АФК на поверхности ФК.

Описание конструкции и методов измерения концентрации $^1\text{O}_2$ для генератора

На основании полученных выше фундаментальных данных разработана схема прибора для генерации $^1\text{O}_2$ под действием видимого излучения (длина волны 405 нм) из кислорода атмосферного воздуха. Генератор $^1\text{O}_2$ (рис. 9) представляет собой контейнер, содержащий массив пластинок из боросиликатного стекла с нанесенным фотосенсибилизирующим покрытием $\text{ZnO-SnO}_2\text{-Fe}_2\text{O}_3$. В качестве вентилятора был выбран коммерчески доступный вентилятор для системного блока с геометрическими размерами $7 \times 7 \text{ см}$ и питанием 12 В. В качестве источника использован светодиодный источник для фотоотверждения полимерных материалов с длиной волны 405 нм и размерами излучающей поверхности $85 \times 80 \text{ мм}$. Покрытие наносилось на предметные стекла стандартного размера $75 \times 25 \text{ мм}$. Геометрические размеры генератора выбирались таким образом, чтобы все излучение источника попадало на

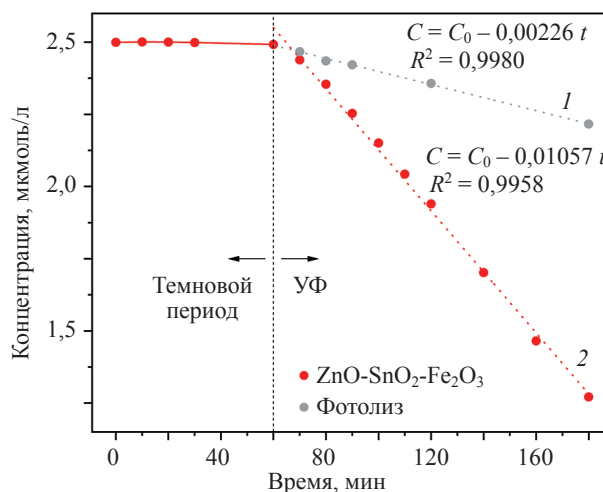
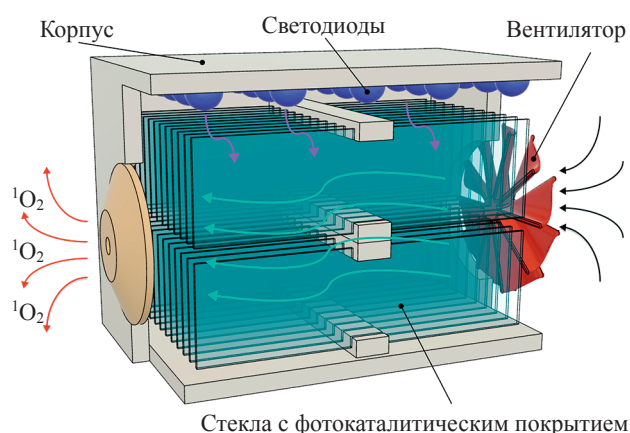


Рис. 8. Фотообесцвечивание RNO в присутствии Imd в процессе фотолиза (кривая 1) и в результате генерации $^1\text{O}_2$ фотокаталитическими покрытиями (кривая 2) в темновой период (без ультрафиолетового облучения) и под ультрафиолетовым воздействием

Fig. 8. Photobleaching of RNO in the presence of Imd during photolysis (curve 1) and as a result of singlet oxygen generation by photocatalytic coatings (curve 2) in the dark period (without ultraviolet irradiation) and under ultraviolet exposure

Рис. 9. 3D-модель разработанного генератора $^1\text{O}_2$ Fig. 9. 3D model of a $^1\text{O}_2$ generator

пластины. Пластины располагались перпендикулярно излучательной поверхности источника. Количество пластин выбиралось из критерия максимального заполнения рабочего объема генератора и расстояния между пластинами 2 мм. Количество пластинок получилось равным 22. Площадь каждой пластины составила 18 см^2 . В качестве источника излучения использовались светодиоды с длиной волны 405 нм и общей потребляемой мощностью 60 Вт. Поток воздуха обеспечивался вентилятором с рабочим потоком воздуха 230 л/мин.

На основании предлагаемой конструктивной модели разработан прототип прибора для изучения влияния генерируемого $^1\text{O}_2$ в экспериментальных условиях на животных (крысах). Прибор, адаптированный для индивидуально-вентилируемой клетки в виварии барьерного типа, представлен на рис. 10. Полученные данные по изучению системных эффектов экзогенно-генериру-

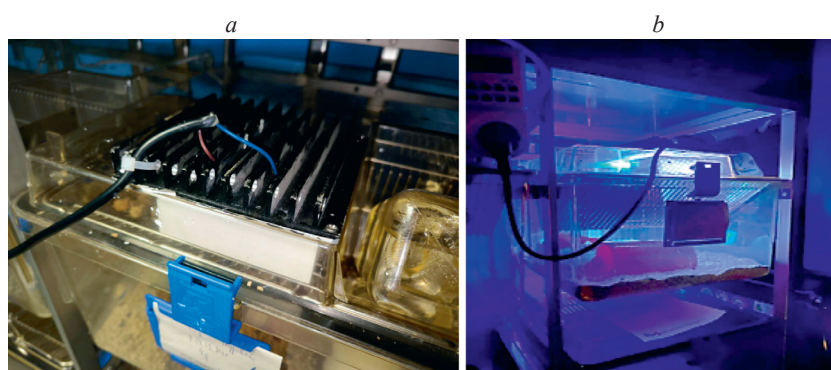
Рис. 10. Вид генератора $^1\text{O}_2$, адаптированного для индивидуально-вентилируемой клетки (ИВК) содержания животных вивария барьерного типа. Прибор, установленный в решетку ИВК для решения задач экспериментальной медицины (a).Генератор $^1\text{O}_2$ в работающем режиме в ИВК содержания крыс (b)

Fig. 10. Appearance of the singlet oxygen generator adapted for IVC animal housing in a barrier-type animal facility. The device installed in the IVC rack for experimental medicine purposes (a). Singlet oxygen generator in operating mode in the IVC housing for rats (b)

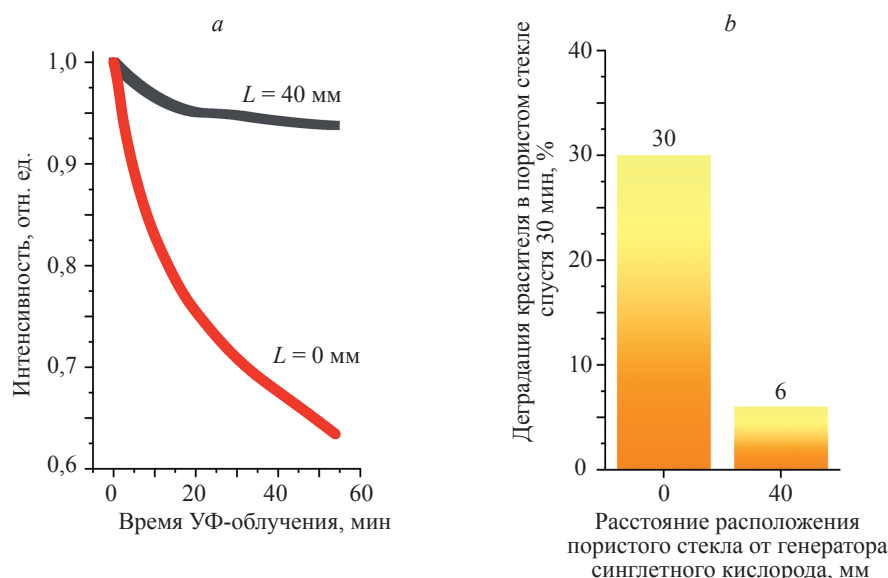
Рис. 11. Зависимость интенсивности флуоресценции для родамина 6Ж в пористом стекле от расстояния (L) и расположения пористого стекла от генератора синглетного кислорода (a); дегградация в процентах красителя в зависимости от расстояния (b)

Fig. 11. Fluorescence intensity of Rhodamine 6G in porous glass versus the distance (L) from the singlet oxygen generator (a), and dye degradation percentage as a function of distance (b)

емого $^1\text{O}_2$ в эксперименте *in vivo* представлены в работе [22]. Показана биологическая безопасность экзогенно-генерируемого $^1\text{O}_2$, попадающего в организм в ходе дыхания с атмосферным воздухом, выявлены первые биологические эффекты длительного воздействия, рассмотрены возможные механизмы их развития и области медицинского применения [22].

Концентрация $^1\text{O}_2$ на выходе генератора, оцененная способом расчета концентрации по фотодеградации органического красителя родамина 6Ж, представлена на рис. 11.

Уровень интенсивности пика флуоресценции через 30 мин экспозиции составил 0,7 от исходного значения. Таким образом, можно предположить, что за 30 мин произошло обесцвечивание 30 % молекул. Величину потока $^1\text{O}_2$ можно оценить как:

$$0,3C_0/t = 10 \text{ (мкмоль/л)/мин,}$$

где C_0 — начальная концентрация красителя в пористом стекле, 10^{-3} моль/л; t — время экспозиции.

Для образца, расположенного на расстоянии 40 мм от выхода генератора, снижение интенсивности флуоресценции наблюдалось только на 6 %, что соответствует потоку молекул $^1\text{O}_2$ с величиной 2 (мкмоль/л)/мин.

Заключение

На основании разработанного фотокаталитического покрытия системы $\text{ZnO-SnO}_2\text{-Fe}_2\text{O}_3$ разработан прототип генератора многофункционального использования для получения синглетного кислорода из кислорода атмосферного воздуха. Отличительными технологическими особенностями конструкции выступают много-разовое фотокаталитическое покрытие, экологическая безопасность материалов, доступность (действующие

поверхности не содержат дорогостоящих металлов платиновой группы), низкий класс энергопотребления, биологическая безопасность, обусловленная использованием излучения видимого диапазона.

Исследование генерации синглетного кислорода с использованием спиновых ловушек и электронного парамагнитного резонанса продемонстрировало, что используемые фотокаталитические материалы эффективно генерируют синглетный кислород при облучении излучением с длиной волны 405 нм. Кинетические характеристики генерации синглетного кислорода соответствовали константе скорости фотогенерации $k = 100$ (мкмоль/л)/мин, в то время как обесцвечивание красителя RNO под действием фотокаталитических покрытий показало заметное увеличение константы скорости до $k_{\text{ФК}} = 0,011$ (мкмоль/л)/мин по сравнению с фотолизом, что подтверждает активное участие синглетного кислорода в процессе деградации.

Конструкция генератора синглетного кислорода, основанная на использовании фотокаталитических покрытий и пористого стекла для детекции, позволила оценить поток синглетного кислорода, который составил 10 (мкмоль/л)/мин на выходе генератора.

Селективность генерации синглетного кислорода без дополнительных агрессивных примесей позволяет использовать генератор для решения задач, связанных с длительным непосредственным контактом с тканями человека, в присутствии человека в закрытых помещениях и безопасно как для потребителя, так и для персонала. Разработанный генератор синглетного кислорода предполагает создание линейки многофункциональных приборов. Это позволяет обсуждать возможности нового применения в медицине по терапевтическим и реабилитационным показаниям, в создании безопасной среды обитания здорового человека.

Литература

- Correia J.H., Rodrigues J.A., Pimenta S., Dong T., Yang Z. Photodynamic Therapy Review: Principles, Photosensitizers, Applications, and Future Directions // *Pharmaceutics*. 2021. V. 13. N 9. P. 1332. <https://doi.org/10.3390/pharmaceutics13091332>
- Мартусевич А.А. Метаболические и гемодинамические эффекты синглетного кислорода: диссертация на соискание ученой степени кандидата биологических наук. Казань, 2019. 167 с.
- Lundberg J., Lindgård A., Elander A., Soussi B. Improved energetic recovery of skeletal muscle in response to ischemia and reperfusion injury followed by in vivo 31P-magnetic resonance spectroscopy // *Microsurgery*. 2002. V. 22. N 4. P. 158–164. <https://doi.org/10.1002/micr.21744>
- Самосюк И.З., Чухраев Н.В., Писанко О.И., Диденко А.А., Бондаренко Л.Н., Курик Л.М. Синглетно-кислородная терапия. Аппараты «МИТ-С» // *Биомедицинская радиоэлектроника*. 2007. № 1. С. 67–75.
- Gunaydin G., Gedik M.E., Ayan S. Photodynamic Therapy—Current Limitations and Novel Approaches // *Frontiers in Chemistry*. 2021. V. 9. P. 691697. <https://doi.org/10.3389/fchem.2021.691697>
- Wang K.-K., Song S., Jung S.-J., Hwang J.-W., Kim M.-G., Kim J.-H., Sung J., Lee J.-K., Kim Y.-R. Lifetime and diffusion distance of singlet oxygen in air under everyday atmospheric conditions // *Physical Chemistry Chemical Physics*. 2020. V. 22. N 38. P. 21664–21671. <https://doi.org/10.1039/d0cp00739k>
- Sun X., Xu K., Chatzidakis A., Norby T. Photocatalytic generation of gas phase reactive oxygen species from adsorbed water: remote action

References

- Correia J.H., Rodrigues J.A., Pimenta S., Dong T., Yang Z. Photodynamic Therapy Review: Principles, Photosensitizers, Applications, and Future Directions. *Pharmaceutics*, 2021, vol. 13, no. 9, pp. 1332. <https://doi.org/10.3390/pharmaceutics13091332>
- Martusevich A.A. *Metabolic and hemodynamic effects of singlet oxygen*. Dissertation for the search of the academic degree of candidate of biological sciences. Kazan, 2019, 167 p. (in Russian)
- Lundberg J., Lindgård A., Elander A., Soussi B. Improved energetic recovery of skeletal muscle in response to ischemia and reperfusion injury followed by in vivo 31P-magnetic resonance spectroscopy. *Microsurgery*, 2002, vol. 22, no. 4, pp. 158–164. <https://doi.org/10.1002/micr.21744>
- Samosiuk I.Z., Chukhraev N.V., Pisanko O.I., Didenko A.A., Bondarenko L.N., Kurik L.M. O.I. Singletno-oxygen therapy. apparatus “MIT — C”. *Biomedical Radioelectronics*, 2007, no. 1, pp. 67–75. (in Russian)
- Gunaydin G., Gedik M.E., Ayan S. Photodynamic Therapy—Current Limitations and Novel Approaches. *Frontiers in Chemistry*, 2021, vol. 9, pp. 691697. <https://doi.org/10.3389/fchem.2021.691697>
- Wang K.-K., Song S., Jung S.-J., Hwang J.-W., Kim M.-G., Kim J.-H., Sung J., Lee J.-K., Kim Y.-R. Lifetime and diffusion distance of singlet oxygen in air under everyday atmospheric conditions. *Physical Chemistry Chemical Physics*, 2020, vol. 22, no. 38, pp. 21664–21671. <https://doi.org/10.1039/d0cp00739k>
- Sun X., Xu K., Chatzidakis A., Norby T. Photocatalytic generation of gas phase reactive oxygen species from adsorbed water: remote action

- and electrochemical detection // *Journal of Environmental Chemical Engineering*. 2021. V. 9. N 2. P. 104809. <https://doi.org/10.1016/j.jece.2020.104809>
8. Khomutinnikova L., Evstropiev S., Meshkovskii I., Bagrov I., Kiselev V. Ceramic ZnO-SnO₂-Fe₂O₃ powders and coatings — effective photogenerators of reactive oxygen species // *Ceramics*. 2023. V. 6. N 2. P. 886–997. <https://doi.org/10.3390/ceramics6020051>
 9. Ronald Dehmlo. Activating and energising water-saturated inhaled air [Электронный ресурс]. URL: [http://www.holmevalleywellbeing.com/files/image/files/Activated Oxygen Therapy/Dr Ronald Dehmlo Study.pdf](http://www.holmevalleywellbeing.com/files/image/files/Activated%20Oxygen%20Therapy/Dr%20Ronald%20Dehmlo%20Study.pdf) (дата обращения: 10.12.2024).
 10. Li B., Lin L., Lin H., Wilson B.C., Photosensitized singlet oxygen generation and detection: Recent advances and future perspectives in cancer photodynamic therapy // *Journal of Biophotonics*. 2016. V. 9. N 11–12. P. 1314–1325. <https://doi.org/10.1002/jbio.201600055>
 11. Гаврилова Д.А., Гаврилова М.А., Хомутичкикова Л.Л., Евстропьев С.К., Мешковский И.К. Оптимизация химического состава и структуры фотокатализаторов системы ZnO-SnO₂-Fe₂O₃ // *Оптика и спектроскопия*. 2024. Т. 132. № 4. С. 413–420. <https://doi.org/10.61011/OS.2024.04.58220.6016-24>
 12. Lachheb H., Ajala F., Hamrouni A., Houas A., Parrino F., Palmisano L. Electron transfer in ZnO-Fe₂O₃ aqueous slurry systems and its effects on visible light photocatalytic activity // *Catalysis Science and Technology*. 2017. V. 7. N 18. P. 4041–4047. <https://doi.org/10.1039/c7cy01085k>
 13. Evstropiev S.K., Karavaeva A.V., Vasilyev V.N., Nikonorov N.V., Aseev V.A., Dukelskii K.V., Lesnykh L.L. Bactericidal properties of ZnO-SnO₂ nanocomposites prepared by polymer-salt method // *Materials Science and Engineering B*. 2021. V. 264. P. 114877. <https://doi.org/10.1016/j.mseb.2020.114877>
 14. Saratovskii A.S., Senchik K.Yu., Karavaeva A.V., Evstropiev S.K., Nikonorov N.V. Photo-oxygenation of water media using photoactive plasmonic nanocomposites // *The Journal of Chemical Physics*. 2022. V. 156. N 20. P. 201103. <https://doi.org/10.1063/5.0094408>
 15. Shelemanov A.A., Nuryev R.K., Evstropiev S.K., Kiselev V.M., Nikonorov N.V. The influence of polyvinylpyrrolidone on the structure and optical properties of ZnO-MgO nanocomposites synthesized by the polymer-salt method // *Optics and Spectroscopy*. 2021. V. 129. N 12. P. 1300–1305. <https://doi.org/10.1134/S0030400X21090198>
 16. Kraljić I., El Mohsni S. A new method for the detection of singlet oxygen in aqueous solutions // *Photochemistry and Photobiology*. 1978. V. 28. N 4–5. P. 577–581. <https://doi.org/10.1111/j.1751-1097.1978.tb06972.x>
 17. Evstropiev S.K., Lesnykh L.L., Nikonorov N.V., Karavaeva A.V., Kolobkova E.V., Oreshkina K.V., Mironov L.Yu., Bagrov I.V. Transparent ZnO-SnO₂ photocatalytic nanocoatings prepared by polymer-salt method // *Optics and Spectroscopy*. 2019. V. 126. N 4. P. 431–438. <https://doi.org/10.1134/S0030400X19040064>
 18. Yi S., Cui J., Li S., Zhang L., Wang D., Lin Y. Enhanced visible-light photocatalytic activity of Fe/ZnO for rhodamine B degradation and its photogenerated charge transfer properties // *Applied Surface Science*. 2014. V. 319. N 1. P. 230–236. <https://doi.org/10.1016/j.apsusc.2014.06.151>
 19. Guo L., Yang S., Yang C., Yu P., Wang J., Ge W., Wong G.K.L. Highly monodisperse polymer-capped ZnO nanoparticles: Preparation and optical properties // *Applied Physics Letters*. 2000. V. 76. N 20. P. 2901–2903. <https://doi.org/10.1063/1.126511>
 20. Khomutinnikova L.L., Evstropiev S.K., Meshkovskii I.K., Moskalenko I.V., Bagrov I.V., Skorb E.V. Intensive singlet oxygen photogeneration and photocatalytic activity of Sn,Fe-doped ZnO-based composites // *Journal of Photochemistry and Photobiology, A: Chemistry*. 2025. V. 462. P. 116254. <https://doi.org/10.1016/j.jphotochem.2024.116254>
 21. Herman J., Neal S.L. Efficiency comparison of the imidazole plus RNO method for singlet oxygen detection in biorelevant solvents // *Analytical and Bioanalytical Chemistry*. 2019. V. 411. N 20. P. 5287–5296 <https://doi.org/10.1007/s00216-019-01910-2>
 22. Журавский С.Г., Мешковский И.К., Подъячева Е.Ю., Шульмейстер Г.А., Зелинская И.А., Плясцов С.А., Хомутичкикова Л.Л., Евстропьев С.К., Баиндурашвили А.Г., Васильев В.Н. Системные эффекты хронической ингаляции воздушной смеси с синглетным кислородом у крыс // *Медицинский академический журнал*. 2025. Т. 25. № 2. С. 41–50.
 - and electrochemical detection. *Journal of Environmental Chemical Engineering*, 2021, vol. 9, no. 2, pp. 104809. <https://doi.org/10.1016/j.jece.2020.104809>
 8. Khomutinnikova L., Evstropiev S., Meshkovskii I., Bagrov I., Kiselev V. Ceramic ZnO-SnO₂-Fe₂O₃ powders and coatings — effective photogenerators of reactive oxygen species. *Ceramics*, 2023, vol. 6, no. 2, pp. 886–997. <https://doi.org/10.3390/ceramics6020051>
 9. Ronald Dehmlo. *Activating and energising water-saturated inhaled air*. Available at: [http://www.holmevalleywellbeing.com/files/image/files/Activated Oxygen Therapy/Dr Ronald Dehmlo Study.pdf](http://www.holmevalleywellbeing.com/files/image/files/Activated%20Oxygen%20Therapy/Dr%20Ronald%20Dehmlo%20Study.pdf) (accessed: 10.12.2024).
 10. Li B., Lin L., Lin H., Wilson B.C., Photosensitized singlet oxygen generation and detection: Recent advances and future perspectives in cancer photodynamic therapy. *Journal of Biophotonics*, 2016, vol. 9, no. 11–12, pp. 1314–1325. <https://doi.org/10.1002/jbio.201600055>
 11. Gavrilova D.A., Gavrilova M.A., Khomutinnikova L.L., Evstropiev S.K., Meshkovskii I.K. Optimization of chemical composition and structure of the photocatalyst. *Optics and Spectroscopy*, 2024, vol. 132, no. 4, pp. 380–386. <https://doi.org/10.61011/EOS.2024.04.58882.6016-24>
 12. Lachheb H., Ajala F., Hamrouni A., Houas A., Parrino F., Palmisano L. Electron transfer in ZnO-Fe₂O₃ aqueous slurry systems and its effects on visible light photocatalytic activity. *Catalysis Science and Technology*, 2017, vol. 7, no. 18, pp. 4041–4047. <https://doi.org/10.1039/c7cy01085k>
 13. Evstropiev S.K., Karavaeva A.V., Vasilyev V.N., Nikonorov N.V., Aseev V.A., Dukelskii K.V., Lesnykh L.L. Bactericidal properties of ZnO-SnO₂ nanocomposites prepared by polymer-salt method. *Materials Science and Engineering B*, 2021, vol. 264, pp. 114877. <https://doi.org/10.1016/j.mseb.2020.114877>
 14. Saratovskii A.S., Senchik K.Yu., Karavaeva A.V., Evstropiev S.K., Nikonorov N.V. Photo-oxygenation of water media using photoactive plasmonic nanocomposites. *The Journal of Chemical Physics*, 2022., vol. 156, no. 20, pp. 201103. <https://doi.org/10.1063/5.0094408>
 15. Shelemanov A.A., Nuryev R.K., Evstropiev S.K., Kiselev V.M., Nikonorov N.V. The influence of polyvinylpyrrolidone on the structure and optical properties of ZnO-MgO nanocomposites synthesized by the polymer-salt method. *Optics and Spectroscopy*, 2021, vol. 129, no. 12, pp. 1300–1305. <https://doi.org/10.1134/S0030400X21090198>
 16. Kraljić I., El Mohsni S. A new method for the detection of singlet oxygen in aqueous solutions. *Photochemistry and Photobiology*, 1978, vol. 28, no. 4–5, pp. 577–581. <https://doi.org/10.1111/j.1751-1097.1978.tb06972.x>
 17. Evstropiev S.K., Lesnykh L.L., Nikonorov N.V., Karavaeva A.V., Kolobkova E.V., Oreshkina K.V., Mironov L.Yu., Bagrov I.V. Transparent ZnO-SnO₂ photocatalytic nanocoatings prepared by polymer-salt method. *Optics and Spectroscopy*, 2019, vol. 126, no. 4, pp. 431–438. <https://doi.org/10.1134/S0030400X19040064>
 18. Yi S., Cui J., Li S., Zhang L., Wang D., Lin Y. Enhanced visible-light photocatalytic activity of Fe/ZnO for rhodamine B degradation and its photogenerated charge transfer properties. *Applied Surface Science*, 2014, vol. 319, no. 1, pp. 230–236. <https://doi.org/10.1016/j.apsusc.2014.06.151>
 19. Guo L., Yang S., Yang C., Yu P., Wang J., Ge W., Wong G.K.L. Highly monodisperse polymer-capped ZnO nanoparticles: Preparation and optical properties. *Applied Physics Letters*, 2000, vol. 76, no. 20, pp. 2901–2903. <https://doi.org/10.1063/1.126511>
 20. Khomutinnikova L.L., Evstropiev S.K., Meshkovskii I.K., Moskalenko I.V., Bagrov I.V., Skorb E.V. Intensive singlet oxygen photogeneration and photocatalytic activity of Sn,Fe-doped ZnO-based composites. *Journal of Photochemistry and Photobiology, A: Chemistry*, 2025, vol. 462, pp. 116254. <https://doi.org/10.1016/j.jphotochem.2024.116254>
 21. Herman J., Neal S.L. Efficiency comparison of the imidazole plus RNO method for singlet oxygen detection in biorelevant solvents. *Analytical and Bioanalytical Chemistry*, 2019, vol. 411, no. 20, pp. 5287–5296 <https://doi.org/10.1007/s00216-019-01910-2>
 22. Zhuravskii S.G., Meshkovskii I.K., Podiacheva E.Iu., Shulmeister G.A., Zelinskaia I.A., Pliastcov S.A., Khomutinnikova L.L., Evstropiev S.K., Baindurashvili A.G., Vasilev V.N. Systemic effects of chronic inhalation of air mixture with singlet oxygen in rats. *Medical Academic Journal*, 2025, vol. 25, no. 2, pp. 41–50. (in Russian)

Авторы

Хомутичкина Лариса Львовна — кандидат технических наук, инженер, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, [sc 57958865600](https://orcid.org/0000-0003-0918-6350), <https://orcid.org/0000-0003-0918-6350>, larahlesnyh@yandex.ru

Быков Егор Павлович — инженер, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, [sc 57376288300](https://orcid.org/0000-0002-5829-2590), <https://orcid.org/0000-0002-5829-2590>, egor1999b@gmail.com

Плясов Семён Алексеевич — кандидат технических наук, заведующий лабораторией, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, [sc 57195587476](https://orcid.org/0000-0002-5764-0960), <https://orcid.org/0000-0002-5764-0960>, s.plyastsov@gmail.com

Евстропьев Сергей Константинович — доктор химических наук, ведущий инженер, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, [sc 6507317768](https://orcid.org/0000-0002-0160-8443), <https://orcid.org/0000-0002-0160-8443>, evstropiev@bk.ru

Мешковский Игорь Касьянович — доктор технических наук, профессор, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, [sc 6603640937](https://orcid.org/0000-0003-3470-1000), <https://orcid.org/0000-0003-3470-1000>,

Журавский Сергей Григорьевич — доктор медицинских наук, ведущий научный сотрудник, Национальный медицинский исследовательский центр имени В.А. Алмазова, Санкт-Петербург, 197341, Российская Федерация; ведущий инженер, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, [sc 8244733500](https://orcid.org/0000-0002-5960-068X), <https://orcid.org/0000-0002-5960-068X>, s.jour@mail.ru

Баушев Владимир Николаевич — доцент, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, <https://orcid.org/0009-0004-3116-3563>, vb@npp17.ru

Authors

Larisa L. Khomutinnikova — PhD, Engineer, ITMO University, Saint Petersburg, 197101, Russian Federation, [sc 57958865600](https://orcid.org/0000-0003-0918-6350), <https://orcid.org/0000-0003-0918-6350>, larahlesnyh@yandex.ru

Egor P. Bykov — Engineer, ITMO University, Saint Petersburg, 197101, Russian Federation, [sc 57376288300](https://orcid.org/0000-0002-5829-2590), <https://orcid.org/0000-0002-5829-2590>, egor1999b@gmail.com

Semyon A. Plyastsov — PhD, Head of Laboratory, ITMO University, Saint Petersburg, 197101, Russian Federation, [sc 57195587476](https://orcid.org/0000-0002-5764-0960), <https://orcid.org/0000-0002-5764-0960>, s.plyastsov@gmail.com

Sergey K. Evstropiev — D.Sc. (Chemistry), Leading Engineer, ITMO University, Saint Petersburg, 197101, Russian Federation, [sc 6507317768](https://orcid.org/0000-0002-0160-8443), <https://orcid.org/0000-0002-0160-8443>, evstropiev@bk.ru

Igor K. Meshkovskii — D.Sc., Professor, ITMO University, Saint Petersburg, 197101, Russian Federation, [sc 6603640937](https://orcid.org/0000-0003-3470-1000), <https://orcid.org/0000-0003-3470-1000>,

Sergei G. Zhuravskii — D.Sc. (Medicine), Leading Researcher, Almazov National Medical Research Centre, Saint Petersburg, 197341, Russian Federation; Leading Engineer, ITMO University, Saint Petersburg, 197101, Russian Federation, [sc 8244733500](https://orcid.org/0000-0002-5960-068X), <https://orcid.org/0000-0002-5960-068X>, s.jour@mail.ru

Vladimir N. Baushev — Associate Professor, ITMO University, Saint Petersburg, 197101, Russian Federation, <https://orcid.org/0009-0004-3116-3563>, vb@npp17.ru

Статья поступила в редакцию 16.01.2025

Одобрена после рецензирования 22.04.2025

Принята к печати 28.05.2025

Received 16.01.2025

Approved after reviewing 22.04.2025

Accepted 28.05.2025



Работа доступна по лицензии
Creative Commons
«Attribution-NonCommercial»

КОМПЬЮТЕРНЫЕ СИСТЕМЫ И ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ COMPUTER SCIENCE

doi: 10.17586/2226-1494-2025-25-3-417-427

УДК 621.396.9, 004.932.4, 550.8.05

Двухэтапный алгоритм восстановления подводных изображений для морских геологоразведочных работ

Иван Владимирович Семерник¹✉, Кристина Владимировна Самонова²

^{1,2} АО «Южморгеология», Геленджик, 353461, Российская Федерация

¹ ione7nick@yandex.ru✉, <https://orcid.org/0000-0002-0238-4154>

² kristi-rosa@yandex.ru, <https://orcid.org/0000-0002-2449-9921>

Аннотация

Введение. Исследованы вопросы восстановления подводных изображений, подверженных искажениям в виде отклонения цветности и контрастности, наличия дымки и других, возникающим в связи с взаимодействием оптического излучения с водной средой. Восстановление подводных изображений является нетривиальной задачей в связи с большой вариативностью параметров водной среды и условий съемки. **Метод.** Предлагаемый метод, в отличие от других алгоритмов восстановления подводных изображений, базирующихся на модели формирования изображения, основан не на упрощенной оценке затухания оптического излучения в воде в виде экспоненциального закона Бера–Ламберта, а на более точном физическом подходе, заключающемся в численном моделировании распространения оптических лучей в воде с использованием метода Монте-Карло, учитывающем основные параметры водной среды и камеры. Результаты моделирования применяются для обработки изображения в пространственной области путем редактирования гистограмм каждого канала изображения в цветовом пространстве RGB. **Основные результаты.** Для тестирования разработанного алгоритма были выбраны 6 реальных подводных изображений, полученных при различных условиях освещения (естественное и искусственное) и различных параметрах водной среды (чистая океанская и мутная прибрежная вода). С целью качественного и количественного анализов полученных результатов использованы следующие аналогичные методы обработки подводных изображений: Fusion, UDCP IATP, Retinex, HE и UWB VCSE. Для количественной оценки полученных результатов использованы показатели Underwater Colour Image Quality Evaluation Metric (UCIQE) и Underwater Image Quality Measure (UIQM). Результаты качественной оценки демонстрируют высокую эффективность предлагаемого метода: вне зависимости от изображения применение метода обеспечивает улучшение визуального восприятия и не приводит к чрезмерному усилению контрастности, искажению цветности, потере детализации, появлению артефактов и пр. Количественная оценка результатов обработки подводных изображений демонстрирует сопоставимые и превосходящие результаты при сравнении с аналогичными методами. Для параметра UCIQE разработанный метод обеспечил улучшение от 9 % до 51 % относительно значения параметра для исходного изображения, при этом аналогичные методы продемонстрировали результаты от минус 10 % до 82 %. Для параметра UIQM разработанный метод обеспечил улучшение от 24 % до 99 % относительно значения параметра для исходного изображения, при этом аналогичные методы продемонстрировали результаты от минус 10 % до 123 %. В отличие от аналогов разработанный метод не продемонстрировал наихудшего значения параметров UCIQE и UIQM ни для одного обработанного изображения, что свидетельствует о стабильности метода вне зависимости от параметров водной среды и условий съемки. Благодаря разделению разработанного метода на предварительный и основной этапы обеспечивается высокая скорость обработки изображений: 0,073 с для изображения с разрешением 400 × 300 пикселей и от 8,02 с до 8,23 с для изображения с разрешением 5184 × 3456 пикселей. Аналогичные методы продемонстрировали значения от 0,19 с до 10,81 с для изображения с разрешением 400×300 пикселей и от 7,65 с до 937,83 с для изображения с разрешением 5184 × 3456 пикселей. **Обсуждение.** Внедрение предлагаемого метода в геологоразведочные работы повысит их эффективность и достоверность, позволит получить более точные данные для дальнейшей разработки месторождений твердых полезных ископаемых. Подобная методика, встроенная в систему машинного зрения подводных аппаратов, позволит существенно расширить их функциональные возможности за счет обеспечения возможности автоматизации операций, повышения эффективности систем распознавания.

Ключевые слова

восстановление подводных изображений, улучшение подводных изображений, обработка подводных изображений, реальные глубоководные фотографии, восстановление цвета, метод Монте-Карло, моделирование распространения света в воде

Благодарности

Исследование выполнено за счет средств гранта Российского научного фонда № 23-79-01253, <https://rscf.ru/project/23-79-01253/>.

Ссылка для цитирования: Семерник И.В., Самонова К.В. Двухэтапный алгоритм восстановления подводных изображений для морских геологоразведочных работ // Научно-технический вестник информационных технологий, механики и оптики. 2025. Т. 25, № 3. С. 417–427. doi: 10.17586/2226-1494-2025-25-3-417-427

Two-stage algorithm for underwater image recovery for marine exploration

Ivan V. Semernik^{1✉}, Christina V. Samonova²

^{1,2} JSC Yuzhmoregeologia, Gelendzhik, 353461, Russian Federation

¹ ione7nick@yandex.ru ✉, <https://orcid.org/0000-0002-0238-4154>

² kristi-rosa@yandex.ru, <https://orcid.org/0000-0002-2449-9921>

Abstract

The paper explores the problems of restoring underwater images exposed to distortions in the form of color and contrast deformations, the presence of haze, etc., arising from the interaction of optical radiation with the aquatic environment. Restoring underwater images is a non-trivial task due to the large variability of the parameters of the aquatic environment and photography conditions. The proposed method, unlike other underwater image recovery algorithms based on an imaging model, is not based on a simplified exponential Beer-Lambert law for estimating optical radiation attenuation in water, but on a more accurate physical approach that simulates the propagation of optical rays in water using the Monte Carlo method, taking into account the main parameters the water environment and the camera. The results of numerical simulation of optical ray propagation in an aquatic environment are used for image processing in the spatial domain by editing the histograms of each image channel in the RGB color space. To test the developed algorithm, 6 real underwater images were selected obtained under various lighting conditions (natural and artificial) and various parameters of the aquatic environment (clear ocean and turbid coastal water). For the purpose of qualitative and quantitative analysis of the obtained results, the following similar underwater image processing methods were used: Fusion, UDCP IATP, Retinex, HE, and UWB VCSE. The Underwater Colour Image Quality Evaluation Metric (UCIQE) and Underwater Image Quality Measure (UIQM) indicators were used to quantify the results obtained. The results of the qualitative assessment demonstrate the high efficiency of the proposed method: regardless of the conditions of the initial image parameters, the application of the developed method improves visual perception and does not lead to excessive contrast enhancement, color distortion, loss of detail, the appearance of artifacts, etc. Quantification of underwater image processing results demonstrates comparable and superior results when comparing the efficiency of the algorithm with similar methods. For the UCIQE parameter, the developed method provided an improvement from 9 % to 51 % relative to the parameter value for the original image, while similar methods demonstrated results from minus 10 % to 82 %. For the UIQM parameter, the developed method provided an improvement from 24 % to 99 % relative to the parameter value for the original image, while similar methods demonstrated results from minus 10 % to 123 %. Unlike analogues, the developed method did not demonstrate the worst values of the UCIQE and UIQM parameters for any processed image, which indicates the stability of the method regardless of the parameters of the aquatic environment and shooting conditions. By dividing the developed method into preliminary and main stages, high image processing speed is ensured: 0.073 seconds for images with a resolution of 400×300 pixels and from 8.02 to 8.23 seconds for images with a resolution of 5184×3456 pixels. Similar methods demonstrated values from 0.19 to 10.81 seconds for an image with a resolution of 400×300 pixels and from 7.65 to 937.83 seconds for an image with a resolution of 5184×3456 pixels. The introduction of the proposed method into the geological exploration will increase their efficiency and reliability, and will provide more accurate data for further exploration of solid mineral deposits. Such technique integrated into the machine vision system of underwater vehicles will significantly expand their functionality by enabling automation of operations and improving the efficiency of recognition systems.

Keywords

underwater image recovery, underwater image enhancement, underwater image processing, true deep-sea pictures, image natural color recovery, Monte Carlo method, propagation of light in water modeling

Acknowledgements

The research was carried out at the expense of the grant of the Russian Science Foundation No. 23-79-01253, <https://rscf.ru/project/23-79-01253/>.

For citation: Semernik I.V., Samonova Ch.V. Two-stage algorithm for underwater image recovery for marine exploration. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2025, vol. 25, no. 3, pp. 417–427 (in Russian). doi: 10.17586/2226-1494-2025-25-3-417-427

Введение

В течение последних нескольких лет наблюдается возрастающий интерес к вопросу подводного машинного зрения и обработки подводных изображений [1, 2]. Одной из важнейших прикладных задач, решение которой требует получения высококачественных результатов фото- и видеосъемки морского дна, является проведение глубоководных геологоразведочных работ [3]. Из-за особенностей водной среды, сложного характера распространения оптического излучения в воде и условий освещения, устранение искажений (восстановление) или улучшение качества (повышение контрастности, яркости, устранение дымки и других) является нетривиальной задачей [4, 5].

Существующие в настоящее время методы обработки подводных изображений можно разделить на следующие основные группы: основанные на применении физических моделей [6–8] (методы восстановления изображений), не основанные на применении физических моделей [1, 9–11] (методы улучшения изображений) и методы на основе машинного обучения нейронных сетей [12].

Так, в работе [9] предложено осуществлять растягивание гистограмм в RGB и HSV цветовых пространствах для улучшения контрастности и насыщенности изображения. В работе [10] представлен подход на основе нормализации гистограмм изображения, направленный на повышение контрастности и балансировку цвета. В [11] представлен метод на основе изменения цветности, контрастности и резкости изображения путем пошагового изменения указанных параметров и оценки количественных параметров для выбора оптимального результата. Указанные методы обеспечивают высокую скорость обработки, но относятся к категории методов улучшения изображений, так как они не основаны на физических моделях и направлены исключительно на повышение визуального восприятия, что может привести к некорректным результатам обработки подводных изображений в части нарушения цветокоррекции.

Основным достоинством методов улучшения качества подводных изображений является возможность повышения визуального восприятия подводного объекта. Существенный недостаток — подобные алгоритмы не опираются на физическую модель распространения света в воде, что приводит к искажению цвета подводных объектов.

Одними из основных подходов, основанных на физических принципах, является метод темного канала (DCP) [6], метод подводного темного канала (UDCP), метод максимальной интенсивности (MIP), метод красного канала (RCP) и другие. Метод DCP позволяет, в первую очередь, устранить дымку на подводных изображениях и основан на допущении, что подводные изображения хорошего качества имеют некоторое количество пикселей с очень малой интенсивностью как минимум в канале одного цвета. Методы на основе DCP и MIP могут приводить к неверным результатам в случае, если условия съемки и освещения не соответствуют основным допущениям, лежащим в основе этих

методов [13]. Применение искусственного освещения приводит к наличию более светлых объектов на переднем плане и более темных объектов на заднем плане изображения, что отличается от результатов съемки при естественном освещении (наличие светлого фона на изображении) и приводит к неверной оценке глубины сцены методами на основе DCP и MIP [13].

Вторым основным направлением в числе методов, основанных на физических принципах, является применение модели формирования изображения (IFM) и оптических свойств водной среды [12]. В работе [7] предложен метод на основе априорной информации с элементами методов обработки в пространственной области и корректировкой контрастности путем нормализации гистограммы. В [8] представлен итерационный метод на основе оптической модели IFM с учетом условий освещенности.

Основное преимущество методов, основанных на применении физических моделей: восстановление естественных цветов подводных объектов. Недостатки: относительная сложность реализации, более высокая длительность обработки изображений по сравнению с методами улучшения изображений, высокая вероятность некорректной работы при использовании искусственного освещения при проведении съемки.

Методы, основанные на применении глубокого обучения, обладают неоспоримыми преимуществами (простота применения, высокое быстродействие, улучшение визуального восприятия), но существенным недостатком нейронных сетей является их обучение на парах изображений (с добавлением дымки, сдвига цвета в зеленую или синюю область спектра и другие). При этом зачастую подобные методы после обучения демонстрируют худшие результаты при обработке реальных подводных изображений [14].

Таким образом, задача реализации относительно простого в применении и эффективного метода обработки изображений, основанного на физической модели распространения света в воде и обеспечивающего восстановление естественных цветов подводных объектов в том числе и при использовании искусственного освещения (что является неизбежным при осуществлении глубоководной съемки), является актуальной.

Модель формирования изображения

В большинстве алгоритмов обработки подводных изображений используется модель формирования изображения, предложенная в [15] и показанная на рис. 1:

$$I_c(x) = t_c(x)J_c(x) + (1 - t_c(x))A_c,$$

где $c \in \{R, G, B\}$ — индекс цветового канала; x — координаты пиксела; I_c — изображение объекта в цветовом канале c , фиксируемое камерой; t_c — коэффициент передачи для цветового канала c ; J_c — излучение от объекта, которое необходимо зафиксировать на изображении; A_c — фоновая освещенность части изображения, на которой отсутствуют объекты. Наличие фоновой освещенности, а также рассеяние прямого и фонового света приводит к снижению контрастности

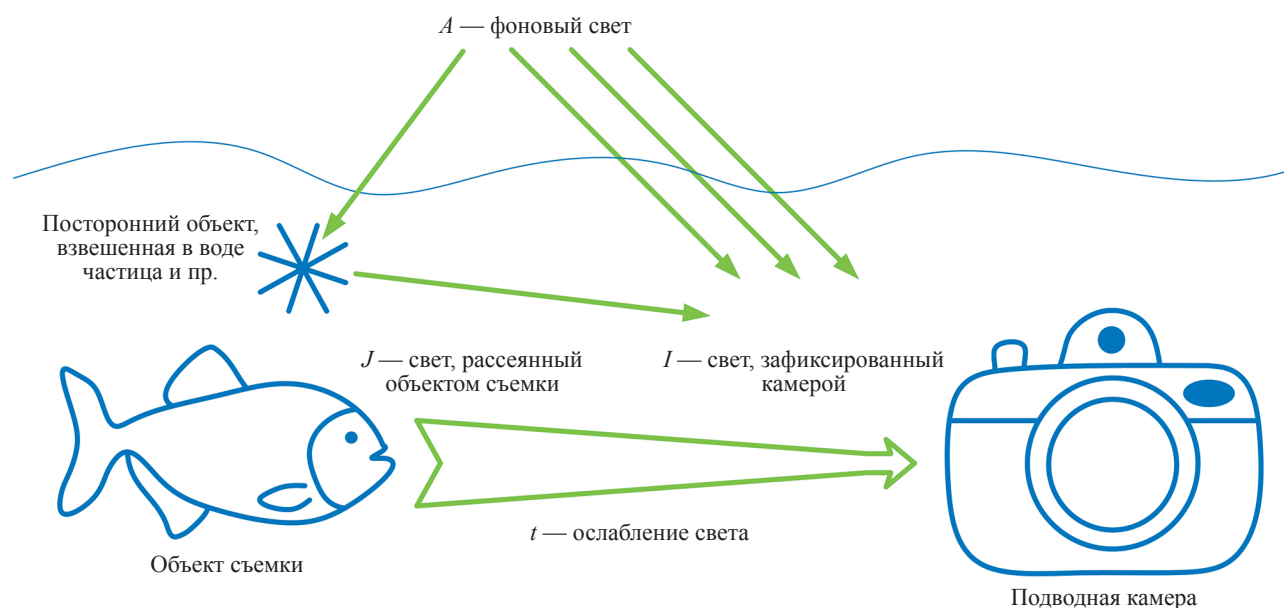


Рис. 1. Упрощенная модель формирования изображения

Fig. 1. Simplified image formation model

кадра, появлению дымки на изображении. $I_c(x)$ и $J_c(x)$ нормированы и их значения находятся в пределах диапазона от 0 до 1.

Карта пропускания t , характеризующая ослабление интенсивности света в каждом канале изображения, в упрощенном виде может быть определена на основе закона Бера–Ламберта [4, 5], характеризующего экспоненциальный закон затухания оптического излучения при распространении в воде:

$$t_c(x) = \exp(-\beta_c z(x)),$$

где β_c — коэффициент ослабления оптического излучения в воде для соответствующего цветового канала (длины волны оптического излучения); $z(x)$ — расстояние от объектива камеры до объекта на изображении с координатами x .

Экспоненциальный закон широко применяется в различных методиках восстановления подводных изображений для описания процесса затухания света в воде.

Применение закона Бера–Ламберта является наиболее простым, но наименее точным методом моделирования распространения света в водной среде.

Предпринимаются попытки уточнить закон ослабления света в воде. Так, в работе [16] предложен более точный закон, описывающий распространение света в воде за счет отдельной оценки коэффициента ослабления для прямого света, рассеянного объектом съемки, и фонового света. Однако для осуществления оценки коэффициентов ослабления используются цветные таблицы, размещаемые возле объекта съемки. Подобный подход позволяет точно оценить и в дальнейшем восстановить естественные цвета подводных объектов, но требует проведения подготовки к съемке, которая трудно осуществима в случае глубоководной съемки с применением подводных аппаратов и невозможна в случае осуществления площадной съемки при осмотре

состояния подводного объекта (трубопровода, подводного сооружения и других), выполнении глубоководных геологоразведочных работ и других.

Более совершенный вариант алгоритма [16] предложен в работе [17] и позволяет за счет применения более разветвленного алгоритма осуществлять настройку метода в зависимости от особенностей условий съемки. Однако данный алгоритм неприменим к результатам глубоководной съемки в связи с отсутствием фоновой освещенности и использованием искусственных источников света для освещения подводной сцены, так как это приводит к некорректной оценке глубины сцены и, как следствие, неверному восстановлению цветов подводных объектов.

В настоящей работе предлагается вместо упрощенного экспоненциального закона использовать более точный подход, заключающийся в численном решении уравнения переноса излучения с применением метода Монте-Карло [18, 19]. Подход, заключающийся в моделировании распространения оптических лучей в воде, нашел широкое применение при решении задачи моделирования беспроводных подводных систем оптической связи, однако в задачах восстановления подводных изображений до настоящего времени не использовался. При этом подобный подход основан на физической модели, учитывающей основные параметры водной среды (замутненность, соленость, температура воды, концентрация хлорофилла, соленость и т. д.), а также параметры камеры (угол обзора, диаметр объектива и другие), и позволяет учитывать влияние на распространение света не только ослабления, но также явления многократного рассеяния, дисперсии оптического излучения и параметров объектива фотоприемного устройства [20].

Алгоритм моделирования распространения оптического излучения в воде с использованием метода Монте-Карло приведен на рис. 2.

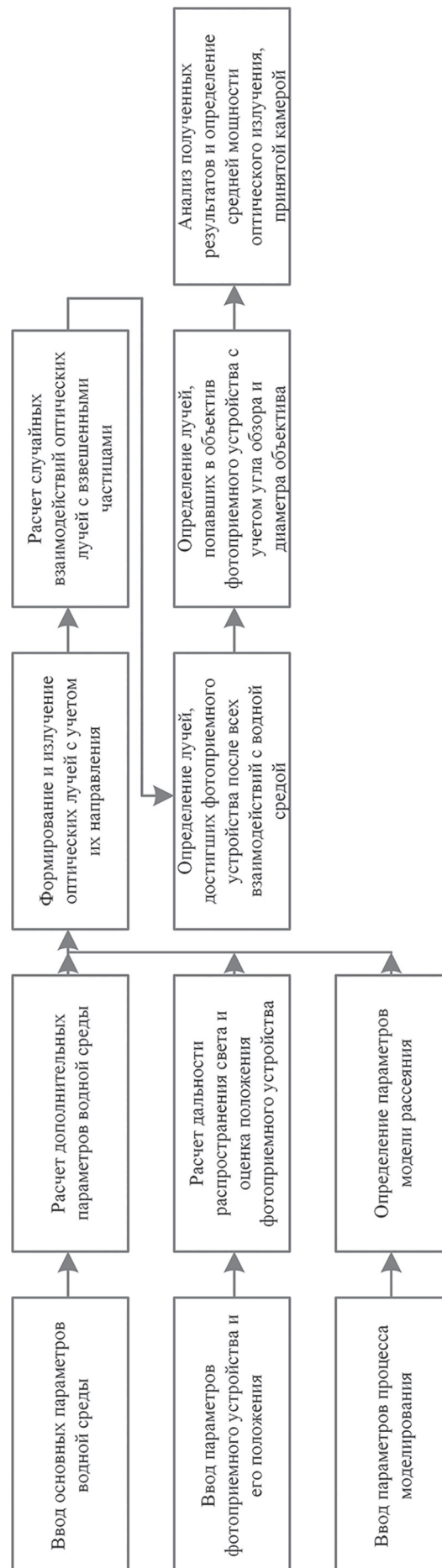


Рис. 2. Алгоритм моделирования распространения оптического излучения в воде с использованием метода Монте-Карло
 Fig. 2. An algorithm for modeling the propagation of optical radiation in water using the Monte Carlo method

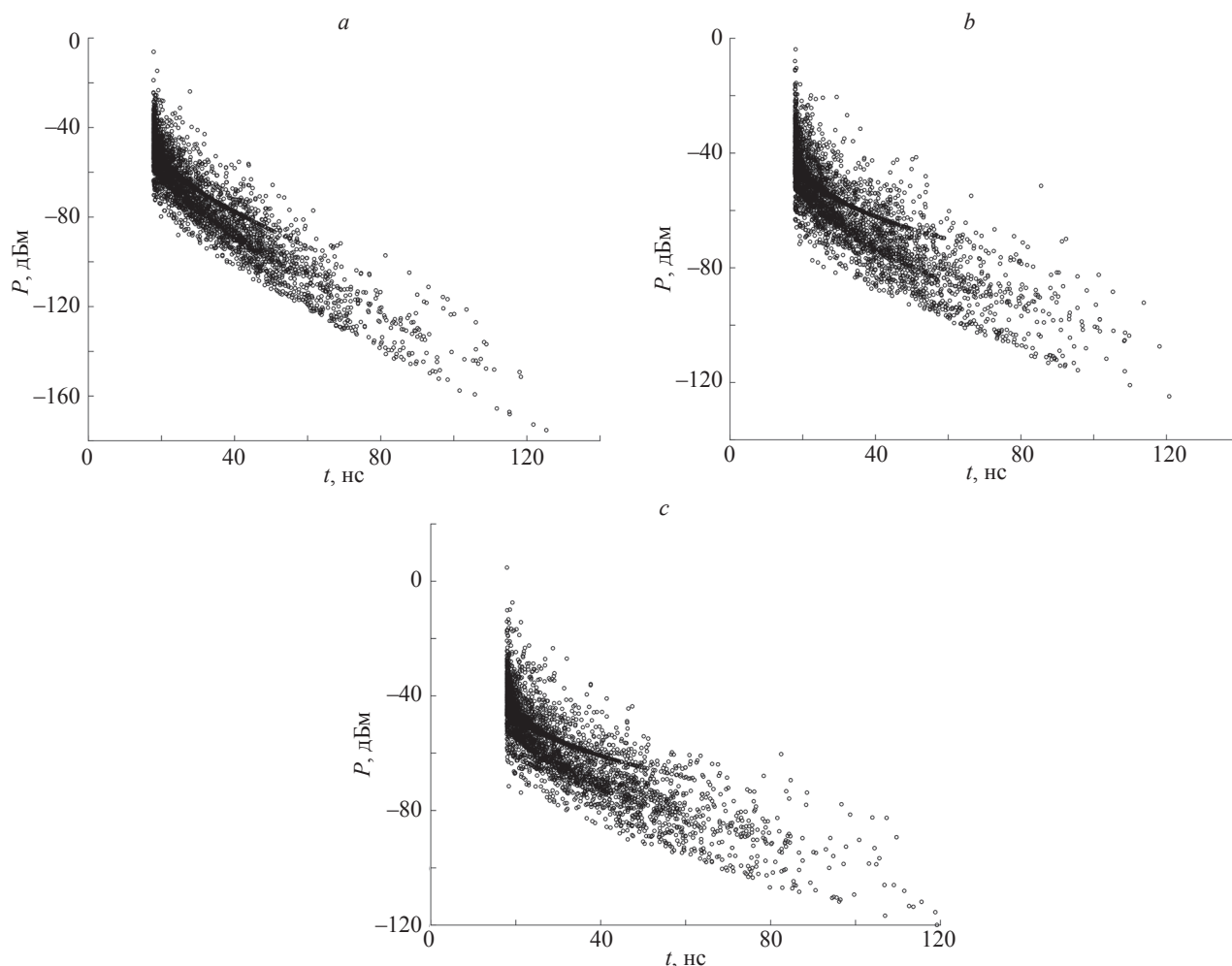


Рис. 3. Результаты моделирования распространения оптических лучей в воде с применением метода Монте-Карло для длин волн: 670 нм (a), 530 нм (b) и 485 нм (c). P — мощность луча, t — время прихода луча

Fig. 3. The results of modeling the propagation of optical rays in water using the Monte Carlo method for wavelengths of 670 nm (a), 530 nm (b) and 485 nm (c). The vertical axis is the ray power, the horizontal is the time of arrival of the ray

В результате моделирования распространения оптических лучей в воде с применением метода Монте-Карло получается матрица, содержащая сведения о мощности и времени приема каждого луча фотоприемным устройством. Результаты моделирования для чистой океанской воды и дальности 4 м приведены на рис. 3.

В связи с тем, что метод Монте-Карло является вероятностным методом, то с целью получения достоверного результата в качестве мощности оптического излучения используется значение средней мощности оптических лучей, принятых фотоприемным устройством. Это позволяет исключить влияние вероятности некорректной оценки мощности отдельных лучей.

Так как параметры модели, используемой для трассировки оптических лучей методом Монте-Карло, имеют четкий физический смысл, то в случае необходимости может быть проведен расчет для наборов параметров, характеризующих водную среду в месте проведения подводной съемки.

Методы количественной оценки качества подводных изображений

Методы количественной оценки качества подводных изображений можно разделить на две основные группы: методы на основе сравнения с эталоном и методы, не требующие эталона. В связи с тем, что при восстановлении реальных подводных изображений отсутствует эталонные версии этих изображений, то методы количественной оценки на основе сравнения с эталоном не могут быть применены. Основные методы количественной оценки качества подводных изображений, не требующих эталона: Underwater Colour Image Quality Evaluation Metric (UCIQE) [21], Underwater Image Quality Measure (UIQM) [22] и Patch-based Contrast Quality Index (PCQI) [23]. Наиболее часто применяемыми для анализа качества подводных изображений методами количественной оценки являются UCIQE и UIQM [4].

Индекс UCIQE основан на комбинации цветности, насыщенности и контрастности изображения в модели CIE-Lab:

$$UCIQE = c_1 \cdot \sigma_c + c_2 \cdot \text{con}_1 + c_3 \cdot \mu_s,$$

где σ_c , con_1 и μ_s — отклонения цветности, контрастности и насыщенности от средних значений; $c_1 = 0,4859$, $c_2 = 0,2745$, $c_3 = 0,2576$ — весовые коэффициенты [4]. Большее значение UCIQE соответствует лучшему балансу между компонентами.

Индекс UIQM представляет собой комбинацию трех показателей: UICM для оценки цветности, UISM для оценки четкости и UIConM для оценки контрастности:

$$UIQM = c_1 \cdot UICM + c_2 \cdot UISM + c_3 \cdot UIConM,$$

где $c_1 = 0,0282$, $c_2 = 0,2953$, $c_3 = 3,5753$ — весовые коэффициенты [4]. Большее значение индекса UIQM характеризует лучшее визуальное восприятие изображения.

Индексы UCIQE и UIQM являются безразмерными.

Описание предлагаемого метода

Предлагаемый в настоящей работе метод обработки подводных изображений основан на уточненной модели в виде численного моделирования трассировки оптических лучей методом Монте-Карло с последующей корректировкой гистограмм каждого канала изображения в цветовом пространстве RGB. Алгоритм обработки изображения разделяется на два основных этапа: предварительный и основной.

Предварительный этап заключается в численном моделировании трассировки оптических лучей методом Монте-Карло [18, 24, 25] с длинами волн, соответствующими трем основным цветам (красному, зеленому и синему) с использованием алгоритма описанного в разделе «Модель формирования изображения» с целью определения матрицы соотношения коэффициентов ослабления оптического излучения в воде для света красного (R), синего (B) и зеленого (G) цветов для различных параметров водной среды и условий съемки:

$$\beta_{RG} = P_R/P_G, \beta_{RB} = P_R/P_B, \beta_{BG} = P_B/P_G.$$

Каждый элемент полученной матрицы представляет собой отношение мощностей оптических сигналов указанных цветов для определенного набора параметров водной среды и дальности между камерой и объектом съемки.

Основной этап алгоритма обработки заключается в восстановлении естественного соотношения яркостей цветовых компонентов изображения на основании сформированной на предварительном этапе матрицы, а также имеющихся сведений об условиях съемки: степени мутности воды (чистая вода, прибрежная вода, мутная вода) и глубины сцены (дальности до объекта съемки) путем коррекции гистограммы каждого канала изображения (R, G, B) с применением значений β_{RG} , β_{RB} , β_{BG} .

Глубина сцены может быть определена с применением альтиметра, дальнометра или иного устройства, установленного на подводном аппарате, используемом для подводной съемки.

Восстановление цветности изображения осуществляется путем растягивания гистограмм каждого канала изображения следующим образом:

$$\begin{aligned} H'_R &= H_R \left\{ 0 \dots \left(\frac{1}{\beta_{RG}\beta_{RB}} \right) \right\}, \\ H'_G &= H_G \left\{ \left(1 - \frac{1}{\beta_{RG}\beta_{BG}} \right) \dots 1 \right\}, \\ H'_B &= H_B \left\{ \left(1 - \frac{1}{\beta_{RB}\beta_{BG}} \right) \dots 1 \right\}, \end{aligned}$$

где R, G, B — индекс цветового канала; $H_{R,G,B}$ — гистограмма соответствующего канала исходного изображения; $H'_{R,G,B}$ — гистограмма соответствующего канала обработанного изображения; в фигурных скобках указаны пределы контрастности изображения в диапазоне $[0, 1]$. С целью повышения контрастности обрабатываемого подводного изображения применяется гамма-коррекция исходного изображения на основании соотношений коэффициентов ослабления оптического излучения в воде (β_{RG} , β_{RB} , β_{BG}):

$$\begin{aligned} I'_c(x) &= I_c(x)^{\gamma_c}, \\ \gamma_R &= \frac{1}{\beta_{RG}\beta_{RB}}, \gamma_G = \frac{1}{\beta_{RG}\beta_{BG}}, \gamma_B = \frac{1}{\beta_{RB}\beta_{BG}}, \end{aligned}$$

где $c \in \{R, G, B\}$ — индекс цветового канала; I_c — исходное изображение; I'_c — обработанное изображение.

Разработанный метод основан на физической модели водной среды, имеет большую вариативность при настройке параметров водной среды на предварительном этапе, а также лишен недостатков методов на основе DCP и MIP, связанных с некорректной обработкой изображений, полученных при использовании искусственного освещения из-за неверной оценки глубины сцены, приводящей к перекомпенсации цветовых компонентов изображения и, следовательно, искажению цветности подводных объектов.

Предварительный этап обработки является более трудоемким в части использования вычислительных ресурсов, но выполняется единожды на этапе подготовки и после формирования матрицы данных не повторяется.

Основной этап обработки прост в реализации и не требователен к наличию вычислительных ресурсов, что позволяет его применять на базе различных платформ, что значительно расширяет область применения предлагаемого метода. Кроме того, разделение алгоритма на две части обеспечивает существенное ускорение процесса за счет того, что при обработке изображения выполняется только основной этап алгоритма.

Разработанный алгоритм объединяет точность восстановления подводных изображений (так как основан на точной физической модели распространения света в воде и учитывает эффекты ослабления, рассеяния и дисперсии света) и быстроту алгоритмов улучшения качества изображений. При этом точность восстановления цветов подводного объекта определяется достоверностью оценки глубины сцены (дальности до

объекта) и корректностью задания набора параметров физической модели при осуществлении моделирования распространения световых лучей на предварительном этапе алгоритма.

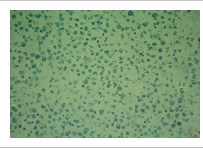
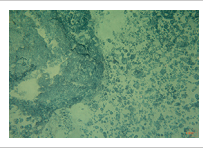
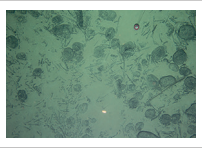
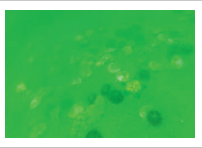
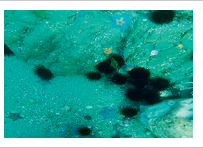
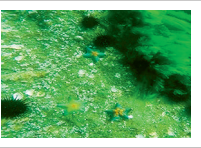
Полученные результаты

Для тестирования выбрано 6 изображений, представленных в табл. 1. Все расчеты выполнены на персональном компьютере на базе операционной системы

Windows 10 с процессором Intel(R) Core(TM) i5-7400 и 16 ГБ оперативной памяти.

Для оценки предлагаемого метода обработки подводных изображений произведено сравнение качественных и количественных показателей со следующими методами обработки: Fusion [27], UDCP+IATP [28], Retinex [29], HE [30] и UWB VCSE [31]. Указанные методы являются широко применяемыми подходами к восстановлению и улучшению качества подводных изображений [4, 5].

Таблица 1. Исходные изображения
Table 1. Original images

Номер изображения					
1	2	3	4	5	6
Размер изображения, пиксели					
5184 × 3456			720 × 405	400 × 300	
					

Примечание. Изображения 1–3 получены в ходе геологоразведочных работ в районе Кларион-Клиппертон в северной приэкваториальной области Северо-Восточной котловины Тихого океана с глубинами более 5500 м с использованием искусственного освещения; 4–6 взяты из базы изображений LSUI [26].

Таблица 2. Результаты качественной оценки эффективности методов обработки
Table 2. The results of a qualitative assessment of the effectiveness of processing methods

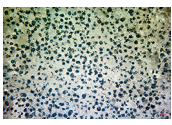
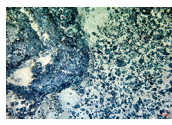
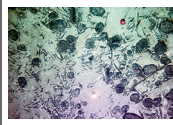



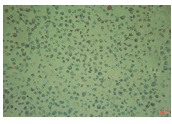
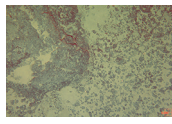
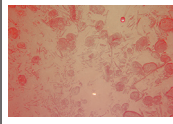
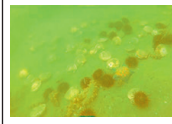
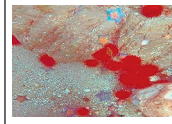
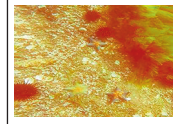
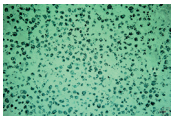
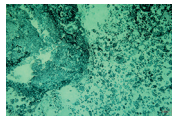
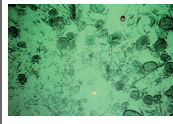
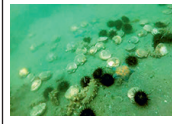
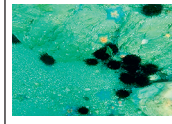
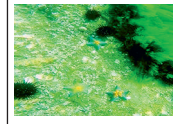
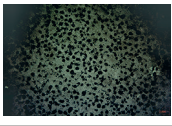
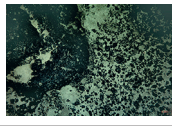
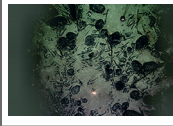
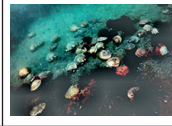
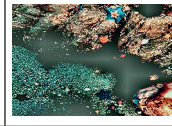

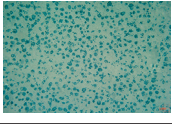
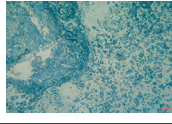
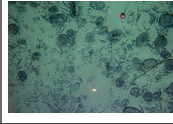


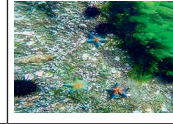
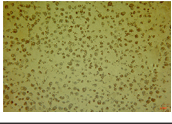
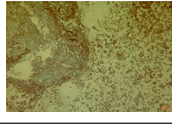
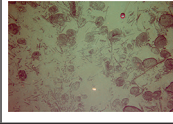
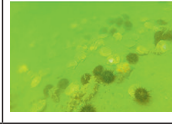
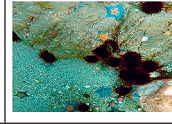
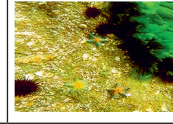
Метод обработки	Номер изображения					
	1	2	3	4	5	6
Fusion						
UDCP + IATP						
Retinex						
HE						
UWB VCSE						
Предлагаемый						

Таблица 3. Результаты количественной оценки эффективности методов обработки
 Table 3. Results of a quantitative assessment of the effectiveness of processing methods

Метод обработки	Номер изображения						Номер изображения					
	1	2	3	4	5	6	1	2	3	4	5	6
	UCIQE						UIQM					
IN	0,373	0,397	0,359	0,338	0,516	0,541	2,978	3,185	2,93	−1,372	2,32	3,561
Fusion	0,576	0,59	0,556	0,614	0,605	0,589	4,946	4,635	4,878	2,469	3,895	4,626
UDCP + IATP	0,358	0,374	0,362	0,352	0,505	0,515	2,998	3,587	4,451	0,581	5,076	5,562
Retinex	0,483	0,499	0,485	0,513	0,559	0,573	2,690	3,046	3,035	1,177	2,143	3,392
HE	0,452	0,489	0,462	0,52	0,599	0,558	4,390	4,274	3,950	2,408	4,127	3,592
UWB VCSE	0,391	0,422	0,402	0,478	0,551	0,542	2,930	3,369	3,729	2,368	5,179	5,106
Предлагаемый	0,513	0,519	0,489	0,511	0,560	0,631	3,683	4,142	4,586	1,206	4,614	4,656

Примечание. Зеленым цветом выделено два наилучших результата, красным — два наихудших значения для каждого показателя.

Таблица 4. Результаты оценки времени обработки подводных изображений, с
 Table 4. The results of estimating the processing time of underwater images, s

Метод обработки	Номер изображения					
	1	2	3	4	5	6
	Размер изображения, пиксели					
	5184 × 3456			720 × 405	400 × 300	
Fusion	22,83	22,83	27,37	2,23	1,19	1,16
UDCP + IATP	84,17	82,77	83,31	4,09	1,63	1,61
Retinex	7,75	7,65	7,76	0,26	0,19	0,19
HE	15,06	15,15	15,99	1,35	1,09	1,14
UWB VCSE	925,57	937,83	927,49	18,77	9,38	10,81
Предлагаемый	8,23	8,05	8,02	0,11	0,07	0,073

Примечание. Зеленым цветом выделено два наилучших результата, красным — два наихудших значения для каждого показателя.

В табл. 2 приведены результаты восстановления исходных изображений с помощью предлагаемого и указанных методов обработки.

В табл. 3 и 4 приведены результаты количественной оценки эффективности предлагаемого метода.

Из полученных результатов видно, что предлагаемый метод обработки подводных изображений демонстрирует эффективность, качественная и количественная оценка которой близка к наиболее эффективным методам улучшения подводных изображений, либо превосходит их.

При использовании метода Fusion для изображений 1–3 (табл. 2) наблюдается чрезмерное увеличение контрастности и искажение цветности изображения. В результате применения метода UDCP+IATP наблюдается искажение цветности темных участков изображений вплоть до их окрашивания в красные оттенки. Применение метода Retinex обеспечивает недостаточное устранение смещения цветов в сине-зеленую область спектра видимого света. В результате применения метода HE на темных участках изображений возникают серые области с равномерным фоном и отсутствием объектов. Применение метода UWB VCSE для изображений 1–3 привело к недостаточно-

му устранению синей дымки и низкой контрастности изображений.

Необходимо отметить высокую стабильность предлагаемого метода в части эффективности обработки подводных изображений (табл. 3 и 4). Значения количественных показателей оценки для предлагаемого метода сопоставимы, либо превосходят значения аналогичных показателей для остальных методов. Предлагаемый метод продемонстрировал второй результат среди анализируемых методов по скорости обработки изображений. Ускорение достигается разделением алгоритма на два этапа — подготовительный и основной. Подготовительный этап выполняется заранее единожды и далее при обработке изображений повторно не выполняется.

Заключение

В работе представлен метод обработки подводных изображений, направленный на восстановление естественных цветов подводных объектов и улучшение контрастности изображений.

В результате качественной и количественной оценок с использованием реальных глубоководных фотографий

морского дна, полученных в ходе геологоразведочных работ, показана высокая эффективность и быстродействие предлагаемого метода, сопоставимая или превосходящая в отдельных случаях показатели известных аналогичных алгоритмов обработки подводных изображений.

Преимущество предлагаемого метода заключается в применении физической модели высокой точности на

основе численного моделирования трассировки оптических лучей методом Монте-Карло.

Основное направление дальнейших исследований заключается в реализации алгоритма восстановления подводных изображений в том числе для любого типа воды, а также замутненной воды с высоким содержанием хлорофилла.

Литература

1. Zhou J., Pang L., Zhang D., Zhang W. Underwater image enhancement method via multi-interval subhistogram perspective equalization // *IEEE Journal of Oceanic Engineering*. 2023. V. 48. N 2. P. 474–488. <https://doi.org/10.1109/joe.2022.3223733f>
2. Zhang W., Liu W., Li L. Underwater single-image restoration with transmission estimation using color constancy // *Journal of Marine Science and Engineering*. 2022. V. 10. N 3. P. 430. <https://doi.org/10.3390/jmse10030430>
3. Анисимов И.М., Римский-Корсаков Н.А., Тронза С.Н. Развитие глубоководных технологий визуальных наблюдений рельефа дна и подводных объектов // *Международный журнал прикладных и фундаментальных исследований*. 2019. № 10. С. 149–153. <https://doi.org/10.17513/mjpf.12883>
4. Zhou J., Zhang D., Zhang W. The Classical and state-of-the-art approaches for underwater image defogging: a comprehensive survey // *Frontiers of Information Technology & Electronic Engineering*. 2020. V. 21. N 12. P. 1745–1769. <https://doi.org/10.1631/FITEE.2000190>
5. Zhang W., Dong L., Pan X., Zou P., Qin L., Xu W. A Survey of restoration and enhancement for underwater images // *IEEE Access*. 2019. V. 7. P. 182259–182279. <https://doi.org/10.1109/ACCESS.2019.2959560>
6. He K., Sun J., Tang X. Single image haze removal using dark channel prior // *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 2011. V. 33. N 12. P. 2341–2353. <https://doi.org/10.1109/TPAMI.2010.168>
7. Chang H.H., Chen P.F., Guo J.K., Sung C.C. A self-adaptive single underwater image restoration algorithm for improving graphic quality // *Eurasip Journal on Image and Video Processing*. 2020. V. 2020. N 1. P. 41. <https://doi.org/10.1186/s13640-020-00528-0>
8. Zhang T., Li Q., Li Y., Liu X. Underwater Optical Image Restoration Method for Natural/Artificial Light // *Journal of Marine Science and Engineering*. 2023. V. 11. N 3. P. 470. <https://doi.org/10.3390/jmse11030470>
9. Iqbal K., Odetayo M., James A. Enhancing the low quality images using unsupervised colour correction method // *Proc. of the IEEE International Conference on Systems, Man and Cybernetics*. 2010. P. 1703–1709. <https://doi.org/10.1109/icsmc.2010.5642311>
10. Pandey V., Padmakumar A., Padmakumar A., Rasal T., Khandelwal H. Underwater image enhancement and restoration using AI & ML // *International Journal of Research Publication and Reviews*. 2023. V. 4. N 4. P. 3372–3377.
11. Joel Fathimson, J. Bibis S., Aswanth R., Gayatri S. Underwater image restoration using UICCS Method in Matlab // *International Journal of New Technology and Research (IJNTR)*. 2018. V. 4. N 2. P. 1–6.
12. Li C., Guo C., Ren W., Cong R., Hou J., Kwong S., Tao D. An underwater image enhancement benchmark dataset and beyond // *IEEE Transactions on Image Processing*. 2020. V. 29. P. 4376–4389. <https://doi.org/10.1109/TIP.2019.2955241>
13. Peng Y-T., Cosman P.C. Underwater image restoration based on image blurriness and light absorption // *IEEE Transactions on Image Processing*. 2017. V. 26. N 4. P. 1579–1594. <https://doi.org/10.1109/TIP.2017.2663846>
14. Xu Y., Wang H., Cooper G.D., Rong S., Sun W. Learning-based dark and blurred underwater image restoration // *Complexity*. 2020. V. 2020. P. 6549410. <https://doi.org/10.1155/2020/6549410>
15. Schechner Y.Y., Karpel N. Recovery of underwater visibility and structure by polarization analysis // *IEEE Journal of Oceanic Engineering*. 2005. V. 30. N 3. P. 570–587. <https://doi.org/10.1109/JOE.2005.850871>

References

1. Zhou J., Pang L., Zhang D., Zhang W. Underwater image enhancement method via multi-interval subhistogram perspective equalization. *IEEE Journal of Oceanic Engineering*. 2023. V. 48. N 2. P. 474–488. <https://doi.org/10.1109/joe.2022.3223733>
2. Zhang W., Liu W., Li L. Underwater single-image restoration with transmission estimation using color constancy. *Journal of Marine Science and Engineering*. 2022. V. 10. N 3. P. 430. <https://doi.org/10.3390/jmse10030430>
3. Anisimov I.M., Rimskiy-Korsakov N.A., Tronza S.N. Bottom relief and underwater objects visual deep-water observations technologies development. *International Journal of Applied and Fundamental Research*, 2019, no. 10, pp. 149–153. (in Russian). <https://doi.org/10.17513/mjpf.12883>
4. Zhou J., Zhang D., Zhang W. The Classical and state-of-the-art approaches for underwater image defogging: a comprehensive survey. *Frontiers of Information Technology & Electronic Engineering*, 2020, vol. 21, no. 12, pp. 1745–1769. <https://doi.org/10.1631/FITEE.2000190>
5. Zhang W., Dong L., Pan X., Zou P., Qin L., Xu W. A Survey of restoration and enhancement for underwater images. *IEEE Access*, 2019, vol. 7, pp. 182259–182279. <https://doi.org/10.1109/ACCESS.2019.2959560>
6. He K., Sun J., Tang X. Single image haze removal using dark channel prior. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2011, vol. 33, no. 12, pp. 2341–2353. <https://doi.org/10.1109/TPAMI.2010.168>
7. Chang H.H., Chen P.F., Guo J.K., Sung C.C. A self-adaptive single underwater image restoration algorithm for improving graphic quality. *Eurasip Journal on Image and Video Processing*, 2020, vol. 2020, no. 1, pp. 41. <https://doi.org/10.1186/s13640-020-00528-0>
8. Zhang T., Li Q., Li Y., Liu X. Underwater Optical Image Restoration Method for Natural/Artificial Light. *Journal of Marine Science and Engineering*, 2023, vol. 11, no. 3, pp. 470. <https://doi.org/10.3390/jmse11030470>
9. Iqbal K., Odetayo M., James A. Enhancing the low quality images using unsupervised colour correction method. *Proc. of the IEEE International Conference on Systems, Man and Cybernetics*, 2010, pp. 1703–1709. <https://doi.org/10.1109/icsmc.2010.5642311>
10. Pandey V., Padmakumar A., Padmakumar A., Rasal T., Khandelwal H. Underwater image enhancement and restoration using AI & ML. *International Journal of Research Publication and Reviews*, 2023, vol. 4, no. 4, pp. 3372–3377.
11. Joel Fathimson, J. Bibis S., Aswanth R., Gayatri S. Underwater image restoration using UICCS Method in Matlab. *International Journal of New Technology and Research (IJNTR)*, 2018, vol. 4, no. 2, pp. 1–6.
12. Li C., Guo C., Ren W., Cong R., Hou J., Kwong S., Tao D. An underwater image enhancement benchmark dataset and beyond. *IEEE Transactions on Image Processing*, 2020, vol. 29, pp. 4376–4389. <https://doi.org/10.1109/TIP.2019.2955241>
13. Peng Y-T., Cosman P.C. Underwater image restoration based on image blurriness and light absorption. *IEEE Transactions on Image Processing*, 2017, vol. 26, no. 4, pp. 1579–1594. <https://doi.org/10.1109/TIP.2017.2663846>
14. Xu Y., Wang H., Cooper G.D., Rong S., Sun W. Learning-based dark and blurred underwater image restoration. *Complexity*, 2020, vol. 2020, pp. 6549410. <https://doi.org/10.1155/2020/6549410>
15. Schechner Y.Y., Karpel N. Recovery of underwater visibility and structure by polarization analysis. *IEEE Journal of Oceanic Engineering*, 2005, vol. 30, no. 3, pp. 570–587. <https://doi.org/10.1109/JOE.2005.850871>

16. Akkaynak D., Treibitz T. A revised underwater image formation model // *Proc. of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2018. P. 6723–6732. <https://doi.org/10.1109/CVPR.2018.00703>
17. Pei S.C., Chen C.Y. Underwater images enhancement by revised underwater images formation model // *IEEE Access*. 2022. V. 10. P. 108817–108831. <https://doi.org/10.1109/ACCESS.2022.3213340>
18. Mobley C.D., Gentili B., Gordon H.R., Jin Z., Kattawar G.W., Morel A., Reinersman P., Stamnes K., Stavn R.H. Comparison of numerical models for computing underwater light fields // *Applied Optics*. 1993. V. 32. N 36. P. 7484–7504. <https://doi.org/10.1364/AO.32.007484>
19. Semernik I.V., Demyanenko A.V., Samonova C.V., Bender O.V., Tarasenko A. Modelling of an underwater wireless optical communication channel // *Proc. of the Radiation and Scattering of Electromagnetic Waves (RSEMW)*. 2023. P. 468–471. <https://doi.org/10.1109/rsemw58451.2023.10202133>
20. Ramley I., Alzayed H.M., Al-Hadeethi Y., Chen M., Barasheed A.Z. An overview of underwater optical wireless communication channel simulations with a focus on the Monte Carlo method // *Mathematics*. 2024. V. 12. N 24. P. 3904. <https://doi.org/10.3390/math12243904>
21. Yang M., Sowmya A. An underwater color image quality evaluation metric // *IEEE Transactions on Image Processing*. 2015. V. 24. N 12. P. 6062–6071. <https://doi.org/10.1109/TIP.2015.2491020>
22. Panetta K., Gao C., Agaian S. Human-visual-system-inspired underwater image quality measures // *IEEE Journal of Oceanic Engineering*. 2016. V. 41. N 3. P. 541–551. <https://doi.org/10.1109/JOE.2015.2469915>
23. Wang S.Q., Ma K.D., Yeganeh H., Wang Z., Lin W. A patch-structure representation method for quality assessment of contrast changed images // *IEEE Signal Processing Letters*. 2015. V. 22. N 12. P. 2387–2390. <https://doi.org/10.1109/LSP.2015.2487369>
24. Семерник И.В. Модель водной среды для точного моделирования распространения оптического излучения // *Кузнечно-штамповочное производство. Обработка материалов давлением*. 2024. № 7. С. 61–66.
25. Семерник И.В. Моделирование распространения оптического излучения для восстановления изображений в рамках морских геологоразведочных работ // *Международный научно-исследовательский журнал*. 2024. № 8 (146). С. 16. <https://doi.org/10.60797/IRJ.2024.146.24>
26. Peng L., Zhu C., Bian L. U-Shape transformer for underwater image enhancement // *IEEE Transactions on Image Processing*. 2023. V. 32. P. 3066–3079. <https://doi.org/10.1109/TIP.2023.3276332>
27. Ancuti C., Ancuti C.O., Haber T., Bekaert P. Enhancing underwater images and videos by fusion // *Proc. of the IEEE Conference on Computer Vision and Pattern Recognition*. 2012. P. 81–88. <https://doi.org/10.1109/cvpr.2012.6247661>
28. Zhang M., Peng J. Underwater image restoration based on a new underwater image formation Model // *IEEE Access*. 2018. V. 6. P. 58634–58644. <https://doi.org/10.1109/ACCESS.2018.2875344>
29. Fu X., Zhuang P., Huang Y., Liao Y., Zhang X.-P., Ding X. A Retinex-based enhancing approach for single underwater image // *Proc. of the IEEE International Conference on Image Processing (ICIP)*. 2014. P. 4572–4576. <https://doi.org/10.1109/icip.2014.7025927>
30. Fu X., Fan Z., Ling M., Huang Y., Ding X. Two-step approach for single underwater image enhancement // *Proc. of the International Symposium on Intelligent Signal Processing and Communication Systems (ISPACS)*. 2017. P. 789–794. <https://doi.org/10.1109/ispacs.2017.8266583>
31. Li X., Hou G., Tan L., Liu W. A hybrid framework for underwater image enhancement // *IEEE Access*. 2020. V. 8. P. 197448–197462. <https://doi.org/10.1109/10.1109/ACCESS.2020.3034275>
16. Akkaynak D., Treibitz T. A revised underwater image formation model. *Proc. of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2018, pp. 6723–6732. <https://doi.org/10.1109/CVPR.2018.00703>
17. Pei S.C., Chen C.Y. Underwater images enhancement by revised underwater images formation model. *IEEE Access*, 2022, vol. 10, pp. 108817–108831. <https://doi.org/10.1109/ACCESS.2022.3213340>
18. Mobley C.D., Gentili B., Gordon H.R., Jin Z., Kattawar G.W., Morel A., Reinersman P., Stamnes K., Stavn R.H. Comparison of numerical models for computing underwater light fields. *Applied Optics*, 1993, vol. 32, no. 36, pp. 7484–7504. <https://doi.org/10.1364/AO.32.007484>
19. Semernik I.V., Demyanenko A.V., Samonova C.V., Bender O.V., Tarasenko A. Modelling of an underwater wireless optical communication channel. *Proc. of the Radiation and Scattering of Electromagnetic Waves (RSEMW)*, 2023, pp. 468–471. <https://doi.org/10.1109/rsemw58451.2023.10202133>
20. Ramley I., Alzayed H.M., Al-Hadeethi Y., Chen M., Barasheed A.Z. An overview of underwater optical wireless communication channel simulations with a focus on the Monte Carlo method. *Mathematics*, 2024, vol. 12, no. 24, P. 3904. <https://doi.org/10.3390/math12243904>
21. Yang M., Sowmya A. An underwater color image quality evaluation metric. *IEEE Transactions on Image Processing*, 2015, vol. 24, no. 12, pp. 6062–6071. <https://doi.org/10.1109/TIP.2015.2491020>
22. Panetta K., Gao C., Agaian S. Human-visual-system-inspired underwater image quality measures. *IEEE Journal of Oceanic Engineering*, 2016, vol. 41, no. 3, pp. 541–551. <https://doi.org/10.1109/JOE.2015.2469915>
23. Wang S.Q., Ma K.D., Yeganeh H., Wang Z., Lin W. A patch-structure representation method for quality assessment of contrast changed images. *IEEE Signal Processing Letters*, 2015, vol. 22, no. 12, pp. 2387–2390. <https://doi.org/10.1109/LSP.2015.2487369>
24. Semernik I.V. Aquatic environment model for accurate simulation of optical radiation propagation. *KSHP-OMD*, 2024, no. 7. pp. 61–66. (in Russian)
25. Semernik I.V. Modelling optical propagation for image reconstruction in marine mineral exploration. *International Research Journal*, 2024, no. 8 (146), pp. 16. (in Russian). <https://doi.org/10.60797/IRJ.2024.146.24>
26. Peng L., Zhu C., Bian L. U-Shape transformer for underwater image enhancement. *IEEE Transactions on Image Processing*, 2023, vol. 32, pp. 3066–3079. <https://doi.org/10.1109/TIP.2023.3276332>
27. Ancuti C., Ancuti C.O., Haber T., Bekaert P. Enhancing underwater images and videos by fusion. *Proc. of the IEEE Conference on Computer Vision and Pattern Recognition*, 2012, pp. 81–88. <https://doi.org/10.1109/cvpr.2012.6247661>
28. Zhang M., Peng J. Underwater image restoration based on a new underwater image formation Model. *IEEE Access*, 2018, vol. 6, pp. 58634–58644. <https://doi.org/10.1109/ACCESS.2018.2875344>
29. Fu X., Zhuang P., Huang Y., Liao Y., Zhang X.-P., Ding X. A Retinex-based enhancing approach for single underwater image. *Proc. of the IEEE International Conference on Image Processing (ICIP)*, 2014, pp. 4572–4576. <https://doi.org/10.1109/icip.2014.7025927>
30. Fu X., Fan Z., Ling M., Huang Y., Ding X. Two-step approach for single underwater image enhancement. *Proc. of the International Symposium on Intelligent Signal Processing and Communication Systems (ISPACS)*, 2017, pp. 789–794. <https://doi.org/10.1109/ispacs.2017.8266583>
31. Li X., Hou G., Tan L., Liu W. A hybrid framework for underwater image enhancement. *IEEE Access*, 2020, vol. 8, pp. 197448–197462. <https://doi.org/10.1109/10.1109/ACCESS.2020.3034275>

Авторы

Семерник Иван Владимирович — кандидат технических наук, главный конструктор, АО «Южморгеология», Геленджик, 353461, Российская Федерация, 55990676800, <https://orcid.org/0000-0002-0238-4154>, ione7nick@yandex.ru

Самонова Кристина Владимировна — кандидат экономических наук, ведущий специалист, АО «Южморгеология», Геленджик, 353461, Российская Федерация, 57194240990, <https://orcid.org/0000-0002-2449-9921>, kristi-rosa@yandex.ru

Authors

Ivan V. Semernik — PhD, Chief Design Engineer, JSC Yuzhmorgeologia, Gelendzhik, 353461, Russian Federation, 55990676800, <https://orcid.org/0000-0002-0238-4154>, ione7nick@yandex.ru

Christina V. Samonova — PhD (Economics), Leading Specialist, JSC Yuzhmorgeologia, Gelendzhik, 353461, Russian Federation, 57194240990, <https://orcid.org/0000-0002-2449-9921>, kristi-rosa@yandex.ru

doi: 10.17586/2226-1494-2025-25-3-428-437

УДК 004.056

Анализ криптографической стойкости хеш-функции SHA-256 при помощи SAT-подхода

Вадим Валерьевич Давыдов¹✉, Михаил Денисович Пихтовников²,
Анастасия Павловна Кирьянова³, Олег Сергеевич Заикин⁴

¹ ООО «КуАпп», Москва, 121205, Российская Федерация

¹ Санкт-Петербургский государственный университет аэрокосмического приборостроения, 190000, Санкт-Петербург, Российская Федерация

^{1,3} Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация

² Южный федеральный университет, Таганрог, 347922, Российская Федерация

⁴ Новосибирский государственный университет, Новосибирск, 630090, Российская Федерация

⁴ Институт динамики систем и теории управления имени В.М. Матросова Сибирского отделения Российской академии наук, Иркутск, 664033, Российская Федерация

¹ vadimdavydov@outlook.com✉, <https://orcid.org/0000-0002-5544-2434>

² pikhtovnikov347@mail.ru, <https://orcid.org/0009-0000-2173-0844>

³ anastaciakosanovskaya@gmail.com, <https://orcid.org/0009-0006-0344-5111>

⁴ oleg.zaikin@icc.ru, <https://orcid.org/0000-0002-0145-5010>

Аннотация

Введение. В современных системах обеспечения информационной безопасности криптографические хеш-функции играют значительную роль и выполняют такие важные задачи, как обеспечение целостности данных и их эффективное сжатие. Одной из наиболее значимых и широко применяемых криптографических хеш-функций является SHA-256 из семейства SHA-2. Исследование криптографической стойкости SHA-256 является актуальной научной задачей и решается с применением современных подходов криптоанализа к атакам нахождения прообразов и коллизий с акцентом на практическую осуществимость таких атак. **Метод.** В представленной работе для поиска прообразов неполнораундовых версий функции сжатия SHA-256 применен логический криптоанализ, согласно которому исходная задача криптоанализа сводится к проблеме булевой выполнимости (SAT). Для поиска коллизий совместно применены логический и дифференциальный криптоанализы. **Основные результаты.** Выполнено сравнение эффективности различных способов сведения функции сжатия SHA-256 к SAT. Впервые найдены прообразы для 17- и 18-раундовых функций сжатия SHA-256, а также прообразы для ослабленной 19-раундовой функции сжатия. Построены базовые дифференциальные пути, с помощью которых быстрее найдены коллизии 18-раундовой функции сжатия. В результате сведения к SAT известных дифференциальных путей найдены коллизии 19-раундовой функции сжатия. **Обсуждение.** Продемонстрирована возможность комбинирования двух методов криптоанализа с целью повышения эффективности анализа криптографических алгоритмов. Результаты исследования подтвердили, что полнораундовая хеш-функция SHA-256 остается устойчивой к атакам, направленным на нахождение прообразов и коллизий, в рамках примененного SAT-подхода.

Ключевые слова

криптографическая хеш-функция, SHA-256, SAT, логический криптоанализ, дифференциальный криптоанализ

Благодарности

Работа Вадима Валерьевича Давыдова и Анастасии Павловны Кирьяновой выполнена в рамках государственного задания (проект FSER-2025-0003). Олег Сергеевич Заикин выполнил свою часть работы при поддержке Математического центра в Академгородке, соглашение с Министерством науки и высшего образования Российской Федерации № 075-15-2025-349. Работа является расширенной версией результатов, полученных в рамках летней школы-конференции «Криптография и информационная безопасность» в 2024 году.

Ссылка для цитирования: Давыдов В.В., Пихтовников М.Д., Кирьянова А.П., Заикин О.С. Анализ криптографической стойкости хеш-функции SHA-256 при помощи SAT-подхода // Научно-технический вестник информационных технологий, механики и оптики. 2025. Т. 25, № 3. С. 428–437. doi: 10.17586/2226-1494-2025-25-3-428-437

Analysis of the cryptographic strength of the SHA-256 hash function using the SAT approach

Vadim V. Davydov¹✉, Michail D. Pikhtovnikov², Anastasia P. Kiryanova³, Oleg S. Zaikin⁴

¹ QApp, Moscow, 121205, Russian Federation

¹ Saint Petersburg State University of Aerospace Instrumentation (SUAI), 190000, Saint Petersburg, Russian Federation

^{1,3} ITMO University, Saint Petersburg, 197101, Russian Federation

² Southern Federal University, Taganrog, 347922, Russian Federation

⁴ Novosibirsk State University, Novosibirsk, 630090, Russian Federation

⁴ Matrosov Institute for System Dynamics and Control Theory of Siberian Branch of RAS, Irkutsk, 664033, Russian Federation

¹ vadmadvydov@outlook.com✉, <https://orcid.org/0000-0002-5544-2434>

² pikhtovnikov347@mail.ru, <https://orcid.org/0009-0000-2173-0844>

³ anastaciakosanovskaya@gmail.com, <https://orcid.org/0009-0006-0344-5111>

⁴ oleg.zaikin@icc.ru, <https://orcid.org/0000-0002-0145-5010>

Abstract

Cryptographic hash functions play a significant role in modern information security systems by ensuring data integrity and enabling efficient data compression. One of the most important and widely used cryptographic hash functions is SHA-256 that belongs to the SHA-2 family. In this regard, the study of SHA-256 cryptographic resistance using modern cryptanalysis approaches to preimage and collision attacks with an emphasis on the practical feasibility of such attacks is an urgent scientific task. To search for preimages of round-reduced versions of the SHA-256 compression function, the logical cryptanalysis was applied, i.e., cryptanalysis problems were reduced to the Boolean satisfiability problem (SAT). For collision attacks, a combination of logical and differential cryptanalysis was utilized. The work presents a comparison between various methods for reducing the SHA-256 compression function to SAT and evaluates their efficiency. As a result of the work, preimages for 17- and 18-round SHA-256 compression functions were found for the first time as well as preimages for a weakened 19-round compression function. Basic differential paths were constructed, which facilitated faster finding of collisions for the 18-round compression function. Known differential paths were reduced in SAT leading to finding collisions for the 19-round compression function. The work demonstrates the possibility of combining two cryptanalysis methods to enhance the efficiency of analyzing cryptographic algorithms. The results of the study confirm that the full-round SHA-256 hash function remains resistant to preimage and collision attacks within the scope of the applied SAT-based approach.

Keywords

cryptographic hash function, SHA-256, SAT, logical cryptanalysis, differential cryptanalysis

Acknowledgements

The work of Vadim Davydov and Anastasia Kiryanova was performed within the framework of the State Assignment (project No. FSER-2025-0003). Oleg Zaikin was funded by the Mathematical Center in Akademgorodok under the Agreement No. 075-15-2025-349 with the Ministry of Science and Higher Education of the Russian Federation. The present paper is an extension of results obtained during the summer school-conference “Cryptography and information security” in 2024.

For citation: Davydov V.V., Pikhtovnikov M.D., Kiryanova A.P., Zaikin O.S. Analysis of the cryptographic strength of the SHA-256 hash function using the SAT approach. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2025, vol. 25, no. 3, pp. 428–437 (in Russian). doi: 10.17586/2226-1494-2025-25-3-428-437

Введение

Хеш-функция — алгоритм, преобразовывающий входные данные и позволяющий получить их «отпечаток» (хеш; в большинстве случаев — битовую строку меньшей длины, чем входная). Для использования в задачах криптографии хеш-функция должна обладать свойствами стойкости к поиску первого прообраза, второго прообраза и коллизий. Криптографические хеш-функции всегда детерминированы, при этом небольшое изменение входа очень сильно влияет на выход — такое свойство называется «лавинным эффектом». На практике очень сложно найти коллизию, т. е. два разных сообщения с одинаковым хеш-значением,

однако, так как пространство хешей много меньше пространства сообщений (если нет ограничений по длине входа), коллизии у любой хеш-функции существуют всегда.

Актуальность анализа криптографических хеш-функций обуславливается их широкой применимостью и популярностью: с их помощью можно генерировать и проверять электронные подписи, формировать блокчейн, обеспечивать безопасное хранение и выполнять проверку верности паролей, проверять целостность файлов, использовать в качестве основы для генератора псевдослучайных чисел и многое другое. Из широкой распространенности вытекает необходимость в анализе уязвимостей используемых криптографических

хеш-функций. Криптоаналитики используют различные методы и атаки для нахождения прообразов и коллизий, такие как атака дней рождения, метод «встречи посередине», дифференциальные пути, логический криптоанализ и многое другое. В логическом криптоанализе исходная задача сводится к проблеме булевой выполнимости (satisfiability или SAT), поэтому он также называется SAT-криптоанализом.

В представленной работе рассмотрена и проанализирована криптографическая хеш-функция SHA-256 [1]. Несмотря на то, что существует более новый стандарт США, хеш-функция SHA-256 все еще используется в различных приложениях, например, в криптовалюте Bitcoin в процессе майнинга и генерации bitcoin-адресов, для аутентификации пакетов программного обеспечения Debian, а также в протоколе TLS 1.3.

Криптографическая стойкость SHA-256 является предметом обширных исследований в научных работах [2]. Атаки направлены на поиск прообраза и поиск коллизий (и их вариаций) функции сжатия SHA-256. На сегодняшний день полнораундовая функция сжатия (64 раунда) остается криптографически стойкой; основной исследовательский интерес сосредоточен на попытках обращения и поиска коллизий неполнораундовых версий. Атаки можно условно разделить на теоретические и практические. Теоретическая атака свидетельствует о снижении асимптотической сложности по сравнению со сложностью полного перебора, однако, на практике такие атаки остаются неприменимыми. Практическая атака подтверждает возможность осуществления реального нахождения прообраза или коллизии в условиях ограниченных вычислительных мощностей.

Для поиска прообраза лучшей теоретической атакой на сегодняшний день является атака, предложенная в работе [3], служащая для нахождения прообраза 45-раундовой версии SHA-256 со сложностью $2^{255,5}$ вызовов функции сжатия. Лучшей практической атакой является работа [4], где был применен логический криптоанализ и найден прообраз 16-раундовой версии функции сжатия.

Лучшей теоретической атакой по поиску коллизий является работа [5], где рассмотрена возможность использования противником квантового компьютера, верхняя граница сложности атаки для нахождения коллизии 38 раундов SHA-256 — 2^{116} вызовов функции сжатия. Лучший практический результат по поиску коллизий был представлен в 2024 году на конференции Eurocrypt [6], где продемонстрировано нахождение коллизии для 31-раундовой SHA-256.

В настоящей работе приведены описание принципа работы SHA-256 и краткий аналитический обзор атак. Представлены результаты применения логического криптоанализа для поиска прообразов и коллизий неполнораундовых версий SHA-256. Впервые найденные прообразы 17- и 18-раундовых функций сжатия SHA-256. При поиске коллизий для достижения большей эффективности продемонстрировано применение гибридной атаки, использующей дифференциальные пути и логический криптоанализ. Приведены ключевые

результаты проведенного исследования, также показаны направления дальнейших исследований.

Хеш-функция SHA-256

Описание и принцип работы. Хеш-функция SHA-256 принадлежит семейству SHA-2. Помимо нее в семейство также входят еще пять различных вариантов, полное описание которых можно найти в [1]. На вход принимается сообщение произвольной длины, выходом является хеш-значение, длина которого варьируется от 160 до 512 бит. Все хеш-функции семейства SHA-2 строятся на основе структуры Меркла–Дамгора [7, 8]. Входное сообщение дополняется до определенной длины, после чего разбивается на блоки заданного размера, затем происходит итеративная обработка этих блоков с применением функции сжатия. В зависимости от используемого алгоритма функция сжатия может состоять из 64 или 80 раундов.

Самой первой функцией SHA, созданной Агентством Национальной Безопасности США и Национальным институтом стандартов и технологий, была SHA-0. Она разрабатывалась в качестве стандарта для безопасной генерации и проверки цифровой подписи, а также в целях обеспечения надежного алгоритма хеширования для федеральных приложений. Однако первая открытая публикация SHA-0 была быстро отозвана из-за неназванного недостатка, и была выпущена обновленная версия алгоритма хеширования — SHA-1. Через несколько лет на смену SHA-1 пришла SHA-256.

В качестве входного сообщения SHA-256 принимает битовую строку произвольного размера (максимальная длина — $(2^{64} - 1)$ бит), которая после дополнения разбивается на N блоков по 512 бит, выходом является 256-битное значение на последнем этапе. Функция сжатия, являющаяся основой SHA-256, состоит из 64 раундов и работает с 512-битными блоками сообщения, которые представлены в виде 16 слов по 32 бита. Основные операции для работы со словами: логические операции сложения, умножения, XOR, отрицания; сложение по модулю 2^{32} ; циклический и нециклический битовые сдвиги.

Первым этапом функции сжатия является расширение 16-ти 32-битных входных слов m_i до 64 слов:

$$W_i \leftarrow \begin{cases} m_i, & 0 \leq i < 16 \\ \sigma_1(W_{i-2}) + W_{i-7} + \sigma_0(W_{i-15}) + W_{i-16}, & 16 \leq i < 64, \end{cases}$$

где функции σ работают следующим образом:

$$\sigma_0(x) \leftarrow (x \ggg 7) \oplus (x \ggg 18) \oplus (x \gg 3),$$

$$\sigma_1(x) \leftarrow (x \ggg 17) \oplus (x \ggg 19) \oplus (x \gg 10),$$

где \ggg и \gg — циклический и нециклический сдвиги вправо.

Основа работы функции сжатия с сообщениями W_i заключена в 8-ми 32-битных инициализационных значениях $H_{(i)}^0 (i = 0, \dots, 7)$ и восьми внутренних регистрах a, b, c, d, e, f, g, h . Изначально эти переменные равны инициализационным значениям $H_{(i)}^0$, на каждом из по-

следующих 64 раундов функции сжатия выполняются следующие операции:

$$\begin{aligned}h &= g, d = c, \\g &= f, c = b, \\f &= e, b = a, \\e &= d + T_1, a = T_1 + T_2,\end{aligned}$$

где

$$\begin{aligned}T_1 &= h + \Sigma_1(h) + Ch(e, f, g) + K_t + W_t, \\T_2 &= \Sigma_0(a) + Maj(a, b, c), \\Ch(x, y, z) &= (x \wedge y) \oplus (\bar{a} \wedge z), \\Maj(x, y, z) &= (x \wedge y) \oplus (y \wedge z) \oplus (x \wedge z), \\ \Sigma_1(x) &= (x \ggg 6) \oplus (x \ggg 11) \oplus (x \ggg 25), \\ \Sigma_0(x) &= (x \ggg 2) \oplus (x \ggg 13) \oplus (x \ggg 22),\end{aligned}$$

где K_t — раундовая константа; W_t — слово на раунде t , $t = 0, \dots, 63$. В конце каждого раунда происходит обновление промежуточных значений $H(i)$, $i = 0, \dots, N$.

Таким образом, на последнем 64-ом раунде, после вычисления значений регистров $a - h$ и обновления состояний $H^{(N)}$, финальное значение хеша равно конкатенации:

$$H_0^{(N)} \| H_1^{(N)} \| H_2^{(N)} \| H_3^{(N)} \| H_4^{(N)} \| H_5^{(N)} \| H_6^{(N)} \| H_7^{(N)}.$$

Методы атак на SHA-256. На рассматриваемую хеш-функцию существует ряд теоретических и практических атак. SHA-256, как и любая другая криптографическая хеш-функция, должна обладать определенными свойствами, которые необходимо проанализировать в первую очередь при доказательстве безопасности, а именно — стойкость к поиску первого и второго прообразов и коллизий.

- Поиск первого прообраза (односторонность): при наличии хеш-значения $H(m)$ сообщения m необходимо вычислить само сообщение m .
- Поиск второго прообраза: при наличии сообщения m_1 и его хеш-значения $H(m_1)$ необходимо найти другое сообщение m_2 , такое что $H(m_1) = H(m_2)$.
- Поиск коллизии: необходимо найти два различных сообщения $m_1 \neq m_2$, таких что их хеш-значения совпадают $H(m_1) = H(m_2)$.

На криптографические хеш-функции существует множество атак, но так или иначе все они являются вариациями поиска прообразов или коллизий с дополнительными требованиями или упрощениями. Так, при поиске псевдопрообразов или псевдоколлизий необходимо наложить условие, что атакующий может изменять значение вектора инициализации IV , а при поиске частичного прообраза атакующему будет достаточно восстановить часть сообщения [9, 10]. При реализации биклик-атаки (одна из разновидностей атак «встречи посередине») необходимо использовать структуру полного двудольного графа для того, чтобы увеличить количество атак «человека посередине». Еще одна уязвимость, присущая SHA-256, выявляется атаккой удлинением сообщения. Данной атаке подвержены

все итеративные конструкции, у которых финальное преобразование внутреннего состояния не отличается от всех предыдущих итераций. Зная длину исходного сообщения и его хеш-значение, злоумышленник может реализовать дополнительную $n + 1$ итерацию функции сжатия, подставив хеш-значение исходного сообщения и дополнив его своим сообщением с дополнением (падингом). После проведения всех манипуляций хеш-функция выдаст валидное новое хеш-значение. Данная атака работает, даже если атакующий не знает самого исходного сообщения.

Не все эти атаки являются общими и не могут применяться ко всем хеш-функциям. Для оценки безопасности любой криптографической хеш-функции, вне зависимости от ее конструкции, можно использовать различные виды криптоанализа, например, алгебраический, логический, дифференциальный или линейный криптоанализы.

В работе [11] предложена алгебраическая атака на основе сбоя для анализа безопасности шифров или хеш-функций. При таком подходе происходит объединение атак по сторонним каналам с алгебраическими методами — для начала нужно внести сбой в расчетах аппаратного устройства, после чего построить алгебраические уравнения для хеш-функции, а в качестве новых значений использовать новые данные с ошибками. После этого применяются автоматические инструменты для решения уравнений и восстановления необходимой информации. В работе [12] алгебраическая атака на основе сбоя совмещена с SAT-решателями для восстановления секретных битов из аппаратных реализаций семейства хеш-функций SHA-1 и SHA-2.

В настоящей работе были применены методы дифференциального и логического криптоанализов. В основе логического криптоанализа лежит проблема выполнимости булевых формул, записанных в конъюнктивной нормальной форме (КНФ).

Необходимо определить, принимает ли формула, записанная в КНФ, значение «истина», и если да, то найти значения переменных, входящих в нее. На практике это означает найти прообраз имеющегося хеш-значения (т. е. обратить хеш-функцию), либо найти несколько сообщений с одним и тем же хеш-значением (коллизия) — все зависит от формулировки и кодирования задачи.

Логический криптоанализ является универсальным средством криптоанализа хеш-функций. Он не позволяет получить теоретические оценки сложности, однако, дает возможность оценить сложность с практической точки зрения. Примером применения логического криптоанализа для поиска прообразов криптографических хеш-функций является работа [13], где использован SAT-решатель Kissat для анализа стойкости криптографических хеш-функций — финалистов конкурса National Institute of Standards and Technology. В работе [14] применен SAT-решатель совместно с системой компьютерной алгебры CAS для поиска коллизий с полусвободным стартом неполнораундовой версии SHA-256. Интересный подход с применением SAT-решателя представлен в работе [15], где были найдены так называемые «квантовые» коллизии (поиск

коллизий с помощью алгоритма Гровера) хеш-функций семейства SHA-3; ученые использовали решатели SAT для более эффективного поиска коллизий, а также для оптимизации подхода в общем.

Идея дифференциального криптоанализа выдвинута в работе [16], где был предложен оригинальный метод анализа, позволяющий на тот момент времени взломать неполнораундовый шифр DES. Суть дифференциального анализа заключается в исследовании зависимости, как небольшая разница между двумя входными сообщениями влияет на их выходные значения. В случае криптографических хеш-функций это помогает при поиске коллизий: необходимо найти два различных сообщения, которые дают один и тот же хеш, при условии наложения ограничения на разницу внутренних состояний на каждом раунде. В качестве примера применения дифференциального криптоанализа можно привести работы [17–19], где был представлен анализ хеш-функций SHA-0, MD5, MD4 и RIPEMD.

Оценка стойкости SHA-256. Ральф Меркл и Иван Дамгар на конференции CRYPTO в 1989 году предложили структуру, в рамках которой стойкость криптографической хеш-функции основывается на стойкости ее функции сжатия. Потому зачастую при оценке безопасности всей хеш-функции достаточно проанализировать стойкость функции сжатия в ее основе. Верхняя граница сложности поиска коллизий на полную версию SHA-256 — около 2^{128} операций вызова функции (полный перебор требует 2^{256} операций вызова, парадокс дней рождения позволяет сократить это количество). Верхняя оценка сложности поиска прообраза в худшем случае — полный перебор, т. е. 2^{256} операций вызова функции. С помощью различных методов и атак можно понизить верхние оценки сложности, однако, в общем случае сделать это затруднительно, а такие оценки для SHA-256 сегодня носят чисто теоретический характер. Это означает, что, хотя с помощью различных атак сложность поиска прообраза и коллизий снижается в сравнении с методом полного перебора, она все еще остается за пределами вычислительных возможностей современных компьютеров.

Атаки нахождения прообраза. Первая атака на поиск первого и второго прообразов SHA-256 основывалась на технике «встречи посередине»; авторы реализовали атаку на урезанную версию, состоящую из 24 раундов, и смогли снизить сложность до 2^{240} вызовов функции сжатия [20]. Впоследствии было проведено множество других атак и их вариаций для поиска прообраза или псевдопрообраза. В работе [21] предложена теоретическая вариация атаки «встречи посередине» для нахождения прообраза 42-раундовой версии SHA-256 со сложностью $2^{248,4}$ вызовов функции сжатия по времени. В [4] применен логический криптоанализ и найден прообраз 16-раундовой версии функции сжатия. Последней успешной на сегодняшний день работой является [3], в которой предложены атаки нахождения прообраза 45-раундовой версии SHA-256 и псевдопрообраза 52-раундовой версии SHA-256. В обоих случаях это были теоретические биклик-атаки по двудольному графу: сложность была снижена до $2^{255,5}$ и 2^{255} вызовов функции сжатия соответственно. Отметим,

что в настоящий момент большинство атак на поиск прообраза SHA-256 являются теоретическими.

Атаки нахождения коллизий. На практике получить коллизии хеш-функции SHA-256 впервые удалось в работе [22], в результате были получены коллизии для 18 раундов. В дальнейшем подходы к поиску коллизий были улучшены. В [23] была реализована вариация атаки «встречи посередине», в результате которой удалось найти коллизию для 28-раундовой функции сжатия и коллизию с полусвободным стартом для 31-раундовой функции сжатия SHA-256 со сложностью $2^{65,5}$ по времени. В работе [6] использован инструмент на основе SAT/SMT для эффективного поиска дифференциальных характеристик SHA-256; в итоге были найдены коллизии для 31 раунда (сложность по времени — $2^{49,5}$ операций вызова); псевдоколлизии для 52 раундов (сложность по времени — $2^{127,5}$ операций вызова) и коллизии с полусвободным стартом для 39 раундов (сложность по времени зависит от реализации атаки и начальных условий, например, если рассматривать только одну стартовую точку, то сложность по времени будет равна $2^{115,6}$ операциям вызова). Также в 2024 году в работе [14] был использован логический криптоанализ для нахождения коллизий с полусвободным стартом для 38-раундовой версии SHA-256.

Поиск прообразов неполнораундовых версий функции сжатия SHA-256

Применим метод логического криптоанализа для поиска прообразов усеченных (неполнораундовых) версий функции сжатия SHA-256. Для проведения экспериментов была выделена функция сжатия SHA-256 и построены ее SAT-кодировки, т. е. были сгенерированы соответствующие КНФ (четыре случая). Характеристики КНФ и время нахождения прообразов представлены в табл. 1, 2. Во всех случаях задачей являлось нахождение прообраза нулевого выхода функции сжатия, т. е. выхода, состоящего из 256 нулевых битов. Все эксперименты проводились на персональном компьютере со следующими характеристиками: 16-ядерный процессор AMD Ryzen 3950X, оперативная память — 64 ГБ.

При проведении экспериментов для получения КНФ использовались следующие программы: C Bounded Model Checker (CBMC), Transalg и SAT-encoding.

Транслятор CBMC [24] — транслятор общего назначения, предназначенный для преобразования программ на языках C и C++ в булевы формулы, которые затем анализируются для выявления ошибок и нарушения корректности исполнения. Процесс создания SAT-кодировок криптографических хеш-функций с помощью CBMC подробно описан в работе [13]. Транслятор Transalg [25] — специализированный транслятор, предназначенный для описания криптографических алгоритмов на предметно-ориентированном языке и формирующий файл с соответствующей КНФ. SAT-encoding — специальная программа, доступная в открытом репозитории [26], предназначенная для построения SAT-кодировок прикладных задач. SAT-encoding позволяет формировать SAT-кодировки функ-

ций сжатия криптографических хеш-функций MD4, SHA-1 и SHA-256. Программа также предоставляет возможность настройки дополнительных параметров: выбора алгоритма хеширования, типа кодирования целочисленного сложения, указания количества раундов для кодировки, определения цели генерации КНФ (поиск прообраза или коллизии), а также задания фиксированного размера входного сообщения (по умолчанию размер сообщения не фиксирован).

Для решения полученных КНФ применялся SAT-решатель Kissat [27], реализующий алгоритм CDCL [28].

В первом случае использована готовая программная реализация хеш-функции SHA-256 [1] на языке C. Для создания кодировки программный код был адаптирован: выделена функция сжатия, убран весь ненужный функционал, проверена корректность реализации и сформулирована задача поиска прообраза неполнораундовой версии функции сжатия. После этого файл с программным кодом подавался на вход транслятору CBMC для получения файла, содержащего КНФ. Таким образом, были сделаны КНФ для нахождения прообразов 16-, 17-, 18- и 19-раундовых функций сжатия SHA-256.

На построенных файлах был запущен SAT-решатель Kissat. Установлено ограничение по времени решения в 5000 с.

Во втором случае использован транслятор Transalg. Программа для функции сжатия SHA-256 доступна в репозитории этого транслятора [29]. Для экспериментов в этой программе варьировалось число раундов и задавался нулевой выход. Результаты, полученные с помощью Kissat, представлены в табл. 1.

В третьем случае была использована программа SAT-encoding. Для работы программы нужно задать вход и выход функции сжатия. При этом выход функ-

ции сжатия либо вводится вручную, либо автоматически генерируется случайным образом. Программа была модифицирована таким образом, чтобы была возможность задать нулевой выход.

В программе SAT-encoding предусмотрено использование различных алгоритмов кодирования в SAT целочисленного сложения: counter chain — использование счетчиков (суммирование битов операндов в одном столбце) в режиме каскадного суммирования, dot matrix — обработка и уменьшение матрицы операндов при помощи дерева Уоллеса, espresso — аналог counter chain с использованием логического минимизатора Espresso, two operand — одновременное внесение двух операндов. В программе SAT-encoding отсутствовала реализация Espresso, поэтому был подключен данный минимизатор в качестве внешней программы, взятой в открытом репозитории¹. После сборки и запуска модифицированной программы были получены КНФ для 16, 17, 18 и 19 раундов функции сжатия для каждого вида кодирования сложения — итого 16 КНФ. Характеристики КНФ и результаты работы Kissat на них представлены в табл. 2.

В результате выполненных экспериментов можно сделать вывод, что нахождение прообраза для 18-раундовой функции сжатия SHA-256 на сегодняшний день — простая задача, и быстрее всего SAT-решатель Kissat ее решил на кодировке, сделанной с помощью SAT-encoding с использованием Espresso.

Так как задача нахождения прообраза 19-раундовой функции сжатия SHA-256 оказалась слишком сложной для решателя Kissat, были рассмотрены ослабленные версии данной задачи. Ослабление заключалось в том, что был известен не весь 256-битный выход, а только его часть. В такой постановке у каждого выхода функции сжатия увеличивается число соответствующих прообразов, что упрощает поиск хотя бы одного из них. Была построена кодировка 19 раундов функции сжатия SHA-256 с помощью SAT-encoding и Espresso. Затем сделаны 32 КНФ: в первой известен только первый бит выхода и он равен нулю; во второй известны первые два нулевых бита и так далее. На каждой из этих КНФ запущен Kissat с ограничением времени в 5000 с. В итоге удалось найти прообраз максимум для 24-битного нулевого выхода. Для сравнения, если закодированы все 64 раунда функции сжатия SHA-256, то в тех же условиях Kissat находит решение максимум для КНФ, в которой заданы нулями первые 17 битов выхода.

Существуют разные подходы к распараллеливанию трудных SAT-задач. Некоторые из них основаны на варьировании значений множества переменных КНФ [30]. Для решения трудных задач был выбран другой подход, который называется Cube-and-Conquer. В рамках этого подхода на первом этапе с помощью lookahead-решателя исходная задача упрощается и разбивается на независимые подзадачи, а затем на втором этапе эти подзадачи решаются с помощью CDCL-решателя. На

Таблица 1. Характеристики КНФ, построенные с помощью CBMC и Transalg, а также время нахождения прообразов с помощью решателя Kissat

Table 1. Characteristics of CNF constructed using CBMC and Transalg as well as the time for finding preimages using the Kissat solver

Раунды	Переменные	Дизъюнкты	Время, с
CBMC			
16	35 136	148 609	0,05
17	35 966	152 848	0,25
18	36 794	157 080	12,53
19	37 622	161 312	не решено
Transalg			
16	22 212	143 478	0,06
17	22 774	147 154	0,36
18	23 334	150 823	5,76
19	23 894	154 492	не решено

Примечание. «не решено» — решатель не смог найти выполняющий набор за выделенное время.

¹ University of California/Espresso-logic-minimizer [Электронный ресурс]. Режим доступа: <https://github.com/classabby/espresso-logic>, свободный. Яз. англ. (дата обращения: 24.07.2024).

Таблица 2. Характеристики КНФ, построенные с помощью алгоритмов SAT-encoding, а также время нахождения прообразов с помощью решателя Kissat

Table 2. Characteristics of CNF constructed using algorithms SAT-encoding, as well as the time for finding preimages using the Kissat solver

Раунды	Перем.	Дизъюнкты	Время, с	Перем.	Дизъюнкты	Время, с
	Алгоритм Counter chain			Алгоритм Dot matrix		
16	11 936	70 504	0,04	10 760	66 088	0,05
17	12 918	76 651	0,14	11 604	71 727	0,13
18	13 900	82 798	7,52	12 448	77 366	8,89
19	14 882	88 945	не решено	13 292	83 005	не решено
Раунды	Алгоритм Two operand			Алгоритм Espresso		
16	11 400	66 536	0,02	7920	14 336	0,01
17	12 286	72 210	0,14	8495	140 442	0,76
18	13 172	77 884	20,60	9070	151 508	2,38
19	14 058	83 558	не решено	9645	162 574	не решено

Примечание. «не решено» — решатель не смог найти выполняющий набор за выделенное время; «Перем.» — число переменных.

8-ми 19-раундовых КНФ был запущен параллельный SAT-решатель EnCnC, реализующий метод Cube-and-Conquer [31]. Решатель был запущен с ограничением времени в 24 часа на 16-ядерном процессоре. На втором этапе использован Kissat, т. е. он был запущен на соответствующих подзадачах. В итоге для всех КНФ, кроме самой сложной (с 32-битным нулевым выходом), удалось найти решение. Все полученные КНФ доступны онлайн [32].

Можно сделать вывод, что нахождение прообраза для 19 раундов функции сжатия SHA-256 является слишком сложной задачей даже если тщательно подобрать SAT-кодировку и использовать параллельные вычисления. Отметим также, что по сравнению с работой [4], в которой был найден прообраз максимум для 16 раундов функции сжатия SHA-256 — достигнут явный прогресс.

Поиск коллизий для неполнораундовых версий функции сжатия SHA-256

Объединим методы и подходы логического криптоанализа для поиска коллизий в неполнораундовых версиях функции сжатия SHA-256. Выполним проверку эффективности совмещения двух предложенных подходов.

Задача поиска коллизий в данном случае формулируется следующим образом. Пусть M и M' — произвольные сообщения, $h(\cdot)$ — функция сжатия. Необходимо найти такие $M \neq M'$ что $H(M) = H(M')$.

Сначала применим только логический криптоанализ для поиска коллизий. Для этого закодируем поставленную задачу следующим образом. Зададим два произвольных входных сообщения M_1 и M_2 при этом наложим на них ограничения таким образом, чтобы они были различны. Также укажем, что выход функции сжатия после применения к обоим сообщениям будет одинаков: $H(M_1) = H(M_2)$. После чего попробуем решить задачу поиска коллизий с помощью Kissat.

Задача относительно легко решается для 18 раундов функции сжатия — поиск коллизии занял 34 с. Для 19 раундов задачу данным методом решить не удалось. Для более эффективного поиска коллизий можно попробовать объединить логический и дифференциальный криптоанализы. На каждом раунде используются 8 регистров: a, b, c, d, e, f, g, h , где каждый имеет размерность 32 бит, при этом выходом функции сжатия являются эти же регистры. Чтобы сократить пространство возможных сообщений для поиска коллизий, можно наложить ограничения на разность части регистров на некоторых раундах: таким образом можно построить дифференциальный путь. Например, для регистра a дифференциальный путь в общем виде будет выглядеть следующим образом:

$$\Delta a_{r_1} \rightarrow \Delta a_{r_2} \rightarrow \dots \rightarrow \Delta a_{r_{n-1}} \rightarrow \Delta a_{r_n},$$

где $\{r_1, \dots, r_n\}$ — раунды (всего n); на последнем раунде $\Delta a_{r_n} = 0$, так как выход регистра должен совпадать для сообщений M и M' . Таким образом, ограничения на значения регистров позволяют уменьшить сложность перебора. Однако следует отметить, что сам поиск нужных ограничений — нетривиальная задача.

Для проверки зададим единственное значение дифференциального пути $\Delta a_{r_1} = 1$, сократим число раундов до 18, закодируем задачу поиска коллизии, получим КНФ с помощью СВМС и попробуем применить Kissat для решения задачи. Задача решается за 19 с. Если задано единственное значение дифференциального пути $\Delta e_{r_1} = 1$, то Kissat находит решение за 3 с. Таким образом, путем добавления всего лишь одного условия удалось достичь значительной оптимизации.

Важно заметить, что это не всегда так. В некоторых случаях, напротив, ограничения могут приводить к ухудшению производительности, а в некоторых коллизии найти и вовсе не удастся. Иногда совместное применение ограничений может привести к моментальному UNSAT, что означает — для таких ограничений

Таблица 3. Коллизия для 19-раундовой функции сжатия SHA-256

Table 3. Collision for the 19-round SHA-256 compression function

0x71646011	0x5479452	0x3791dfb5	0x1d1338de
0x9900a4c7	0x60745f6b	0x1e0ac5de	0x9b286266
0xef7f4755	0x65ad1cba	0xab32b24	0x9dc00b20
0xb6f13681	0xcd20048c	0xdc967d4e	0xb1e8b4b
0x71646011	0x5479452	0x3791dfb5	0xf51f13c9
0xd4c569f1	0xb9e206a2	0x5df3b728	0xfad4f80e
0xd46a748	0x65ad1cba	0xab32b24	0xc5b43035
0xb6f13681	0xcd20048c	0xdc967d4e	0xb1e8b4b

Примечание. Совпадающие 32-битные слова в сообщениях выделены полужирным шрифтом.

формула невыполнима, и они не могут быть использованы. Например, если совместно задать $\Delta a_{r_1} = 1$ и $\Delta e_{r_1} = 5$ (для 19-раундовой функции сжатия), Kissat выдает UNSAT (время выполнения 0,06 с), что доказывает — коллизий при указанных двух ограничениях не существует.

Для проверки корректности подхода закодируем все дифференциальные пути из работы [33] для 19-раундовой функции сжатия SHA-256. Kissat решает данную задачу за 2 мин 2 с. Найденная коллизия представлена в табл. 3.

Для 20-раундовой функции сжатия с теми же дифференциальными путями коллизию найти не удалось. Отметим, что полученный результат уступает существующим научным работам, например, в [23] была найдена коллизия для 28-раундовой версии SHA-256. Эффективное нахождение дифференциальных путей для полнораундовой версии хеш-функции остается на сегодняшний день открытой проблемой.

Заключение

В работе приведен логический криптоанализ алгоритма криптографической хеш-функции SHA-256. Исследована сама хеш-функция и описаны возможные атаки. В качестве эксперимента для нахождения прообразов и коллизий были использованы различные трансляторы для получения конъюнктивных нормальных форм и SAT-решатели для решения проблемы булевой выполнимости. В результате работы были найдены прообразы и коллизии 18-раундовой функции сжатия SHA-256. Изучена техника совместного использования дифференциального и логического криптоанализов. В результате для 18-раундовой функции сжатия удалось значительно ускорить время выполнения логического криптоанализа, а также получилось найти коллизии для 19-раундовой функции сжатия. В рамках дальнейших исследований планируется усовершенствовать предложенные в работе методы поиска прообразов и коллизий, а также изучить дополнительные подходы к криптоанализу алгоритма SHA-256 для достижения лучших результатов.

Литература

- Secure hash standard (shs) // Fips pub. 2012. V. 180. N 4.
- Alamgir N. Programmatic SAT for SHA-256 Collision Attack. 2024 [Электронный ресурс]. URL: <https://scholar.uwindsor.ca/etd/9525> (дата обращения: 12.07.2024).
- Khovratovich D., Rechberger C., Savelieva A. Bicliques for preimages: attacks on Skein-512 and the SHA-2 family // *Lecture Notes in Computer Science*. 2012. V. 7549. P. 244–263. https://doi.org/10.1007/978-3-642-34047-5_15
- Homsirikamol E., Morawiecki P., Rogawski M., Srebrny M. Security margin evaluation of SHA-3 contest finalists through SAT-based attacks // *Lecture Notes in Computer Science*. 2012. V. 7564. P. 56–67. https://doi.org/10.1007/978-3-642-33260-9_4
- Hosoyamada A., Sasaki Y. Quantum collision attacks on reduced SHA-256 and SHA-512 // *Lecture Notes in Computer Science*. 2021. V. 12825. P. 616–646. https://doi.org/10.1007/978-3-030-84242-0_22
- Li Y., Liu F., Wang G. New records in collision attacks on SHA-2 // *Lecture Notes in Computer Science*. 2024. V. 14651. P. 158–186. https://doi.org/10.1007/978-3-031-58716-0_6
- Damgard I.B. A design principle for hash functions // *Lecture Notes in Computer Science*. 1990. P. V. 435. P. 416–427. https://doi.org/10.1007/0-387-34805-0_39
- Merkle R.C. A certified digital signature // *Lecture Notes in Computer Science*. 1990. V. 435. P. 218–238. https://doi.org/10.1007/0-387-34805-0_21
- Al-Kuwari S., Davenport J.H., Bradford R. J. Cryptographic hash functions: Recent design trends and security notions // *Proc. of the 6th*

References

- Secure hash standard (shs). *Fips pub*, 2012, vol. 180, no. 4.
- Alamgir N. *Programmatic SAT for SHA-256 Collision Attack*. 2024 Available at: <https://scholar.uwindsor.ca/etd/9525> (accessed: 12.07.2024).
- Khovratovich D., Rechberger C., Savelieva A. Bicliques for preimages: attacks on Skein-512 and the SHA-2 family. *Lecture Notes in Computer Science*. 2012, vol. 7549, pp. 244–263. https://doi.org/10.1007/978-3-642-34047-5_15
- Homsirikamol E., Morawiecki P., Rogawski M., Srebrny M. Security margin evaluation of SHA-3 contest finalists through SAT-based attacks. *Lecture Notes in Computer Science*, 2012, vol. 7564, pp. 56–67. https://doi.org/10.1007/978-3-642-33260-9_4
- Hosoyamada A., Sasaki Y. Quantum collision attacks on reduced SHA-256 and SHA-512. *Lecture Notes in Computer Science*, 2021, vol. 12825, pp. 616–646. https://doi.org/10.1007/978-3-030-84242-0_22
- Li Y., Liu F., Wang G. New records in collision attacks on SHA-2. *Lecture Notes in Computer Science*, 2024, vol. 14651, pp. 158–186. https://doi.org/10.1007/978-3-031-58716-0_6
- Damgard I.B. A design principle for hash functions. *Lecture Notes in Computer Science*, 1990, vol. 435, pp. 416–427. https://doi.org/10.1007/0-387-34805-0_39
- Merkle R.C. A certified digital signature. *Lecture Notes in Computer Science*, 1990, vol. 435, pp. 218–238. https://doi.org/10.1007/0-387-34805-0_21
- Al-Kuwari S., Davenport J.H., Bradford R. J. Cryptographic hash functions: Recent design trends and security notions. *Proc. of the 6th*

- China International Conference on Information Security and Cryptology (Inscrypt '10), 2010. P. 133–150.
10. Gauravaram P. Cryptographic Hash Functions: Cryptanalysis, Design and Applications. PhD thesis. Queensland University of Technology, 2007. 298 p.
 11. Courtois N.T., Jackson K., Ware D. Fault-algebraic attacks on inner rounds of DES // *Proc. of the E-Smart'10*, 2010. P. 22–24.
 12. Nejati S., Horacek J., Gebotys C., Ganesh V. Algebraic fault attack on sha hash functions using programmatic SAT solvers // *Lecture Notes in Computer Science*, 2018. V. 11008. P. 737–754. https://doi.org/10.1007/978-3-319-98334-9_47
 13. Заикин О.С., Давыдов В.В., Кирьянова А.П. Применение алгоритмов решения проблемы булевой выполнимости для анализа финалистов конкурса SHA-3 // *Вычислительные методы и программирование*, 2024. Т. 25. С. 259–273. <https://doi.org/10.26089/NumMet.v25r320>
 14. Alamgir N., Nejati S., Bright C. SHA-256 collision attack with programmatic SAT // *CEUR Workshop Proceedings*, 2024. V. 3717. P. 91–110.
 15. Guo J., Liu G., Song L., Tu Y. Exploring SAT for cryptanalysis:(Quantum) collision attacks against 6-round SHA-3 // *Lecture Notes in Computer Science*, 2022. V. 13793. P. 645–674. https://doi.org/10.1007/978-3-031-22969-5_22
 16. Biham E., Shamir A. Differential cryptanalysis of DES-like cryptosystems // *Journal of Cryptology*, 1991. V. 4. N 1. P. 3–72. <https://doi.org/10.1007/BF00630563>
 17. Wang X., Yu H. How to break MD5 and other hash functions // *Lecture Notes in Computer Science*, 2005. V. 3494. P. 19–35. https://doi.org/10.1007/11426639_2
 18. Wang X., Lai X., Feng D., Chen H., Yu X. Cryptanalysis of the hash functions MD4 and RIPEMD // *Lecture Notes in Computer Science*, 2005. V. 3494. P. 1–18. https://doi.org/10.1007/11426639_1
 19. Wang X., Yu H., Yin Y.L. Efficient collision search attacks on SHA-0 // *Lecture Notes in Computer Science*, 2005. V. 3621. P. 1–16. https://doi.org/10.1007/11535218_1
 20. Isobe T., Shibutani K. Preimage attacks on reduced Tiger and SHA-2 // *Lecture Notes in Computer Science*, 2009. V. 5665. P. 139–155. https://doi.org/10.1007/978-3-642-03317-9_9
 21. Guo J., Ling S., Rechberger C., Wang H. Advanced meet-in-the-middle preimage attacks: First results on full Tiger, and improved results on MD4 and SHA-2 // *Lecture Notes in Computer Science*, 2010. V. 6477. P. 56–75. https://doi.org/10.1007/978-3-642-17373-8_4
 22. Mendel F., Pramstaller N., Rechberger C., Rijmen V. Analysis of step-reduced sha-256 // *Lecture Notes in Computer Science*, 2006. V. 4047. P. 126–143. https://doi.org/10.1007/11799313_9
 23. Mendel F., Nad T., Schl  ffer M. Improving local collisions: New attacks on reduced SHA-256 // *Lecture Notes in Computer Science*, 2013. V. 7881. P. 262–278. https://doi.org/10.1007/978-3-642-38348-9_16
 24. Clarke E., Kroening D., Lerda F. A tool for checking ANSI-C programs // *Lecture Notes in Computer Science*, 2004. V. 2988. P. 168–176. https://doi.org/10.1007/978-3-540-24730-2_15
 25. Semenov A., Otpuschennikov I., Griбанова I., Zaikin O., Kochemazov S. Translation of algorithmic descriptions of discrete functions to SAT with applications to cryptanalysis problems // *Logical Methods in Computer Science*, 2020. V. 16. N 1. P. 29. [https://doi.org/10.23638/LMCS-16\(1:29\)2020](https://doi.org/10.23638/LMCS-16(1:29)2020)
 26. Nejati S. SAT Encoding [Электронный ресурс]. URL: <https://github.com/saeednj/SAT-encoding> (дата обращения: 12.07.2024).
 27. Biere A. The Kissat SAT Solver [Электронный ресурс]. URL: <https://github.com/arminbiere/kissat.git> (дата обращения: 12.07.2024).
 28. Marques-Silva J.P., Sakallah K.A. GRASP: A search algorithm for propositional satisfiability // *IEEE Transactions on Computers*, 1999. V. 48. N 5. P. 506–521. <https://doi.org/10.1109/12.769433>
 29. Otpuschennikov I. Programs for SHA-256 [Электронный ресурс]. URL: https://gitlab.com/satencodings/satencodings/-/tree/master/SHA2?ref_type=heads (дата обращения: 12.07.2024).
 30. Semenov A., Zaikin O., Kochemazov S. Finding effective SAT partitionings via black-box optimization // *Springer Optimization and Its Applications*, 2021. V. 170. P. 319–355. https://doi.org/10.1007/978-3-030-66515-9_11
 31. Zaikin O. Inverting cryptographic hash functions via Cube-and-Conquer // *Journal of Artificial Intelligence Research*, 2024. V. 81. P. 359–399. <https://doi.org/10.1613/jair.1.15244>
 32. Заикин О. КНФ для SHA-256 [Электронный ресурс]. URL: <https://github.com/olegzaikin/sha256sat.git> (дата обращения: 20.02.2025).
 33. China International Conference on Information Security and Cryptology (Inscrypt '10), 2010, pp. 133–150.
 34. Gauravaram P. Cryptographic Hash Functions: Cryptanalysis, Design and Applications. PhD thesis. Queensland University of Technology, 2007, 298 p.
 35. Courtois N.T., Jackson K., Ware D. Fault-algebraic attacks on inner rounds of DES. *Proc. of the E-Smart'10*, 2010, pp. 22–24.
 36. Nejati S., Horacek J., Gebotys C., Ganesh V. Algebraic fault attack on sha hash functions using programmatic SAT solvers. *Lecture Notes in Computer Science*, 2018, vol. 11008, pp. 737–754. https://doi.org/10.1007/978-3-319-98334-9_47
 37. Zaikin O.S., Davydov V.V., Kiryanova A.P. SAT-based analysis of SHA-3 competition finalists. *Numerical Methods and Programming*, 2024, vol. 25, pp. 259–273. (in Russian). <https://doi.org/10.26089/NumMet.v25r320>
 38. Alamgir N., Nejati S., Bright C. SHA-256 collision attack with programmatic SAT. *CEUR Workshop Proceedings*, 2024, vol. 3717, pp. 91–110.
 39. Guo J., Liu G., Song L., Tu Y. Exploring SAT for cryptanalysis:(Quantum) collision attacks against 6-round SHA-3. *Lecture Notes in Computer Science*, 2022, vol. 13793, pp. 645–674. https://doi.org/10.1007/978-3-031-22969-5_22
 40. Biham E., Shamir A. Differential collision search attacks on SHA-0. *Journal of Cryptology*, 1991, vol. 4, no. 1, pp. 3–72. <https://doi.org/10.1007/BF00630563>
 41. Wang X., Yu H. How to break MD5 and other hash functions. *Lecture Notes in Computer Science*, 2005, vol. 3494, pp. 19–35. https://doi.org/10.1007/11426639_2
 42. Wang X., Lai X., Feng D., Chen H., Yu X. Cryptanalysis of the hash functions MD4 and RIPEMD. *Lecture Notes in Computer Science*, 2005, vol. 3494, pp. 1–18. https://doi.org/10.1007/11426639_1
 43. Wang X., Yu H., Yin Y.L. Efficient collision search attacks on SHA-0. *Lecture Notes in Computer Science*, 2005, vol. 3621, pp. 1–16. https://doi.org/10.1007/11535218_1
 44. Isobe T., Shibutani K. Preimage attacks on reduced Tiger and SHA-2. *Lecture Notes in Computer Science*, 2009, vol. 5665, pp. 139–155. https://doi.org/10.1007/978-3-642-03317-9_9
 45. Guo J., Ling S., Rechberger C., Wang H. Advanced meet-in-the-middle preimage attacks: First results on full Tiger, and improved results on MD4 and SHA-2. *Lecture Notes in Computer Science*, 2010, vol. 6477, pp. 56–75. https://doi.org/10.1007/978-3-642-17373-8_4
 46. Mendel F., Pramstaller N., Rechberger C., Rijmen V. Analysis of step-reduced sha-256. *Lecture Notes in Computer Science*, 2006, vol. 4047, pp. 126–143. https://doi.org/10.1007/11799313_9
 47. Mendel F., Nad T., Schl  ffer M. Improving local collisions: New attacks on reduced SHA-256. *Lecture Notes in Computer Science*, 2013, vol. 7881, pp. 262–278. https://doi.org/10.1007/978-3-642-38348-9_16
 48. Clarke E., Kroening D., Lerda F. A tool for checking ANSI-C programs. *Lecture Notes in Computer Science*, 2004, vol. 2988, pp. 168–176. https://doi.org/10.1007/978-3-540-24730-2_15
 49. Semenov A., Otpuschennikov I., Griбанова I., Zaikin O., Kochemazov S. Translation of algorithmic descriptions of discrete functions to SAT with applications to cryptanalysis problems. *Logical Methods in Computer Science*, 2020, vol. 16, no. 1, pp. 29. [https://doi.org/10.23638/LMCS-16\(1:29\)2020](https://doi.org/10.23638/LMCS-16(1:29)2020)
 50. Nejati S. SAT Encoding. Available at: <https://github.com/saeednj/SAT-encoding> (accessed: 12.07.2024).
 51. Biere A. The Kissat SAT Solver. Available at: <https://github.com/arminbiere/kissat.git> (accessed: 12.07.2024).
 52. Marques-Silva J.P., Sakallah K.A. GRASP: A search algorithm for propositional satisfiability. *IEEE Transactions on Computers*, 1999, vol. 48, no. 5, pp. 506–521. <https://doi.org/10.1109/12.769433>
 53. Otpuschennikov I. Programs for SHA-256. Available at: https://gitlab.com/satencodings/satencodings/-/tree/master/SHA2?ref_type=heads (accessed: 12.07.2024).
 54. Semenov A., Zaikin O., Kochemazov S. Finding effective SAT partitionings via black-box optimization. *Springer Optimization and Its Applications*, 2021, vol. 170, pp. 319–355. https://doi.org/10.1007/978-3-030-66515-9_11
 55. Zaikin O. Inverting cryptographic hash functions via Cube-and-Conquer. *Journal of Artificial Intelligence Research*, 2024, vol. 81, pp. 359–399. <https://doi.org/10.1613/jair.1.15244>
 56. Заикин О. CNF for SHA-256. Available at: <https://github.com/olegzaikin/sha256sat.git> (accessed: 20.02.2025). (in Russian)

33. Sanadhya S.K., Sarkar P. Attacking reduced round SHA-256 // *Lecture Notes in Computer Science*, 2008. V. 5037. P. 130–143. https://doi.org/10.1007/978-3-540-68914-0_8

33. Sanadhya S.K., Sarkar P. Attacking reduced round SHA-256. *Lecture Notes in Computer Science*, 2008, vol. 5037, pp. 130–143. https://doi.org/10.1007/978-3-540-68914-0_8

Авторы

Давыдов Вадим Валерьевич — кандидат технических наук, криптограф-исследователь, ООО «КуАпп», Москва, 121205, Российская Федерация; доцент, Санкт-Петербургский государственный университет аэрокосмического приборостроения, 190000, Санкт-Петербург, Российская Федерация; научный сотрудник, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, [sc 57203909696](https://orcid.org/0000-0002-5544-2434), <https://orcid.org/0000-0002-5544-2434>, vadimdavydov@outlook.com

Пихтовников Михаил Денисович — студент, Южный федеральный университет, Таганрог, 347922, Российская Федерация, <https://orcid.org/0009-0000-2173-0844>, pikhtovnikov347@mail.ru

Кирьянова Анастасия Павловна — аспирант, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, <https://orcid.org/0009-0006-0344-5111>, anastaciakosanovskaya@gmail.com

Заикин Олег Сергеевич — кандидат технических наук, научный сотрудник, Новосибирский государственный университет, Новосибирск, 630090, Российская Федерация; ведущий научный сотрудник, Институт динамики систем и теории управления имени В.М. Матросова Сибирского отделения Российской академии наук, Иркутск, 664033, Российская Федерация, [sc 56786079600](https://orcid.org/0000-0002-0145-5010), <https://orcid.org/0000-0002-0145-5010>, oleg.zaikin@icc.ru

Authors

Vadim V. Davydov — PhD, Cryptography Researcher, QApp, Moscow, 121205, Russian Federation; Associate Professor, Saint Petersburg State University of Aerospace Instrumentation (SUAI), 190000, Saint Petersburg, Russian Federation; Scientific Researcher, ITMO University, 197101, Saint Petersburg, Russian Federation, [sc 57203909696](https://orcid.org/0000-0002-5544-2434), <https://orcid.org/0000-0002-5544-2434>, vadimdavydov@outlook.com

Michail D. Pikhtovnikov — Student, Southern Federal University, Taganrog, 347922, Russian Federation, <https://orcid.org/0009-0000-2173-0844>, pikhtovnikov347@mail.ru

Anastasia P. Kiryanova — PhD Student, ITMO University, 197101, Saint Petersburg, Russian Federation, <https://orcid.org/0009-0006-0344-5111>, anastaciakosanovskaya@gmail.com

Oleg S. Zaikin — PhD, Researcher, Novosibirsk State University, Novosibirsk, 630090, Russian Federation; Leading Researcher, Matrosov Institute for System Dynamics and Control Theory of Siberian Branch of RAS, Irkutsk, 664033, Russian Federation, [sc 56786079600](https://orcid.org/0000-0002-0145-5010), <https://orcid.org/0000-0002-0145-5010>, oleg.zaikin@icc.ru

*Статья поступила в редакцию 18.02.2025
Одобрена после рецензирования 21.04.2025
Принята к печати 27.05.2025*

*Received 18.02.2025
Approved after reviewing 21.04.2025
Accepted 27.05.2025*



Работа доступна по лицензии
Creative Commons
«Attribution-NonCommercial»

doi: 10.17586/2226-1494-2025-25-3-438-445

УДК 004.852

Исследование возможности применения эволюционных алгоритмов для условной генерации атрибутированных графов

Ирина Юрьевна Деева¹✉, Полина Олеговна Андреева², Егор Николаевич Шиков³,
Анна Владимировна Калюжная⁴

^{1,2,3,4} Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация

¹ iriny.deeva@gmail.com✉, <https://orcid.org/0000-0001-8679-5868>

² polinaspb@ya.ru, <https://orcid.org/0000-0002-3130-1012>

³ egorshikov@itmo.ru, <https://orcid.org/0000-0001-5749-4222>

⁴ kalyuzhnaya.ann@gmail.com, <https://orcid.org/0000-0002-9612-8601>

Аннотация

Введение. Область синтетической генерации атрибутированных графов активно развивается благодаря прогрессу в генеративном моделировании. Однако ключевой проблемой современных методов остается ограниченное разнообразие синтезируемых графов, которое зависит от характеристик реальных данных, используемых для обучения генеративных моделей. Также ограничение связано с топологическими свойствами графов и статистическими характеристиками атрибутов, критически влияющими на эффективность графовых моделей машинного обучения. В данной работе проверена гипотеза о том, что комбинация эволюционных алгоритмов и байесовских сетей может обеспечить гибкий контроль над генерацией как топологии, так и атрибутов графа. **Метод.** Предложенный подход включает два ключевых компонента: эволюционные алгоритмы для управления топологическими характеристиками графа (например, средняя степень вершины, коэффициент кластеризации) и байесовские сети для генерации атрибутов с заданными статистическими параметрами, такими как ассортативность или средняя корреляция между атрибутами. Метод позволяет явно задавать ограничения на свойства графа, обеспечивая вариативность, не зависящую от исходных данных. **Основные результаты.** Эксперименты подтвердили, что подход способен генерировать атрибутированные графы с широким спектром топологических характеристик и заданными статистическими параметрами атрибутов с достаточно низкой ошибкой генерации. **Обсуждение.** Результаты демонстрируют перспективность использования эволюционных и байесовских методов для условной генерации графов. Основное преимущество подхода — возможность декомпозиции задачи на независимое управление топологией и атрибутами, что открывает новые возможности для тестирования алгоритмов машинного обучения в контролируемых условиях. Ограничением является вычислительная сложность эволюционной оптимизации, что требует дальнейшей работы по оптимизации алгоритма. В перспективе метод может быть расширен для генерации динамических графов и интеграции с глубокими генеративными моделями.

Ключевые слова

атрибутированные графы, генерация синтетических графов, эволюционная оптимизация, байесовские сети, топологические характеристики графов, статистические характеристики атрибутов

Благодарности

Исследование выполнено при финансовой поддержке Российского научного фонда, договор № 24-71-10093, <https://rscf.ru/en/project/24-71-10093/>.

Ссылка для цитирования: Деева И.Ю., Андреева П.О., Шиков Е.Н., Калюжная А.В. Исследование возможности применения эволюционных алгоритмов для условной генерации атрибутированных графов // Научно-технический вестник информационных технологий, механики и оптики. 2025. Т. 25, № 3. С. 438–445. doi: 10.17586/2226-1494-2025-25-3-438-445

Investigation of the possibility of using evolutionary algorithms for conditional generation of attributed graphs

Irina Yu. Deeva¹✉, Polina O. Andreeva², Egor N. Shikov³, Anna V. Kalyuzhnaya⁴

^{1,2,3,4} ITMO University, Saint Petersburg, 197101, Russian Federation

¹ iriny.deeva@gmail.com✉, <https://orcid.org/0000-0001-8679-5868>

² polinaspb@ya.ru, <https://orcid.org/0000-0002-3130-1012>

³ egorshikov@itmo.ru, <https://orcid.org/0000-0001-5749-4222>

⁴ kalyuzhnaya.ann@gmail.com, <https://orcid.org/0000-0002-9612-8601>

Abstract

The field of synthetic generation of attributed graphs is actively developing due to advances in generative modeling. However, a key problem of current methods remains the limited diversity of synthesized graphs, due to the dependence on the characteristics of real data used to train generative models. This is a problem because topological properties of graphs and statistical characteristics of attributes critically affect the performance of graph-based machine learning models. In this paper, we test the hypothesis that a combination of evolutionary algorithms and Bayesian networks can provide flexible control over the generation of both graph topology and attributes. The proposed approach includes two key components: evolutionary algorithms to control topological characteristics of the graph (e.g. average vertex degree, clustering coefficient) and Bayesian networks to generate attributes with given statistical parameters such as assortativity or average correlation between attributes. The method allows explicitly setting constraints on graph properties, providing variability independent of the original data. Experiments confirmed that the approach can generate attributed graphs with a wide range of topological characteristics and given statistical parameters of the attributes with sufficiently low generation error. The results demonstrate the promising use of evolutionary and Bayesian methods for conditional graph generation. The main advantage of the approach is the ability to decompose the problem into independent control of topology and attributes, which opens new possibilities for testing machine learning algorithms under controlled conditions. A limitation is the computational complexity of evolutionary optimization, which requires further work to optimize the algorithm. In the future, the method can be extended to generate dynamic graphs and integrate with deep generative models.

Keywords

attributed graphs, synthetic graph generation, evolutionary optimization, Bayesian networks, topological characteristics of graphs, statistical characteristics of attributes

Acknowledgements

This research is financially supported by the Russian Scientific Foundation, Agreement 24-71-10093, <https://rscf.ru/en/project/24-71-10093/>.

For citation: Deeva I.Yu., Andreeva P.O., Shikov E.N., Kalyuzhnaya A.V. Investigation of the possibility of using evolutionary algorithms for conditional generation of attributed graphs. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2025, vol. 25, no. 3, pp. 438–445 (in Russian). doi: 10.17586/2226-1494-2025-25-3-438-445

Введение

Современное развитие методов машинного обучения на графовых данных тесно связано с качеством и разнообразием датасетов-бенчмарков, используемых для их тестирования и валидации. Как отмечено в [1, 2], большинство существующих архитектур графовых нейронных сетей испытываются на ограниченном наборе реальных графов, что снижает обобщающую способность полученных результатов. При этом реальные графы обычно не отражают полного спектра топологических и атрибутивных характеристик, способных существенно влиять на эффективность алгоритмов [3]. Данный факт создает следующую проблему: отсутствует универсальный генератор графов, способный воспроизводить как типичные, так и экстремальные случаи, необходимые для адекватной оценки устойчивости и производительности графовых моделей.

Одним из перспективных направлений решения данной проблемы является условная генерация графов, позволяющая создавать синтетические структуры с заранее заданными характеристиками. В последние годы было предложено множество параметрических генераторов, ориентированных на управление топологическими особенностями, такими как структура

сообществ [4–6] или распределение степеней по закону мощности [7, 8]. Однако данные методы часто ограничены своей способностью моделировать комплексные и многогранные особенности графов, поскольку требуют индивидуальной настройки под каждую характеристику. Аналогично, современные глубокие генеративные модели, включая диффузионные методы [9, 10], авто-регрессионные модели [11], VAE [12, 13] и GAN [14], преимущественно нацелены на воспроизведение обучающих данных, что ограничивает их применение при необходимости генерации редких или специально сконфигурированных структур.

Основная проблема, решаемая в настоящей работе, заключается в разработке методологии, способной объединить эволюционные алгоритмы для оптимизации топологических характеристик графов и байесовские сети (БС) для генерации атрибутов с заданными статистическими свойствами. Существующие подходы часто фокусируются на отдельном аспекте графа — на его структуре или на свойствах атрибутов — что приводит к несоответствию реальным сценариям применения, где необходимо учитывать сложное взаимодействие между структурой и атрибутами [4, 7]. Таким образом, задача состоит в создании гибкого и масштабируемого инструмента для условной генерации атрибути-

рованных графов, способного учитывать корреляции между атрибутами соседних вершин и обеспечивать воспроизводимость экстремальных случаев, имеющих практическое значение для тестирования графовых моделей [5, 6].

Для проверки гипотезы о перспективах условной генерации графов был разработан подход, основанный на эволюционных алгоритмах, оптимизирующих структуру графа посредством минимизации функции стоимости, зависящей от желаемых топологических характеристик. Дополнительно предложена интеграция БС для генерации атрибутов, что позволяет задавать и контролировать статистические зависимости между признаками вершин. Такой комбинированный подход позволяет обеспечить более глубокое соответствие синтетических графов реальным сценариям, где топология и атрибуты взаимодействуют сложным и взаимозависимым образом.

Основные понятия и обозначения

Графы. Граф $G = (V, E)$ можно определить как множества его вершин $V = \{v_1, \dots, v_n\}$ и ребер $E = \{(v_i, v_j)\}$, $v_i, v_j \in V$.

Предполагается, что каждая вершина имеет атрибуты в виде матрицы признаков $X \in \mathbb{R}^{n \times d}$, в которой i -ая строка соответствует вектору признаков x_i вершины v_i , где n — количество вершин в графе; d — размерность пространства признаков. Для простоты предполагается, что графы являются неориентированными, хотя те же методы с минимальными изменениями применимы и к направленным графам.

Топологические инварианты. Инвариантом графа G называется функция $I(G)$, которая характеризует структуру графа и не зависит от способа разметки узлов или образа графа.

Проведенный комплексный анализ свойств графовых моделей машинного обучения, а также результаты тематических научных работ выявляют несколько ключевых инвариантов, которые существенно влияют на результаты: средний коэффициент кластеризации I_c , средняя длина кратчайшего пути I_{asp} [15], средняя степень I_d [16], ассортативности лейблов I_{la} и атрибутов I_{aa} [16]. Однако важно отметить, что предлагаемый метод является адаптивным и может быть расширен для включения других характеристик графа, например степени узла, локальной кластеризации, и другие [17–19].

Выбранные ключевые инварианты обычно используются в сетевых науках, информацию о которых можно найти в известных исследованиях, за исключением атрибутивной ассортативности I_{aa} , которая в настоящей работе определяется аналогично ассортативности меток:

$$I_{aa}(G) = \frac{1}{n} \sum_{i=1}^n \frac{|\{v_j \in \mathcal{N}(v_i): \cos(x_i, x_j) < q\}|}{|\mathcal{N}(v_i)|},$$

где q — пороговое значение; \mathcal{N} — соседи вершины v_i .

Генерация условных графов. В данной работе задача генерации синтетических графов с заданными

характеристиками формулируется как многоцелевая оптимизационная задача:

$$\min_{G \sim p(G)} (\|I_1(G) - \tilde{I}_1\|, \|I_2(G) - \tilde{I}_2\|, \dots, \|I_k(G) - \tilde{I}_k\|),$$

где $\tilde{I}_1, \dots, \tilde{I}_k$ — заранее заданные значения инвариантов; G — граф. В процессе оптимизации графы выбираются из заданного распределения на пространстве графов $p(G)$ и подбирается экземпляр, наиболее подходящий для заданных значений.

Предлагаемый подход

Главная особенность разработанного метода — итеративная модификация графа, постепенно сходящаяся к заданным топологическим характеристикам, предоставляемым в качестве входных данных. Процесс генерации состоит из двух этапов (на рис. 1 — шаги внутри обеих частей рисунка): *инициализации*, когда генерируется исходный случайный граф, и *оптимизации* — эволюционные операторы итеративно применяются для модификации графа. Эволюционные алгоритмы, используемые в представленном подходе, основаны на фреймворке GOLEM¹, который реализует генетический алгоритм NSGA-II [20]. Фреймворк позволяет определять собственные генетические операторы, а также достаточно гибко варьировать параметры генетического алгоритма. Общая схема подхода показана на рис. 1 и представляет собой генерацию синтетического графа и генерацию синтетических атрибутов для синтетического графа.

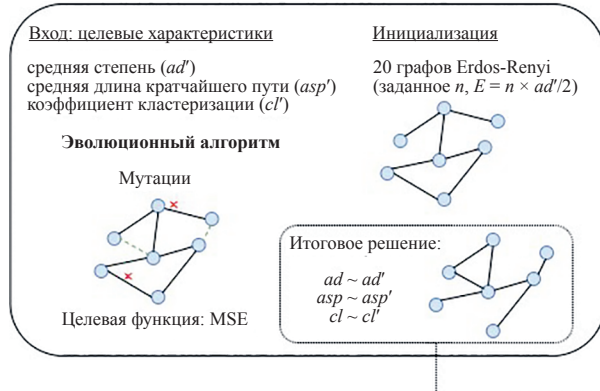
Условная генерация синтетических графов. Исходные графы строятся как случайные графы с помощью генератора Эрдоса–Реньи [21] с заданным числом вершин и числом ребер. В отличие от существующих подходов, где каждое новое изменение характеристик графа требует полной генерации нового образца графа, предлагаемый подход позволяет постепенно изменять существующие образцы. В качестве функции стоимости в данном контексте используется средняя квадратичная ошибка, которая измеряет несоответствие между фактическими характеристиками графа и заранее заданными целевыми значениями. Основная мутация, применяемая к графу, заключается в случайных изменениях ребра внутри графа и блока из пяти ребер.

Условная генерация синтетических атрибутов. Использование БС в качестве модели генерации атрибутов обусловлено двумя ключевыми факторами.

Во-первых, БС [22] способна учитывать зависимости между атрибутами, которые обычно встречаются в реальных данных. Во-вторых, графовая структура БС позволяет использовать тот же подход к оптимизации (как и для генерации графов) и тот же инструмент (фреймворк GOLEM). Важно отметить, что в рамках текущей работы рассматривается только генерация непрерывных атрибутов. Использование БС в качестве генеративной модели для атрибутов требует, чтобы структура и параметры распределений в ее узлах были

¹ [Электронный ресурс]. Режим доступа: <https://github.com/aimclub/GOLEM> (дата обращения: 01.04.2025).

Генерация синтетической структуры графа



Генерация синтетических атрибутов

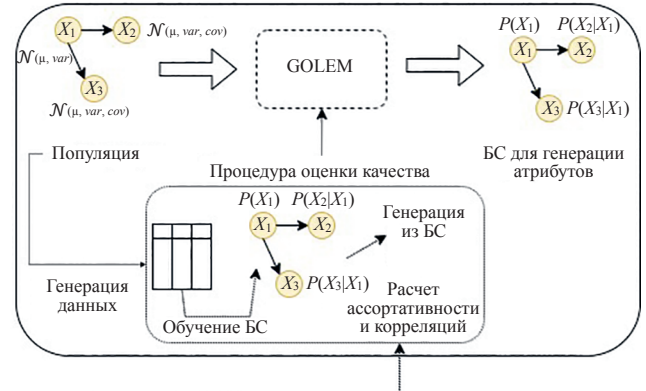


Рис. 1. Общая схема условной генерации атрибутированных графов с помощью эволюционных вычислений.

MSE — средняя квадратическая ошибка между целевыми характеристиками и полученными; X_1, X_2, X_3 — атрибуты; μ, var, cov — среднее, дисперсия и ковариация распределений атрибутов

Fig. 1. General scheme of conditional generation of attributed graphs using evolutionary computation, MSE — mean square error between target features and obtained ones, X_1, X_2, X_3 — attributes, μ, var, cov — mean, variance and covariance of attribute distributions

полностью заданы или обучены. В контексте данного исследования фиксируется структура сети, а параметры распределений определяются с помощью эволюционного подхода к обучению. В частности, распределения в узлах БС моделируются как гауссовы распределения. В процессе обучения вводятся мутации и кроссоверы для изменения параметров этих распределений в узлах. Процедура оценки включает в себя создание обучающей выборки из БС без ребер (только на основе маргинальных распределений) с использованием заранее заданных параметров распределений. Затем эта выборка используется для обучения БС с ребрами, из которой в последствии и семплируются синтетические атрибуты.

Экспериментальные результаты

Метрики качества. Для оценки результатов экспериментов предложено использовать следующие метрики. Оценку качества воспроизведения топологических характеристик предлагается измерять как абсолютную ошибку между значением на сгенерированном графе и целевым значением:

$$\varepsilon_{topo} = |F_{gen} - F_{target}|.$$

Чем меньше значение, тем точнее генерация.

Оценку качества генерируемых атрибутов производят с точки зрения двух метрик. Оценку ассортативности можно также выполнять с помощью расчета абсолютной ошибки. Для оценки корреляций генерируемых атрибутов используется евклидово расстояние между целевым вектором корреляций атрибутов и вектором на сгенерированных атрибутах:

$$\varepsilon_{corr} = \sqrt{\sum_{i=1}^d (Corr_{gen,i} - Corr_{target,i})^2}.$$

Условная генерация синтетических графов. В настоящем исследовании проведен анализ применения

эволюционных алгоритмов для генерации графов с заданными топологическими характеристиками. В качестве критериев оценки рассматривались следующие параметры: коэффициент кластеризации, варьирующийся в пределах от 0,1 до 0,45 с шагом 0,05; средняя длина кратчайших путей, принимающая значения 1,5, 2 и 2,5; а также средняя степень, изменяющаяся от 2 до 35 с шагом 5. Экспериментальная база формировалась с использованием графов, содержащих 20, 30, 40 и 50 узлов. Эффективность предложенного алгоритма оценивалась посредством сопоставления вычисленных характеристик графа с заранее заданными значениями, что позволило выявить степень соответствия синтезированных структур требуемым параметрам. Выполнен анализ вычислительной сложности алгоритма на основе времени обработки (рис. 2). Полученные данные демонстрируют приблизительно линейную зависимость между временем обработки (в секундах) и параметрами, такими как средняя степень и количество узлов, что

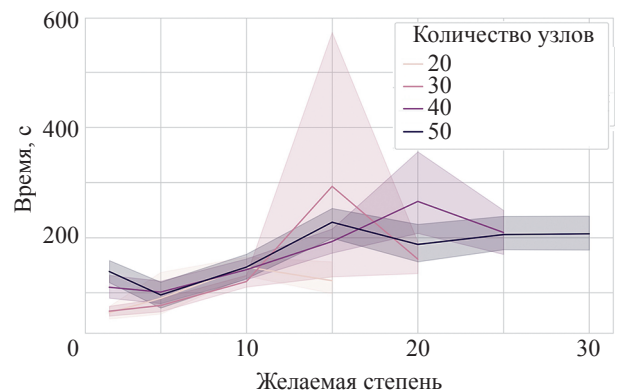


Рис. 2. Зависимость времени генерации от количества узлов и значения желаемой средней степени

Fig. 2. Dependence of generation time on number of nodes and value of desired average degree

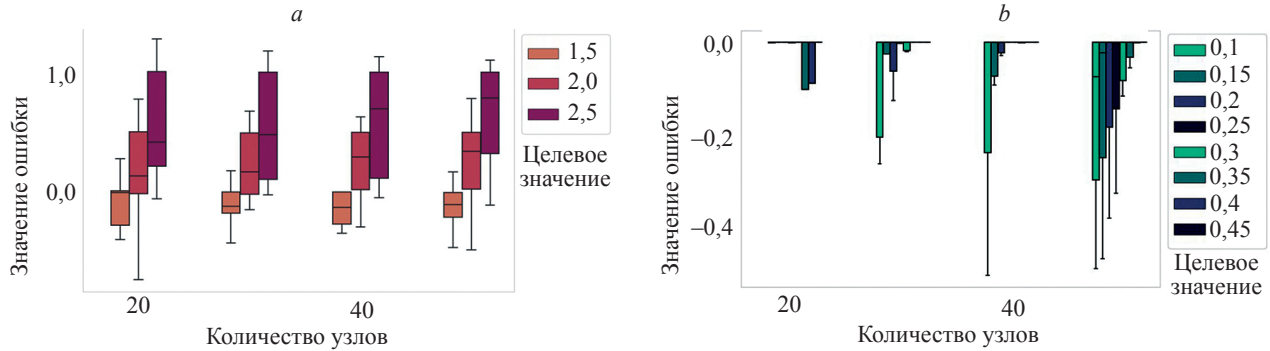


Рис. 3. Зависимости абсолютных ошибок характеристик сгенерированного графа от количества узлов при средней длине кратчайшего пути (a) и для коэффициента кластеризации (b)

Fig. 3. The dependence of the absolute errors of the characteristics of the generated graph vs. the number of nodes for the average length of the shortest path (a) and for the clustering coefficient (b)

свидетельствует о потенциальной масштабируемости предложенного метода.

На рис. 3 представлены различия между заданными и фактическими характеристиками сгенерированных графов. На рис. 3, a видно, что среднее значение ошибки остается на приемлемом уровне вне зависимости от размера графа, однако наблюдается увеличение дисперсии ошибки при больших значениях средней длины кратчайших путей. Рис. 3, b иллюстрирует тенденцию к систематическому занижению коэффициента кластеризации для графов, содержащих свыше 40 узлов, что свидетельствует о склонности алгоритма генерировать структуры с коэффициентами кластеризации, значительно меньшими, чем предусмотренные в параметрах моделирования.

На рис. 4 представлены ошибки в оценке средней длины кратчайшего пути и коэффициента кластеризации относительно целевой средней степени графа. На рис. 4, a видна тенденция к увеличению ошибки в определении средней длины кратчайшего пути для графов с целевой средней степенью, превышающей 5. Следует отметить, что наблюдается общее увеличение ошибки с ростом значения целевой средней степени, за исключением случая наименьшего значения, равного 2.

Это объясняется тем, что средняя степень 2 может быть достигнута только в циклическом графе, вследствие чего предложенный алгоритм обычно переоценивает этот показатель. Примечательно, что при увеличении числа вершин ошибка часто оказывается ниже. Рис. 4, b иллюстрирует ошибку в оценке коэффициента кластеризации по отношению к заданной средней степени. Из представленных данных следует, что ошибка возрастает с увеличением целевой степени, что характерно для всех рассматриваемых размеров графа. Наибольшая ошибка наблюдается при высоких значениях средней степени, что свидетельствует о повышенной точности алгоритма в аппроксимации заданного коэффициента кластеризации для более разреженных графов.

Условная генерация синтетических атрибутов.

В настоящем исследовании была поставлена задача выяснить, могут ли БС генерировать атрибуты с заданным уровнем ассоциативности и корреляции для полносвязных графов, в частности для 500 вершин. Для этого проведены эксперименты с различными конфигурациями, манипулируя количеством атрибутов и целевыми значениями ассоциативности и корреляции. Целевые значения ассортативности были установлены на 0,2 (низкая) и 0,9 (высокая). Для целевого значения

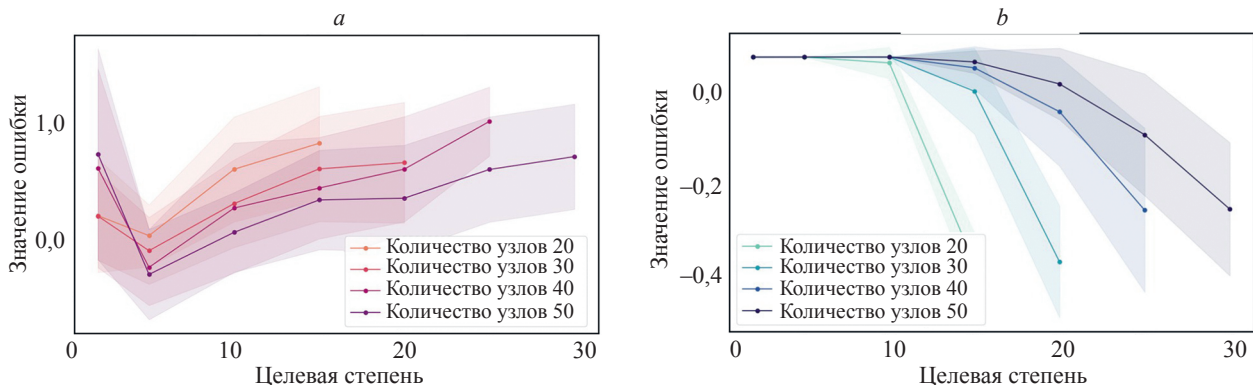


Рис. 4. Зависимости абсолютных ошибок характеристик сгенерированного графа от значений средней степени для различного количества вершин для длины кратчайшего пути (a) и коэффициента кластеризации (b)

Fig. 4. The dependence of the absolute errors of the characteristics of the generated graph vs. the values of the average degree for a different number of vertices for the length of the shortest path (a) and the clustering coefficient (b)

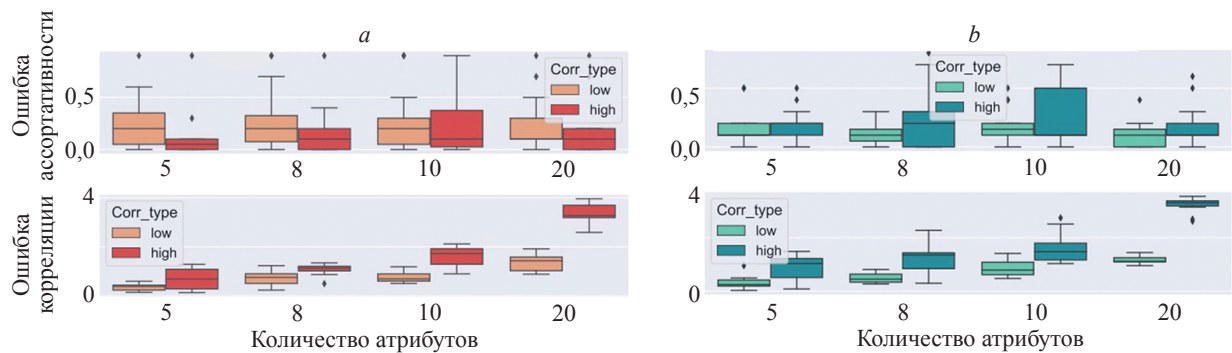


Рис. 5. Сравнение качества генерации атрибутов для различных экспериментальных конфигураций и количества атрибутов с точки зрения их близости к целевым характеристикам: высокая (0,9) (а) и низкая (0,2) (б) целевые ассортативности

Fig. 5. Comparison of the quality of attribute generation for various experimental configurations and the number of attributes in terms of their proximity to the target characteristics: high (0.9) (a) and low (0.2) (b) target assortativeness

корреляции (Corr_type) существует два режима: высокая корреляция (high), когда все значения корреляции для связанных атрибутов находятся в диапазоне от 0,7 до 0,9, и низкая корреляция (low), когда значения варьируются от 0,0 до 0,3. Структура БС всегда за-

дается как разреженная случайная структура с меньшим количеством ребер, чем узлов. На рис. 5 видно, что ошибка в ассортативности (абсолютная разница) в среднем не превышает 0,2 и остается стабильной для разного количества атрибутов. Это говорит о том, что

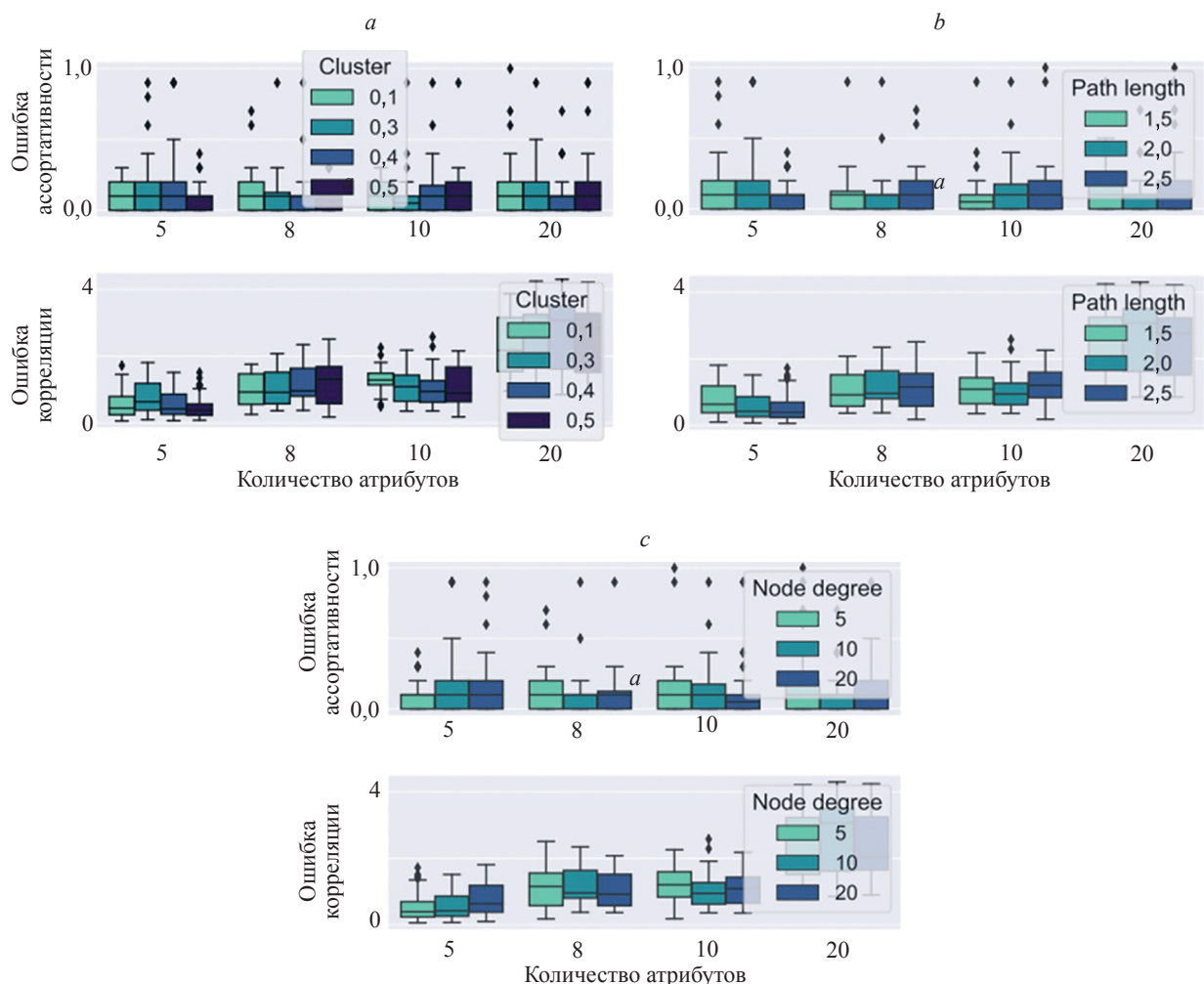


Рис. 6. Зависимости результатов генерации атрибутов от характеристик синтетического графа для различных коэффициентов кластеризации (а), длин кратчайших путей (б) и степеней вершин (с)

Fig. 6. The dependences of the attribute generation results vs. the characteristics of the synthetic graph for various clustering coefficients (a), shortest path lengths (b), and vertex degrees (c)

предложенный подход может генерировать атрибуты с необходимой ассортативностью. Важно отметить, что ошибка корреляции (евклидово расстояние между целевым вектором корреляций и полученным) увеличивается с ростом числа атрибутов, что ожидаемо в связи с ростом размера целевого вектора корреляции. Однако ошибка остается на приемлемом уровне.

В заключительной серии экспериментов были использованы синтетические графы, описанные в подразделе «Условная генерация синтетических атрибутов», обладающие разнообразными характеристиками: коэффициентом кластеризации, длиной кратчайших путей и степенью вершин. Для каждого из этих графов проведены эксперименты по генерации синтетических атрибутов с целевыми значениями ассортативности и корреляции. Структура БС формировалась случайным образом с варьированием степени плотности, а также изменялось число атрибутов. Рис. 6 демонстрирует, что БС способны эффективно генерировать атрибуты на синтетических графах, при этом качество исходных графовых структур не оказывает существенного влияния на данный процесс.

Литература

1. Palowitch J., Tsitsulin A., Mayer B., Perozzi B. Graphworld: Fake graphs bring real insights for GNNs // *Proc. of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*. 2022. P. 3691–3701. <https://doi.org/10.1145/3534678.3539203>
2. Hu W., Fey M., Zitnik M., Dong Y., Ren H., Liu B., Catasta M., Leskovec J. Open graph benchmark: Datasets for machine learning on graphs // *Advances in Neural Information Processing Systems*. 2020. V. 33. P. 22118–22133.
3. Shah N. Scale-free, attributed and class-assortative graph generation to facilitate introspection of graph neural networks // *Proc. of the MLG '20: ACM Symposium on Neural Gaze Detection*. 2020.
4. Maekawa S., Sasaki Y., Fletcher G., Onizuka M. Gencat: Generating attributed graphs with controlled relationships between classes, attributes, and topology // *Information Systems*. 2023. V. 115. P. 102195. <https://doi.org/10.1016/j.is.2023.102195>
5. Kim M., Leskovec J. Multiplicative attribute graph model of real-world networks // *Internet mathematics*. 2012. V. 8. N 1-2. P. 113–160. <https://doi.org/10.1080/15427951.2012.625257>
6. Tsitsulin A., Rozemberczki B., Palowitch J., Perozzi B. Synthetic graph generation to benchmark graph learning // *arXiv*. 2022. arXiv:2204.01376. <https://doi.org/10.48550/arXiv.2204.01376>
7. Largeron C., Mougél P.-N., Rabbany R., Zaïane O.R. Generating attributed networks with communities // *PLoS ONE*. 2015. V. 10. N 4. P. e0122777. <https://doi.org/10.1371/journal.pone.0122777>
8. Barabási A.-L. The Barabási-Albert Model // *Network Science*. 2016. P. 164–202.
9. Tseng A.M., Diamant N., Biancalani T., Scalia G. GraphGUIDE: interpretable and controllable conditional graph generation with discrete bernoulli diffusion // *arXiv*. 2023. arXiv:2302.03790. <https://doi.org/10.48550/arXiv.2302.03790>
10. Li M., Krecacic E., Potluru V.K., Li P. GraphMaker: Can diffusion models generate large attributed graphs? // *arXiv*. 2023. arXiv:2310.13833. <https://doi.org/10.48550/arXiv.2310.13833>
11. Faez F., Dijuij N.H., Baghshah M.S., Rabiee H.R. SCGG: A deep structure-conditioned graph generative model // *PLoS ONE*. 2022. V. 17. N 11. P. e0277887. <https://doi.org/10.1371/journal.pone.0277887>
12. Kipf T.N., Welling M. Variational graph auto-encoders // *arXiv*. 2016. arXiv:1611.07308. <https://doi.org/10.48550/arXiv.1611.07308>
13. Simonovsky M., Komodakis N. Graphvae: Towards generation of small graphs using variational autoencoders // *Lecture Notes in Computer Science*. 2018. V. 11139. P. 412–422. https://doi.org/10.1007/978-3-030-01418-6_41

Заключение

В данной работе исследована возможность условной генерации синтетических графов с использованием эволюционного алгоритма. Экспериментальные результаты подтвердили эффективность предложенного подхода в создании графов с заданными топологическими характеристиками и статистическими свойствами атрибутов. Анализ показал, что алгоритм способен воспроизводить целевые параметры, включая коэффициент кластеризации, среднюю длину кратчайшего пути и среднюю степень вершин, однако его точность снижается в более плотных графах. Это может быть связано с недостижимостью некоторых областей гиперпространства характеристик, что требует дальнейшего изучения. Предложенный метод открывает новые возможности для управляемой генерации графов и может быть полезен в исследованиях графовых нейронных сетей. Дальнейшие исследования могут быть направлены на повышение точности алгоритма и расширение его применимости.

References

1. Palowitch J., Tsitsulin A., Mayer B., Perozzi B. Graphworld: Fake graphs bring real insights for GNNs. *Proc. of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, 2022, pp. 3691–3701. <https://doi.org/10.1145/3534678.3539203>
2. Hu W., Fey M., Zitnik M., Dong Y., Ren H., Liu B., Catasta M., Leskovec J. Open graph benchmark: Datasets for machine learning on graphs. *Advances in Neural Information Processing Systems*, 2020, vol. 33, pp. 22118–22133.
3. Shah N. Scale-free, attributed and class-assortative graph generation to facilitate introspection of graph neural networks. *Proc. of the MLG '20: ACM Symposium on Neural Gaze Detection*, 2020.
4. Maekawa S., Sasaki Y., Fletcher G., Onizuka M. Gencat: Generating attributed graphs with controlled relationships between classes, attributes, and topology. *Information Systems*, 2023, vol. 115, pp. 102195. <https://doi.org/10.1016/j.is.2023.102195>
5. Kim M., Leskovec J. Multiplicative attribute graph model of real-world networks. *Internet mathematics*, 2012, vol. 8, no. 1–2, pp. 113–160. <https://doi.org/10.1080/15427951.2012.625257>
6. Tsitsulin A., Rozemberczki B., Palowitch J., Perozzi B. Synthetic graph generation to benchmark graph learning. *arXiv*, 2022, arXiv:2204.01376. <https://doi.org/10.48550/arXiv.2204.01376>
7. Largeron C., Mougél P.-N., Rabbany R., Zaïane O.R. Generating attributed networks with communities. *PLoS ONE*, 2015, vol. 10, no. 4, pp. e0122777. <https://doi.org/10.1371/journal.pone.0122777>
8. Barabási A.-L. The Barabási-Albert Model. *Network Science*, 2016, pp. 164–202.
9. Tseng A.M., Diamant N., Biancalani T., Scalia G. GraphGUIDE: interpretable and controllable conditional graph generation with discrete bernoulli diffusion. *arXiv*, 2023, arXiv:2302.03790. <https://doi.org/10.48550/arXiv.2302.03790>
10. Li M., Krecacic E., Potluru V.K., Li P. GraphMaker: Can diffusion models generate large attributed graphs? *arXiv*, 2023, arXiv:2310.13833. <https://doi.org/10.48550/arXiv.2310.13833>
11. Faez F., Dijuij N.H., Baghshah M.S., Rabiee H.R. SCGG: A deep structure-conditioned graph generative model. *PLoS ONE*, 2022, vol. 17, no. 11, P. e0277887. <https://doi.org/10.1371/journal.pone.0277887>
12. Kipf T.N., Welling M. Variational graph auto-encoders. *arXiv*, 2016, arXiv:1611.07308. <https://doi.org/10.48550/arXiv.1611.07308>
13. Simonovsky M., Komodakis N. Graphvae: Towards generation of small graphs using variational autoencoders. *Lecture Notes in Computer Science*, 2018, vol. 11139, pp. 412–422. https://doi.org/10.1007/978-3-030-01418-6_41

14. Bojchevski A., Shchur O., Zügner D., Günnemann S. NetGAN: Generating graphs via random walks // *arXiv*. 2018. arXiv:1803.00816. <https://doi.org/10.48550/arXiv.1803.00816>
15. Gasteiger J., Bojchevski A., Günnemann S. Predict then propagate: Graph neural networks meet personalized pagerank // *arXiv*. 2018. arXiv:1810.05997. <https://doi.org/10.48550/arXiv.1810.05997>
16. Kim D., Oh A. How to find your friendly neighborhood: Graph attention design with self-supervision // *arXiv*. 2022. arXiv:2204.04879. <https://doi.org/10.48550/arXiv.2204.04879>
17. Attar N., Aliakbari S., Nezhad Z.H. Automatic generation of adaptive network models based on similarity to the desired complex network // *Physica A: Statistical Mechanics and its Applications*. 2020. V. 545. P. 123353. <https://doi.org/10.1016/j.physa.2019.123353>
18. Verstraaten M., Varbanescu A.L., de Laat C. Synthetic graph generation for systematic exploration of graph structural properties // *Lecture Notes in Computer Science*. 2017. V. 10104. P. 557–570. https://doi.org/10.1007/978-3-319-58943-5_45
19. Barry A., Griffith J., O’Riordan C. An evolutionary and graph-rewriting based approach to graph generation // *Proc. of the 7th International Joint Conference on Computational Intelligence (IJCCI 2015)*. 2015. V. 1. P. 237–243. <https://doi.org/10.5220/0005597102370243>
20. Deb K., Pratap A., Agarwal S., Meyarivan T. A fast and elitist multiobjective genetic algorithm: NSGA-II // *IEEE Transactions on Evolutionary Computation*. 2002. V. 6. N 2. P. 182–197. <https://doi.org/10.1109/4235.996017>
21. Erdos P., Renyi A. On the evolution of random graphs // *The Structure and Dynamics of Networks*. 2011. P. 38–82.
22. Koller D., Friedman N. *Probabilistic Graphical Models: Principles and Techniques*. The MIT Press, 2009. 1231 p.

Авторы

Деева Ирина Юрьевна — кандидат физико-математических наук, старший научный сотрудник, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, [sc 57210416999](https://orcid.org/0000-0001-8679-5868), <https://orcid.org/0000-0001-8679-5868>, iriny.deeva@gmail.com

Андреева Полина Олеговна — аспирант, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, [sc 57223144672](https://orcid.org/0000-0002-3130-1012), <https://orcid.org/0000-0002-3130-1012>, polinaspb@ya.ru

Шиков Егор Николаевич — кандидат технических наук, младший научный сотрудник, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, [sc 57209451139](https://orcid.org/0000-0001-5749-4222), <https://orcid.org/0000-0001-5749-4222>, egorshikov@itmo.ru

Калюжная Анна Владимировна — кандидат технических наук, старший научный сотрудник, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, [sc 56218089100](https://orcid.org/0000-0002-9612-8601), <https://orcid.org/0000-0002-9612-8601>, kalyuzhnaya.ann@gmail.com

Authors

Irina Yu. Deeva — PhD (Physics & Mathematics), Senior Researcher, ITMO University, Saint Petersburg, 197101, Russian Federation, [sc 57210416999](https://orcid.org/0000-0001-8679-5868), <https://orcid.org/0000-0001-8679-5868>, iriny.deeva@gmail.com

Polina O. Andreeva — PhD Student, ITMO University, Saint Petersburg, 197101, Russian Federation, [sc 57223144672](https://orcid.org/0000-0002-3130-1012), <https://orcid.org/0000-0002-3130-1012>, polinaspb@ya.ru

Egor N. Shikov — PhD, Junior Researcher, ITMO University, Saint Petersburg, 197101, Russian Federation, [sc 57209451139](https://orcid.org/0000-0001-5749-4222), <https://orcid.org/0000-0001-5749-4222>, egorshikov@itmo.ru

Anna V. Kalyuzhnaya — PhD, Senior Researcher, ITMO University, Saint Petersburg, 197101, Russian Federation, [sc 56218089100](https://orcid.org/0000-0002-9612-8601), <https://orcid.org/0000-0002-9612-8601>, kalyuzhnaya.ann@gmail.com

Статья поступила в редакцию 01.03.2025
Одобрена после рецензирования 15.04.2025
Принята к печати 23.05.2025

Received 01.03.2025
Approved after reviewing 15.04.2025
Accepted 23.05.2025



Работа доступна по лицензии
Creative Commons
«Attribution-NonCommercial»

doi: 10.17586/2226-1494-2025-25-3-446-456

УДК 004.056.55

Анализ применимости существующих схем разделения секрета в условиях постквантовой эры

Елизар Филаретович Кустов¹✉, Сергей Валентинович Беззатеев²

^{1,2} Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация

² Санкт-Петербургский государственный университет аэрокосмического приборостроения, Санкт-Петербург, 190000, Российская Федерация

¹ elizarkustov@mail.ru✉, <https://orcid.org/0000-0002-0191-1178>

² sergey.bezzateev@gmail.com, <https://orcid.org/0000-0002-0924-6221>

Аннотация

Рассмотрены современные подходы к разделению секрета, включая как классические, так и постквантовые криптографические схемы. Исследованы методы распределения секретной информации между несколькими участниками с использованием математических примитивов, таких как многочлены Лагранжа и Ньютона, китайская теорема об остатках, коды исправляющие ошибки, теория решеток, изогении эллиптических кривых, многомерные уравнения и хэш-функции. Приведен сравнительный анализ различных схем с точки зрения их устойчивости к квантовым атакам, эффективности и соответствия критериям Шамира. Особое внимание уделено оценке устойчивости схем к атакам с использованием квантовых компьютеров, что особенно актуально в условиях развития квантовых технологий. Рассмотрены преимущества и недостатки каждой из схем, включая их вычислительную сложность, гибкость и возможность адаптации к различным условиям. Показано, что классические схемы, такие как схемы Шамира и Ньютона, остаются эффективными и простыми в реализации, но уязвимы к квантовым атакам. В то же время постквантовые схемы, основанные на теории решеток, демонстрируют высокий уровень безопасности, но требуют более сложных вычислений.

Ключевые слова

постквантовая криптография, схема распределение секрета, пороговая схема, криптография с открытым ключом, теория решеток, эллиптические кривые, многомерные уравнения, коды исправляющие ошибки, хэш-функции

Благодарности

Работа выполнена в рамках государственного задания (проект № FSER-2025-0003).

Ссылка для цитирования: Кустов Е.Ф., Беззатеев С.В. Анализ применимости существующих схем разделения секрета в условиях постквантовой эры // Научно-технический вестник информационных технологий, механики и оптики. 2025. Т. 25, № 3. С. 446–456. doi: 10.17586/2226-1494-2025-25-3-446-456

Analysis of the applicability of existing secret separation schemes in the post-quaternary era

Elizar F. Kustov¹✉, Sergey V. Bezzateev²

^{1,2} ITMO University, Saint Petersburg, 197101, Russian Federation

² Saint-Petersburg State University of Aerospace Instrumentation (SUAI), Saint Petersburg, 190000, Russian Federation

¹ elizarkustov@mail.ru✉, <https://orcid.org/0000-0002-0191-1178>

² sergey.bezzateev@gmail.com, <https://orcid.org/0000-0002-0924-6221>

Abstract

Modern approaches to secret sharing have been examined, encompassing both classical and post-quantum cryptographic schemes. The study explores methods for distributing secret information among multiple participants using various mathematical primitives, such as Lagrange and Newton polynomials, the Chinese remainder theorem, error-correcting

© Кустов Е.Ф., Беззатеев С.В., 2025

codes, lattice theory, elliptic curve isogenies, multivariate equations, and hash functions. A comparative analysis of different schemes is provided in terms of their resistance to quantum attacks, efficiency, and compliance with Shamir's criteria. Special attention is given to assessing the schemes resilience against attacks using quantum computers, which is particularly relevant given the advancement of quantum technologies. The advantages and disadvantages of each scheme are discussed, including their computational complexity, flexibility, and adaptability to various conditions. It is shown that classical schemes, such as those by Shamir and Newton, remain efficient and easy to implement but are vulnerable to quantum attacks. Meanwhile, post-quantum schemes based on lattice theory demonstrate a high level of security but require more complex computations.

Keywords

post-quantum cryptography, secret sharing scheme, threshold scheme, public-key cryptography, lattice theory, elliptic curves, multivariate equations, error-correcting codes, hash functions

Acknowledgments

The work was carried out within the framework of the State Assignment (project No. FSER-2025-0003).

For citation: Kustov E.F., Bezzateev S.V. Analysis of the applicability of existing secret separation schemes in the post-quantum era. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2025, vol. 25, no. 3, pp. 446–456 (in Russian). doi: 10.17586/2226-1494-2025-25-3-446-456

Введение

Первый метод разделения секрета был предложен еще в 1979 году Шамиром [1] и Блейкли [2] независимо друг от друга. Схемы разделения секрета применяются в случаях, когда существует значимая вероятность компрометации хранителей секрета, но вероятность недобросовестного сговора значительной части участников считается пренебрежимо малой.

Появление алгоритма Шора для квантового компьютера, поставило под угрозу классические криптосистемы, основанные на проблемах факторизации числа и дискретного логарифма. Возник вопрос о создании новых систем, устойчивых к атаке с использованием квантового компьютера.

Исследование в этой области дало толчок к развитию криптографии на новых постквантовых криптопримитивах, таких как коды исправляющие ошибки, теория решеток, изогении эллиптических кривых, многомерные уравнения и хэш-функции. Активное развитие этой области создало множество новых систем, в частности схемы разделения секрета. В связи с этим возник вопрос о возможной замене классических схем Шамира, Ньютона и Блейкли, которые сами по себе не обеспечивают безопасность от квантовых атак, на новые постквантовые схемы.

В настоящей работе выполнен анализ и сравнение между собой новых постквантовых и классических схем разделения секрета для поиска оптимальных схем. Сравнение происходит по критериям Шамира: размер доли меньше или равен самому секрету, возможность повторного использования секрета, невозможность проанализировать секрет, возможность добавления нового участника, возможность обновить секрет, возможность изменения веса долей.

В результате проведенного анализа различных схем представлена схема разделения секрета, отвечающая наибольшему числу критериев.

Схема разделения секрета по интерполяционному многочлену Лагранжа

Схема разделения секрета по интерполяционному многочлену Лагранжа, также известная как схема

Шамира [1], является одной из самых популярных пороговых схем разделения секрета. Она основана на интерполяции многочленов [3] и позволяет разделить секрет между n участниками так, что для его восстановления требуется как минимум k долей ($k \leq n$).

В этой схеме секретная информация, представленная целым числом S , распределяется между n участниками и может быть восстановлена при наличии как минимум k долей. Доверенный центр, называемый дилером, выбирает простое число $p > S$ и генерирует многочлен $q(x)$ степени $k - 1$ над $GF(p)$:

$$q(x) = S + a_1x + a_2x^2 + \dots + a_{k-1}x^{k-1},$$

где a_1, \dots, a_{k-1} выбираются дилером случайным образом из равномерно распределенных элементов $GF(p)$. Участник P_i ($i = 1, \dots, n$) получает только целое число $D_i = q(i)$ в качестве доли, где i — уникальный идентификатор участника P_i ; $q(i)$ — значение многочлена в точке i . При наличии k долей можно восстановить $q(x)$ с использованием интерполяции Лагранжа и, следовательно, восстановить секрет $S = q(0)$.

В схеме по интерполяционному многочлену Лагранжа, при наличии $k - 1$ долей для любого секрета S' , невозможно восстановить соответствующий ему многочлен. Таким образом, $k - 1$ долей не дают никакой информации о секрете.

Злоумышленник не сможет восстановить секрет S при условии, что он не имеет доступа к k или более долям, так как многочлен степени $k - 1$ однозначно определяется k точками. Схема устойчива к атакам, основанным на переборе, благодаря использованию конечного поля $GF(p)$ при выборе большого значения p . Также благодаря использованию интерполяционного многочлена Лагранжа можно добавить нового участника.

В табл. 1 представлены значения размеров ключей и долей секрета для 128-битной и 256-битной безопасности для схемы Шамира. В качестве ключей были использованы RSA (Rivest, Shamir, Adleman) из рекомендации NIST [4]. 128-битная и 256-битная безопасность — уровни криптографической стойкости, которые показывают, насколько сложно взломать систему перебором или другими атаками и означают, что для

Таблица 1. Сравнение размера ключевых элементов схемы Шамира, бит

Table 1. Comparison of the key elements size of the Shamir scheme, bit

Параметр	128-битная безопасность	256-битная безопасность
Секрет S	3072	15 360
Доля D_i	3072	15 360
Коэффициенты	$3072 \times (k - 1)$	$15\,360 \times (k - 1)$

взлома потребуется в среднем 2^{128} и 2^{256} операций соответственно.

Как видно из табл. 1 размер секрета и его долей равны, что позволяет использовать доли секрета в качестве секрета в криптосхемах без дополнительных модификаций. Отметим, что наиболее затратным для хранения данных, является размер коэффициентов многочлена.

Подчеркнем, что рассматриваемая схема не обеспечивает безопасность самого секрета S и его долей D_i , для этого сегодня используются вариации схемы Шамира с криптосистемами RSA [5], ECDCA [6] и Эль-Гамаль [7]. В этом варианте реализации общая криптосистема становится уязвима к атаке с применением квантового компьютера, в отличие от криптосистем, основанных на постквантовых криптопримитивах. Также к недостаткам схемы можно отнести ненадежность дилера, так как предполагается, что все участники могут доверять выбранному центру, что не всегда верно.

Схема разделения секрета по интерполяционному многочлену Ньютона

Схема разделения секрета на основе интерполяционного многочлена Ньютона — метод разделения секрета, который использует интерполяцию многочлена Ньютона для восстановления секрета. Этот подход похож на схему Шамира, но вместо интерполяции Лагранжа применяется интерполяция Ньютона. Многочлен Ньютона степени $t - 1$ имеет вид:

$$f(x) = a_0 + a_1(x - x_1) + a_2(x - x_1)(x - x_2) + \dots + a_{t-1}(x - x_1)(x - x_2) \dots (x - x_{t-1}),$$

где a_0, a_1, \dots, a_{t-1} — коэффициенты; x_1, x_2, \dots, x_{t-1} — заранее выбранные точки.

В начале дилер выбирает x_1, x_2, \dots, x_n , где n — количество участников. Для каждого участника P_i вычисляется значение многочлена в точке x_i , т. е. $f(x_i)$. Это значение является долей участника.

Для восстановления секрета необходимо собрать как минимум t долей. Используя интерполяцию Ньютона, строится многочлен $f(x)$ по t точкам $(x_i, f(x_i))$. Секрет S восстанавливается как значение многочлена в точке $x = 0$, т. е. $S = f(0)$.

Схема разделения секрета на основе многочлена Ньютона обладает теми же свойствами, что и схема разделения секрета по Лагранжу. Это значит, что она также уязвима к атаке с использованием квантового компьютера, если не добавить в схему посткванто-

вые криптопримитивы. Размеры ключей схемы подобны значениям, что и в схеме разделения секрета по Лагранжу, представленные в табл. 1.

Схема разделения секрета на основе китайской теоремы об остатках

Схема распределения секрета на основе китайской теоремы об остатках [8] — метод разделения секрета, который использует свойства модульной арифметики, и китайскую теорему об остатках для распределения и восстановления секрета. Этот подход позволяет разделить секрет между несколькими участниками так, чтобы только определенное подмножество участников могло его восстановить.

Как и в схеме Шамира, в работе [8] присутствует доверенный дилер, который определяет общий секрет S и попарно взаимно простые числа m_1, \dots, m_n , такие что:

$$m_{n-k+2} \times m_{n-k+3} \times \dots \times m_n < S < m_1 \times m_2 \times \dots \times m_k, \\ m_1 < m_2 < \dots < m_n,$$

где k — пороговое значение.

Каждому участнику выдается значение $S_i = S \bmod m_i$. Для восстановления секрета требуется собрать хотя бы k значений S_i и решить систему сравнений с помощью китайской теоремы об остатках.

Размеры ключей схемы подобны значениям, что и в схеме разделения секрета по Лагранжу, представленные в табл. 1. Отметим, что наиболее затратным является хранение модулей $m_i > 2^{256}$ бит для 128-битной и 256-битной безопасности.

К преимуществам данной схемы можно отнести ее эффективность по сравнению со схемой на основе интерполяционной формулы Лагранжа, так как восстановление секрета требуется только решение системы линейных сравнений, что вычислительно проще, чем интерполяция многочлена в схеме Шамира. Также, выбирая новые модули, можно добавить новых участников.

Важно, что, как и в схеме Шамира, схема на основе китайской теоремы об остатках не обеспечивает безопасность секрета и его долей, что приводит к ее уязвимости к квантовым атакам при использовании классических криптопримитивов.

Схемы разделения секрета, основанные на кодах, исправляющих ошибки

В [9] доказано, что схема Шамира является частным случаем обмена секретами с использованием кодов Рида–Соломона. Позже Мэсси [10] обнаружил взаимосвязь между структурой доступа для совместного использования секретов и кодовыми словами двояственного кода. Построение схем совместного использования секретов с применением кодовых конструкций остается важной темой и в настоящее время. Некоторые исследователи продолжают разрабатывать новые схемы обмена секретами [11–13].

В работе [9] было замечено, что схема разделения секрета Шамира является частным слу-

чаем более общей схемы, основанной на кодах Рида–Соломона. В этой схеме принято, что $(\alpha_1, \dots, \alpha_{r-1})$ — фиксированный список ненулевых элементов некоторого поля \mathbf{F} ($|\mathbf{F}| = r$, $r > n$), информационное слово $a = (S, a_1, \dots, a_{k-1}) \in \mathbf{F}_k$ кодируется в кодовое слово $D = (D_0, \dots, D_{r-2})$ с использованием кода Рида–Соломона:

$$D_i = \sum_{j=0}^{k-1} a_j \alpha_i^j, a_0 = S,$$

где D_i — доля участника P_i . При наличии всех r долей, чтобы восстановить секрет необходимо вычислить:

$$S = - \sum_{i=0}^{r-1} D_i.$$

При наличии s долей ($r \geq s \geq k$), из которых t содержат ошибки $\left(t \leq \frac{s-k}{2}\right)$, можно восстановить кодовое слово D с использованием алгоритма исправления ошибок и, следовательно, информационное слово a и секрет S .

В табл. 2 представлены значения размеров ключей и долей секрета для 128-битной и 256-битной безопасности. В качестве секрета был использован секретный ключ RSA.

Как видно из табл. 2 для хранения всех значений требуется достаточно большой объем памяти.

Данная схема обладает всеми свойствами схемы Шамира, а также новым свойством устойчивости к искажениям или деградации хранилища.

В работе [10] Мэсси показал, что минимальные кодовые слова в дуальном коде полностью определяют структуру схемы разделения секрета, и наоборот. Мэсси также определил, что структура схемы разделения секрета, соответствующая линейному (n, k) коду \mathbf{V} , задается теми минимальными кодовыми словами в дуальном коде, у которых первый компонент равен 1.

В [14] Мэсси расширил ранее полученные результаты, показав, что каждое кодовое слово q -ичного линейного кода может быть записано как линейная комбинация минимальных кодовых слов, которые его покрывают.

Таблица 2. Сравнение размера ключевых элементов схемы, основанной на кодах с исправлением ошибок, бит

Table 2. Comparison of the key elements size of a scheme based on error-correcting codes, bit

Параметр	128-битная безопасность	256-битная безопасность
Кодовое слово D	786 432	7 864 320
Информационное слово a	685 056	6 881 280
Проверочная матрица \mathbf{H}	25 067 520	503 316 480

Примечание. Все параметры зависят от размерности кода Рида–Соломона (n, k) .

Схемы разделения секрета, основанные на теории решеток

Первая схема разделения секрета на основе решеток была предложена в [15], где использовалась задача нахождения ближайшего вектора (Closest Vector Problem, CVP). Современные схемы чаще опираются на задачи нахождения кратчайшего вектора (Shortest Vector Problem, SVP) [16] и обучения с ошибками (Learning With Error, LWE) [17], которые являются NP-трудными, что делает их пригодными для криптографических приложений.

В [18] была предложена пороговая схема разделения секрета, основанная на SVP. В этой схеме для восстановления секрета требуется не менее n_1 участников, где $n_1 = n_2$ (количество участников, получающих доли). Основная идея заключается в восстановлении базисной матрицы, которая необходима для вычисления секретного вектора.

Безопасность схемы основана на использовании криптографических хэш-функций. В работе [19] предложено использовать хэш-функции Айти [20], Любашевского и Миккианчио [21] а также Пейкerta и Розена [22], что делает схему устойчивой к атакам с применением квантового компьютера. Основным недостатком схемы является условие, что для восстановления секрета необходимо участие всех участников P_i со своими долями.

В [23] была предложена пороговая схема (n, n) , основанная на LWE. Обозначим участников как P_i , $1 \leq i \leq n-1$, а секретный вектор как $\mathbf{s} \in \mathbf{Z}_q^m$. Дилер выбирает модуль q , простое число p , генератор g группы $GF(p)$ и распределение ошибок $\chi \in \mathbf{Z}^q$. Для каждого участника дилер выбирает случайный вектор $\mathbf{a}_i \in \mathbf{Z}_q^m$ и ошибку $e_i \in \mathbf{Z}_q$. Доля секрета $D_i = (\mathbf{a}_i, b_i)$, где b_i вычисляется как:

$$b_i = \langle \mathbf{a}_i, \mathbf{s} \rangle + e_i, 1 \leq i \leq n-1,$$

где $\langle \mathbf{a}_i, \mathbf{s} \rangle$ — скалярное произведение. Дилер публикует $V_i = g^{S_i} = (g^{a_i}, g^{b_i})$. Для последнего участника P_n дилер вычисляет $e_n = -e_1 - e_2 - \dots - e_{n-1}$ и соответственно

$$b_n = \langle \mathbf{a}_n, \mathbf{s} \rangle + e_n.$$

В фазе восстановления участники предоставляют агрегатору свои доли D_i . Если все доли валидные, для восстановления секрета \mathbf{s} необходимо просуммировать доли:

$$D_1 + \dots + D_n = \left(\sum_{i=1}^n \mathbf{a}_i, \sum_{i=1}^n \langle \mathbf{a}_i, \mathbf{s} \rangle \right).$$

В работе [24] выполнена модификация фазы восстановления с помощью добавления порога t , тем самым превратив схему из (n, n) в (t, n) . Используя $t \leq n$ долей, решена система линейных уравнений для нахождения секрета \mathbf{s} . Поскольку полученная система уравнений является недоопределенной, было предложено использовать метод наименьших квадратов (МНК) [25] для нахождения приближенного решения. Однако применение МНК не всегда возможно или эффективно, что накладывает на схему дополнительные ограничения.

Таблица 3. Сравнение размера ключевых элементов схемы на решетках, бит

Table 3. Comparison of the key elements size of a lattice-based scheme, bit

Параметр	128-битная безопасность	256-битная безопасность
Секрет s	20 224	38 912
Ошибка e_i	45	64
Доля D_i	20 224	38 912

В табл. 3 представлены значения размеров ключей и долей секрета для 128-битной и 256-битной безопасности. Данные взяты на основе рекомендации NIST.

Как видно из табл. 3 для хранения секрета и его долей, необходимо значительно больше ресурсов, чем в классических схемах. Это может ограничивать применение схемы в системах с ограниченными ресурсами.

Схемы разделения секрета, основанные на изогениях эллиптических кривых

Эллиптические кривые, традиционно использующиеся в классических подходах, основаны на задаче дискретного логарифма для эллиптических кривых, демонстрируют уязвимость к квантовым атакам с применением алгоритма Шора [26]. Это послужило стимулом для развития нового направления — криптографии на изогениях, где вместо работы с точками на кривых используются сами эллиптические кривые и морфизмы между ними (изогении) [27].

Криптография на изогениях базируется на вычислительной сложности задачи нахождения изогении между заданными изогенными кривыми над конечным полем $GF(p)$ [28, 29]. Этот подход породил множество научных работ в этой области [30, 31].

Одной из прорывных работ является разработка протоколов обмена ключами SIDH [32] и CSIDH [33]. Последний протокол предлагает более компактные размеры ключей за счет использования суперсингулярных кривых. Особого внимания заслуживает схема цифровых подписей CSI-FiSh [34], демонстрирующая практическую применимость подхода.

В контексте пороговых схем важными разработками стали модификации схем Кувеньес–Ростовцев–Столбунова [35], основанные на классической схеме Шамира и использующие концепцию Hard Homogeneous Spaces [36, 37]. Безопасность этих схем опирается на сложность решения Group Action Inverse Problem [38, 39].

Рассмотрим пороговую схему группового действия, предложенную в [36]. Пусть P_i — участник пороговой схемы и владеет частью секрета $s_i = f(i)$, где сам секрет — сумма всех долей:

$$S = \sum_{i \in S} s_i, S \in \mathbf{Z}_q.$$

Основная идея заключается в вычислении группового действия $[s]E_0$ без раскрытия своих долей. Пусть $S \subset \{1, \dots, n\}$ — множество мощности не менее k . Тогда секрет может быть вычислен как:

$$S = \sum_{i \in S} s_i \times L_{0,i}^S,$$

где $L_{0,i}^S$ — коэффициенты Лагранжа P_i .

Пусть в схеме участвуют k участников. Сначала E_0 отправляется участнику P_1 , который вычисляет:

$$E_1 = [s_1 \times L_{0,1}^S]E_0, \\ [a]E = g^a \times E,$$

где g — элемент порядка q группы $GF(q)$.

Результат E_1 отправляется следующему участнику $P_j, j \leq k$, который вычисляет:

$$E_j = [s_j \times L_{0,j}^S]E_{j-1} = [\sum_{i \in S} s_i \times L_{0,i}^S]E_0 = [s]E_0.$$

Таким образом, общая структура коммуникации выглядит следующим образом:

$$\xrightarrow{E_0, s} P_1 \xrightarrow{E_1, s} P_2 \xrightarrow{E_2, s} \dots \xrightarrow{E_{k-1}, s} P_k \xrightarrow{[s]E_0}.$$

Особенность системы заключается в том, что только последний участник знает секрет, и он должен распределить его между другими участниками. В табл. 4 представлены размеры ключевых значений, основываясь на NIST Level 1 (минимальный уровень стойкости в постквантовых стандартах NIST [4]).

Из табл. 4 можно сделать вывод, что схемы на основе изогений обладают наименьшими размерами по сравнению с постквантовыми аналогами, представленными в данной работе. Основная проблема — выбор параметров для обеспечения безопасности системы. Используя параметры CSIDH, нельзя установить схему разделения секрета из-за недостатка квантовой безопасности. В итоге в [36] сделан вывод, что безопасность схемы соответствует уровню NIST Level 1, но сделать ее более безопасной проблематично [40]. Для обеспечения безопасности необходимо использовать суперсингулярные кривые с параметрами поля $p \geq 2^{254}$.

Схемы разделения секрета, основанные на хэш-функциях

Хэш-функции являются одним из основных элементов при проектировании схем разделения секрета. Их безопасность основывается на свойстве стойкости к прообразу, которое гарантирует, что вычислительно невозможно восстановить исходное сообщение по его хэш-значению [41]. Однако существенным недостатком схем, основанных на хэш-функциях, является то, что их безопасность напрямую зависит от стойкости используемой хэш-функции.

Table 4. Размер ключевых элементов схемы, бит

Table 4. Size of the key elements, bit

Параметр	128-битная безопасность
Секрет s	512
Кривая E_j	512
Доля s_i	512

Одна из первых схем распределения секрета, основанная на хэш-функциях, описана в работе [42], в которой была выдвинута идея использования хэш-функций для проектирования схем распределения секрета. Впоследствии данная схема получила развитие в работах [43–45]. В [43] хэш-функции использовались как дополнительный механизм для повышения безопасности и эффективности схем. Однако в [44, 45] хэш-функции стали основным вычислительным механизмом, что значительно улучшило производительность схем.

Среди последних работ стоит отметить схему Кета и Чжан [46]. При инициализации схемы случайным образом генерируются n долей s_1, \dots, s_n для n участников. Далее задаются авторизованные подмножества $A_i = \{p_{A_i,1}, \dots, p_{A_i,t}\}$, которые определяют группы пользователей, которые могут восстановить секрет S .

Для каждого авторизованного подмножества A_i формируется приватное сообщение PM_i путем конкатенации долей участников подмножества:

$$PM_i = s_{A_i,1} || s_{A_i,2} || \dots || s_{A_i,t}.$$

Для каждого приватного сообщения PM_i вычисляется его хэш-значение:

$$h_i = H(PM_i),$$

где h_i имеет тот же размер, что и секрет s . Для каждого авторизованного подмножества генерируется контрольное значение:

$$c_i = h_i \oplus s,$$

где \oplus — побитовое исключающее ИЛИ (XOR).

Чтобы восстановить общий секрет s для авторизованного подмножества A_i , участники объединяют свои доли:

$$PM_i = s_{A_i,1} || s_{A_i,2} || \dots || s_{A_i,t}.$$

Вычисляется хэш-значение восстановленного приватного сообщения

$$h_i = H(PM_i).$$

С использованием контрольного значения c_i секрет S восстанавливается следующим образом:

$$S = h_i \oplus c_i.$$

Размеры секрета и его долей могут быть любыми так как, они затем преобразуются в хэш-значения.

К преимуществам схемы можно отнести теоретико-информационную идеальность схемы, т. е. частичная информация о долях участников авторизованного подмножества не предоставляет никакой информации о секрете до тех пор, пока их количество не станет равным заданному порогу схемы.

Возможные атаки на схему совпадают с атаками на хэш-функцию. Использование современных хэш-функций таких как Кессак [47], делает схему устойчивой к атакам с применением квантового компьютера.

Однако вычисление хэш-значений для больших авторизованных подмножеств может быть вычислительно затратным. Схема имеет ограниченную гибкость, что делает ее сложной для адаптации в случае динамического изменения авторизованных подмножеств или количества участников.

Схемы разделения секрета на основе многомерных уравнений

Рассмотрим схемы разделения секрета, основанные на многомерных уравнениях. Такие схемы используют многочлены от нескольких переменных для разделения и восстановления секрета.

В работе [48] предложен вариант схемы Шамира, в котором вместо одномерного многочлена использован многомерный многочлен степени n :

$$f(x_1, x_2, \dots, x_m) = \sum_{i=0}^{k-1} a_i x_1^{e_{i1}} x_2^{e_{i2}} \dots x_m^{e_{im}},$$

где $a_i \in GF(p)$; $k = \binom{n+m}{m}$ — биномиальный коэффициент; $e_{im} \in \mathbb{N} \cup \{0\}$, причем $\sum_{j=1}^m e_{ij} \leq n$. Долями секрета являются наборы $x_i = (x_{i1}, \dots, x_{im}, f_i)$. Для секрета необходимо собрать k долей и использовать интерполяционную формулу Лагранжа.

В [49] предложена схема Карнина–Грина–Хеллмана, основанная на решении системы линейных уравнений. Секрет $S \in \mathbb{F}_q$ разделяется с использованием случайных векторов $\mathbf{u}, \mathbf{v}_0 \in \mathbb{F}_q^t$, таких что $\langle \mathbf{u}, \mathbf{v}_0 \rangle = S$. Для вычисления теней выбираются случайные векторы $\mathbf{v}_1, \dots, \mathbf{v}_n \in \mathbb{F}^t$, такие что любые t из них линейно независимы. Тени вычисляются как $S = \langle \mathbf{u}, \mathbf{v}_i \rangle$, для $i \in \{1, \dots, n\}$. Вектор \mathbf{u} забывается, а $\mathbf{v}_1, \dots, \mathbf{v}_n$ становятся публично доступными. Для восстановления секрета необходимо собрать t теней и решить систему линейных уравнений для нахождения \mathbf{u} .

В схеме [50], представляющей собой обобщенную схему Асмута–Блума [51], секретом является многочлен $s(\mathbf{x}) \in \mathbb{F}_q[\mathbf{x}]$ степени меньше d_0 , где $\mathbf{x} = (x_1, \dots, x_m)$. Для разделения секрета генерируется набор взаимно простых многочленов $m_i(\mathbf{x}) \in \mathbb{F}_q[\mathbf{x}]$, где $m_0(\mathbf{x}) = x^{d_0}$. Затем генерируется случайный многочлен $a(x)$, $\deg a(x) \leq \sum_{i=1}^t d_i - d_0 - 1$ и вычисляется $f(x) = s(x) + a(x) \times m_0(x)$. Долями секрета будут являться $s_i(x) = f(x) \bmod m_i(x)$. В фазе комбинации секрета применена китайская теорема об остатках. В [52] предложено обобщение схемы на многомерные многочлены.

Схемы разделения секрета на основе многомерных уравнений обладают свойствами теоретико-информационной идеальности схемы, так как знание менее k долей не дает информации о секрете, где размеры секрета и долей одинаковы. Однако их применение требует учета ограничений, таких как необходимость раскрытия секрета для его использования и сложность добавления новых пользователей. Размеры ключей подобны значениям, что и в схеме разделения секрета по интерполяционным многочленам Лагранжа и представлены в табл. 1.

Сравнительный анализ схем разделения секрета

Для сравнения различных схем разделения секрета использованы критерии Шамира, которые включают: размер доли меньше или равен самому секрету, возможность повторного использования секрета, невозможность проанализировать секрет, возможность добавления нового участника, возможность обновить секрет, возможность изменения веса долей. Результаты представлены в табл. 5.

Как видно из табл. 5 классические схемы (Шамира, Ньютона, на основе китайской теоремы об остатках) остаются актуальными благодаря своей простоте и эффективности, однако они уязвимы к атакам с применением квантовых компьютеров. Схемы на основе теории решеток и кодах, исправляющих ошибки, являются наиболее перспективными для использования в постквантовой криптографии, так как они обеспечивают высокий уровень безопасности и соответствуют большинству критериев Шамира.

В табл. 6 представлены данные по сложности взлома представленных схем, с использованием квантового компьютера и классических способов. Под защитными мерами понимается условия, при которых, указанная атака не может быть реализована за полиномиальное время.

Одной из возможных атак на схему Шамира и схему на основе интерполяции Ньютона, является перебор многочленов, при которой злоумышленник пытается восстановить секретный многочлен $q(x)$ степени $k - 1$, имея менее k долей. Для этого необходимо решить систему уравнений с $k - 1$ неизвестными коэффициентами. Если p будет меньше указанного значения, то

злоумышленник сможет перебрать все возможные варианты многочленов и тем самым восстановить секрет S .

Специфичной атакой на схему на китайской теореме об остатках является перебор модулей, при которой злоумышленник пытается восстановить секрет S , подбирая взаимно простые модули m_1, \dots, m_k при наличии $k - 1$ долей. Если $m_i < 2^{256}$, то злоумышленник сможет восстановить все m_i , и тем самым решить систему уравнений и найти S .

Одной из самых распространенных атак на криптопримитивы на основе кодов, исправляющих ошибки, является атака по информационным совокупностям. Атака представляет собой семейство алгоритмов для решения задачи синдромного декодирования, заключающейся в нахождении вектора ошибок e веса t . При $n < 6960$ злоумышленник сможет построить таблицу синдромов и тем самым восстановить кодовое слово D .

В качестве специальной атаки на схему на решетках, была рассмотрена атака поиска секретного вектора в решетке (SVP, CVP). При $n < 1024$ у злоумышленника появляется возможность перебрать все возможные вектора решетки и тем самым найти кратчайший вектор.

В криптографии на изогениях для нахождения секрета злоумышленнику необходимо найти все возможные пути изогений между кривыми, где p — размер поля. При малом значении $p < 2^{254}$, злоумышленник сможет перебрать все пути в графе изогений кривых, а значит восстановить все кривые E_j и тем самым найти секрет s .

Специфичными атаками на многомерные уравнения являются линеаризация и атаки, использующие базис Грёбнера. Цель данных атак преобразовать нелинейную систему в линейную для решения стандартными мето-

Таблица 5. Сравнение схем разделения секрета по критериям Шамира
Table 5. Comparison of the secret separation schemes according to Shamir's criteria

Критерии Шамира	Схема Лагранжа	Схема разделения секрета по интерполяционному многочлену Ньютона	Схема на китайской теореме об остатках	Схемы на кодах, исправляющих ошибки	Схемы на решетках	Схемы на эллиптических кривых	Схемы на хэш-функциях	Схемы на многомерных уравнениях
Размер доли не больше секрета	+	+	—	+	+	+	+	+
Возможность повторного использования секрета	—	—	—	—	—	—	—	—
Невозможность проанализировать секрет	+	+	+	+	+	+	+	+
Возможность добавления нового участника	+	+	—	+	+	—	—	—
Возможность обновить секрет	+	+	—	+	+	—	+	—
Возможность изменения веса долей	—	—	—	—	—	—	—	—

Таблица 6. Сравнение схем разделения секрета по сложности взлома
 Table 6. Comparison of the secret sharing schemes by hacking complexity

Схема	Квантовый метод	Классический метод	Возможные атаки	Защитные меры
Схема Шамира (RSA)	$O((\log_2 N)^3)$	$O(p^{k-1})$	Квантовый компьютер (алгоритм Шора), перебор многочленов	Замена классической криптографии на постквантовую, $p \geq 2^{256}$
Схема на основе интерполяции Ньютона (RSA)				
Схема на китайской теореме об остатках (RSA)	$O((\log_2 N)^3)$	$O(N)$	Квантовый компьютер (алгоритм Шора), перебор модулей	Замена классической криптографии на постквантовую, $m_i > 2^{256}$
Схемы на кодах, исправляющих ошибки (RSA)	$O((\log_2 N)^3)/O(2^{0,1n})$	$O(n^3)$	Квантовый компьютер (алгоритм Шора, атака на секрет, кодовое слово), атака по информационным совокупностям	Замена классической криптографии на постквантовую, $n \geq 6960$, использовать коды Гоппы
Схемы на решетках	$O(2^{0,1N})$	$O(2^N)$	Поиск секретного вектора в решетке	$n \geq 1024$
Схемы на эллиптических кривых	$O(p^{1/4})$	$O(p^{1/4})$	Перебор всех возможных путей изогений между кривыми	суперсингулярные эллиптические кривые над полем $GF(p)$, $p \geq 2^{254}$
Схемы на хэш-функциях	$O(2^N)$	$O(2^N)$	Коллизии, атаки на прообраз	Использование Кессак (SHA-3), $n \geq 256$
Схемы на многомерных уравнениях	$O(p^m)$	$O(k^x)$	Линеаризация, Грёб-норовы базисы	$x \geq 64, k \geq 5$

Примечание. N — параметр безопасности; n — размер кодового слова; p — размер поля; x — число переменных; k — степень многочленов.

дами. При $x < 64, k < 5$ полученная система уравнений будет слишком простой, что дает злоумышленнику возможность решить ее.

Из табл. 6 видно, что классические схемы разделения секрета уязвимы к квантовым атакам, так как сложность их взлома с использованием квантового компьютера оценивается как $O((\log_2 N)^3)$, что делает их непригодными для долгосрочного применения. В то же время постквантовые схемы, основанные на теории решеток, демонстрируют высокую устойчивость к квантовым атакам.

Схемы на изогениях эллиптических кривых также устойчивы к квантовым атакам, но требуют тщательного выбора параметров. Схемы на хэш-функциях зависят от криптостойкости самой функции, а схемы на многомерных уравнениях сложны в реализации и вычислениях.

Сложность квантовой атаки на схемы, основанные на постквантовых криптопримитивах является теоретической, так как в настоящий момент не существует эффективного квантового алгоритма их решения.

Заключение

В работе проведен анализ современных схем разделения секрета, как классических, так и постквантовых. Рассмотрены схемы, основанные на различных математических примитивах, включая многочлены Лагранжа и Ньютона, китайскую теорему об остатках, коды исправляющие ошибки, теорию решеток, изогении эллиптических кривых, многомерные уравнения и

хэш-функции. Выполнен сравнительный анализ по критериям Шамира, который позволил выделить наиболее перспективные схемы для использования в условиях постквантовой эры.

Классические схемы сохраняют свою актуальность благодаря простоте реализации и высокой эффективности. Однако их уязвимость к атакам с использованием квантовых компьютеров делает их непригодными для долгосрочного применения в условиях постквантовой эры.

Постквантовые схемы, включая схемы на основе теории решеток, изогений эллиптических кривых, многомерных уравнений и хэш-функций, демонстрируют повышенную устойчивость к квантовым атакам. Несмотря на это, они требуют более сложных вычислений, что может ограничивать их применение в системах с ограниченными вычислительными ресурсами, таких как устройства интернета вещей.

Схемы, основанные на теории решеток, выделяются как наиболее перспективные направления в постквантовой криптографии. Они не только обеспечивают высокий уровень безопасности, но и соответствуют большинству критериев Шамира, что делает их пригодными для широкого использования в современных криптографических системах.

Для дальнейшего развития рекомендуется исследовать совместное использование классических схем и новых постквантовых криптопримитивов, особенно на основе теории решеток. Объединение двух подходов сможет сохранить преимущества классических схем и защиту от квантовых атак.

Литература

References

1. Blakley G.R. Safeguarding cryptographic keys // *Proc. of the International Workshop on Managing Requirements Knowledge (MARK)*. 1979. P. 313–313. <https://doi.org/10.1109/mark.1979.8817296>
2. Shamir A. How to share a secret // *Communications of the ACM*. 1979. V. 22. N 11. P. 612–613. <https://doi.org/10.1145/359168.35917>
3. Ильин В.А., Позняк Э.Г. Основы математического анализа. М.: Физматлит, 2005. 646 с.
4. Barker E., Barker W., Burr W., Polk W., Smid M. NIST special publication 800-57 // *NIST Special publication*. 2007. V. 800. N 57. P. 1–142.
5. Desmedt Y., Frankel Y. Threshold cryptosystems // *Lecture Notes in Computer Science*. 1990. V. 435. P. 307–315. https://doi.org/10.1007/0-387-34805-0_28
6. Gennaro R., Jarecki S., Krawczyk H., Rabin T. Secure distributed key generation for discrete-log based cryptosystems // *Lecture Notes in Computer Science*. 1999. V. 1592. P. 295–310. https://doi.org/10.1007/3-540-48910-X_21
7. Menezes A.J. *Handbook of Applied Cryptography*. Taylor & Francis, 1997. 810 p.
8. Feldman P. A practical scheme for non-interactive verifiable secret sharing // *Proc. of the 28th Annual Symposium on Foundations of Computer Science (sfcs 1987)*. 1987. P. 427–438. <https://doi.org/10.1109/SFCS.1987.4>
9. McEliece R.J., Sarwate D.V. On sharing secrets and Reed-Solomon codes // *Communications of the ACM*. 1981. V. 24. N 9. P. 583–584. <https://doi.org/10.1145/358746.358762>
10. Massey J.L. Minimal codewords and secret sharing // *Proc. of the 6th joint Swedish-Russian international workshop on information theory*. 1993. P. 276–279.
11. Martínez-Peñas U. Communication efficient and strongly secure secret sharing schemes based on algebraic geometry codes // *IEEE Transactions on Information Theory*. 2018. V. 64. N 6. P. 4191–4206. <https://doi.org/10.1109/tit.2018.2823326>
12. Sole P., Çalkavur S., Bonnetaze A., Dela Cruz R. *Code Based Secret Sharing Schemes: Applied Combinatorial Coding Theory*. World Scientific Pub Co Inc., 2022. 212 p.
13. Kurihara J., Uyematsu T., Matsumoto R. Secret sharing schemes based on linear codes can be precisely characterized by the relative generalized Hamming weight // *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences*. 2012. V. E95.A. N 11. P. 2067–2075. <https://doi.org/10.1587/transfun.e95.a.2067>
14. Massey J.L. Some applications of coding theory in cryptography // *Codes and Ciphers: Cryptography and Coding IV*. 1995. P. 33–47.
15. Steinfeld R., Wang H., Pieprzyk J. Lattice-based threshold-changeability for standard Shamir secret-sharing schemes // *Lecture Notes in Computer Science*. 2004. V. 3329. P. 170–186. https://doi.org/10.1007/978-3-540-30539-2_13
16. Ajtai M. Generating hard instances of lattice problems (extended abstract) // *Proc. of the 28th Annual ACM Symposium on Theory of Computing (STOC '96)*. 1996. P. 99–108. <https://doi.org/10.1145/237814.237838>
17. Regev O. On lattices, learning with errors, random linear codes, and cryptography // *Journal of the ACM (JACM)*. 2009. V. 56. N 6. P. 1–40. <https://doi.org/10.1145/1568318.1568324>
18. Bansarkhani R.E., Mezzani M. An efficient lattice-based secret sharing construction // *Lecture Notes in Computer Science*. 2012. V. 7322. P. 160–168. https://doi.org/10.1007/978-3-642-30955-7_14
19. Khorasgani H.A., Asaad S., Eghlidos T., Aref M. A lattice-based threshold secret sharing scheme // *Proc. of the 11th International ISC Conference on Information Security and Cryptology*. 2014. P. 173–179. <https://doi.org/10.1109/ISCISC.2014.6994043>
20. Ajtai M., Dwork C. A public-key cryptosystem with worst-case/average-case equivalence // *Proc. of the 29th Annual ACM Symposium on Theory of Computing (STOC '97)*. 1997. P. 284–293. <https://doi.org/10.1145/258533.258604>
21. Lyubashevsky V., Micciancio D. Generalized compact knapsacks are collision resistant // *Lecture Notes in Computer Science*. 2006. V. 4052. P. 144–155. https://doi.org/10.1007/11787006_13
22. Peikert C., Rosen A. Efficient collision-resistant hashing from worst-case assumptions on cyclic lattices // *Lecture Notes in Computer Science*. 2006. V. 3876. P. 145–166. https://doi.org/10.1007/11681878_8
1. Blakley G.R. Safeguarding cryptographic keys. *Proc. of the International Workshop on Managing Requirements Knowledge (MARK)*, 1979, pp. 313–313. <https://doi.org/10.1109/mark.1979.8817296>
2. Shamir A. How to share a secret. *Communications of the ACM*, 1979, vol. 22, no. 11, pp. 612–613. <https://doi.org/10.1145/359168.35917>
3. Ilin V.A., Pozniak E.G. *Fundamentals of Mathematical Analysis*. Moscow, Fizmatlit Publ., 2005. 656 p. (in Russian)
4. Barker E., Barker W., Burr W., Polk W., Smid M. NIST special publication 800-57. *NIST Special publication*, 2007, vol. 800, no. 57, pp. 1–142.
5. Desmedt Y., Frankel Y. Threshold cryptosystems. *Lecture Notes in Computer Science*, 1990, vol. 435, pp. 307–315. https://doi.org/10.1007/0-387-34805-0_28
6. Gennaro R., Jarecki S., Krawczyk H., Rabin T. Secure distributed key generation for discrete-log based cryptosystems. *Lecture Notes in Computer Science*, 1999, vol. 1592, pp. 295–310. https://doi.org/10.1007/3-540-48910-X_21
7. Menezes A.J. *Handbook of Applied Cryptography*. Taylor & Francis, 1997, 810 p.
8. Feldman P. A practical scheme for non-interactive verifiable secret sharing. *Proc. of the 28th Annual Symposium on Foundations of Computer Science (sfcs 1987)*, 1987, pp. 427–438. <https://doi.org/10.1109/SFCS.1987.4>
9. McEliece R.J., Sarwate D.V. On sharing secrets and Reed-Solomon codes. *Communications of the ACM*, 1981, vol. 24, no. 9, pp. 583–584. <https://doi.org/10.1145/358746.358762>
10. Massey J.L. Minimal codewords and secret sharing. *Proc. of the 6th joint Swedish-Russian international workshop on information theory*, 1993, pp. 276–279.
11. Martínez-Peñas U. Communication efficient and strongly secure secret sharing schemes based on algebraic geometry codes. *IEEE Transactions on Information Theory*, 2018, vol. 64, no. 6, pp. 4191–4206. <https://doi.org/10.1109/tit.2018.2823326>
12. Sole P., Çalkavur S., Bonnetaze A., Dela Cruz R. *Code Based Secret Sharing Schemes: Applied Combinatorial Coding Theory*. World Scientific Pub Co Inc., 2022. 212 p.
13. Kurihara J., Uyematsu T., Matsumoto R. Secret sharing schemes based on linear codes can be precisely characterized by the relative generalized Hamming weight. *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences*, 2012, vol. E95.A, no. 11, pp. 2067–2075. <https://doi.org/10.1587/transfun.e95.a.2067>
14. Massey J.L. Some applications of coding theory in cryptography. *Codes and Ciphers: Cryptography and Coding IV*, 1995, pp. 33–47.
15. Steinfeld R., Wang H., Pieprzyk J. Lattice-based threshold-changeability for standard Shamir secret-sharing schemes. *Lecture Notes in Computer Science*. 2004, vol. 3329, pp. 170–186. https://doi.org/10.1007/978-3-540-30539-2_13
16. Ajtai M. Generating hard instances of lattice problems (extended abstract). *Proc. of the 28th Annual ACM Symposium on Theory of Computing (STOC '96)*, 1996, pp. 99–108. <https://doi.org/10.1145/237814.237838>
17. Regev O. On lattices, learning with errors, random linear codes, and cryptography. *Journal of the ACM (JACM)*, 2009, vol. 56, no. 6, pp. 1–40. <https://doi.org/10.1145/1568318.1568324>
18. Bansarkhani R.E., Mezzani M. An efficient lattice-based secret sharing construction. *Lecture Notes in Computer Science*, 2012, vol. 7322, pp. 160–168. https://doi.org/10.1007/978-3-642-30955-7_14
19. Khorasgani H.A., Asaad S., Eghlidos T., Aref M. A lattice-based threshold secret sharing scheme. *Proc. of the 11th International ISC Conference on Information Security and Cryptology*. 2014, pp. 173–179. <https://doi.org/10.1109/ISCISC.2014.6994043>
20. Ajtai M., Dwork C. A public-key cryptosystem with worst-case/average-case equivalence. *Proc. of the 29th Annual ACM Symposium on Theory of Computing (STOC '97)*, 1997, pp. 284–293. <https://doi.org/10.1145/258533.258604>
21. Lyubashevsky V., Micciancio D. Generalized compact knapsacks are collision resistant. *Lecture Notes in Computer Science*, 2006, vol. 4052, pp. 144–155. https://doi.org/10.1007/11787006_13
22. Peikert C., Rosen A. Efficient collision-resistant hashing from worst-case assumptions on cyclic lattices. *Lecture Notes in Computer Science*, 2006, vol. 3876, pp. 145–166. https://doi.org/10.1007/11681878_8

23. Georgescu A. A LWE-based secret sharing scheme // *IJCA Special Issue on Network Security and Cryptography*, NSC. 2011. V. 3. P. 27–29.
24. Dehkordi M.H., Farahi S.T., Mashhadi S. LWE-based verifiable essential secret image sharing scheme $((t, s, k, n) \rightarrow (t, s, k, n))$ -VESIS // *IET Image Processing*. 2024. V. 18. N 4. P. 1053–1072. <https://doi.org/10.1049/ipr2.13006>
25. Golub G.H., Van Loan C.F. *Matrix Computations*. JHU press, 2013. 784 p.
26. Shor P.W. Algorithms for quantum computation: discrete logarithms and factoring // *Proc. of the 35th Annual Symposium on Foundations of Computer Science*. 1994. P. 124–134. <https://doi.org/10.1109/SFCS.1994.365700>
27. Smith B. Pre-and post-quantum Diffie–Hellman from groups, actions, and isogenies // *Lecture Notes in Computer Science*. 2018. V. 11321. P. 3–40. https://doi.org/10.1007/978-3-030-05153-2_1
28. Couveignes J.M. Hard homogeneous spaces: Preprint // *HAL science ouverte*. 2006. hal-04538731
29. Rostovtsev A., Stolbunov A. Public-key cryptosystem based on isogenies // *IACR Cryptology ePrint Archive*. 2006.
30. Teske E. An elliptic curve trapdoor system // *Journal of Cryptology*. 2006. V. 19. P. 115–133. <https://doi.org/10.1007/s00145-004-0328-3>
31. Charles D.X., Lauter K.E., Goren E.Z. Cryptographic hash functions from expander graphs // *Journal of Cryptology*. 2009. V. 22. N 1. P. 93–113. <https://doi.org/10.1007/s00145-007-9002-x>
32. Jao D., De Feo L. Towards quantum-resistant cryptosystems from supersingular elliptic curve isogenies // *Lecture Notes in Computer Science*. 2011. V. 7071. P. 19–34. https://doi.org/10.1007/978-3-642-25405-5_2
33. Castryck W., Lange T., Martindale C., Panny L., Renes J. CSIDH: an efficient post-quantum commutative group action // *Lecture Notes in Computer Science*. 2018. V. 11274. P. 395–427. https://doi.org/10.1007/978-3-030-03332-3_15
34. Beullens W., Kleinjung T., Vercauteren F. CSI-FiSh: efficient isogeny based signatures through class group computations // *Lecture Notes in Computer Science*. 2019. V. 11921. P. 227–247. https://doi.org/10.1007/978-3-030-34578-5_9
35. Stolbunov A. Constructing public-key cryptographic schemes based on class group action on a set of isogenous elliptic curves // *Advances in Mathematics of Communications*. 2010. V. 4. N 2. P. 215–235. <https://doi.org/10.3934/amc.2010.4.215>
36. De Feo L., Meyer M. Threshold schemes from isogeny assumptions // *Lecture Notes in Computer Science*. 2020. V. 12111. P. 187–212. https://doi.org/10.1007/978-3-030-45388-6_7
37. Cozzo D., Smart N.P. Sashimi: cutting up CSI-FiSh secret keys to produce an actively secure distributed signing protocol // *Lecture Notes in Computer Science*. 2020. V. 12100. P. 169–186. https://doi.org/10.1007/978-3-030-44223-1_10
38. Sotáková J. Elliptic curves, isogenies, and endomorphism rings: Preprint. 2020.
39. Kim T. Security analysis of group action inverse problem with auxiliary inputs with application to CSIDH Parameters // *Lecture Notes in Computer Science*. 2020. V. 11975. P. 165–174. https://doi.org/10.1007/978-3-030-40921-0_10
40. Meyer M. *Practical isogeny-based cryptography*: dissertation. Universität Würzburg, 2021. 167 p.
41. Rogaway P., Shrimpton T. Cryptographic hash-function basics: Definitions, implications, and separations for preimage resistance, second-preimage resistance, and collision resistance // *Lecture Notes in Computer Science*. 2004. V. 3017. P. 371–388. https://doi.org/10.1007/978-3-540-25937-4_24
42. Zheng Y., Hardjono T., Seberry J. Reusing shares in secret sharing schemes // *The Computer Journal*. 1994. V. 37. N 3. P. 199–205. <https://doi.org/10.1093/comjnl/37.3.199>
43. Liaojun P., Huixian L., Yumin W. An efficient and secure multi-secret sharing scheme with general access structures // *Wuhan University Journal of Natural Sciences*. 2006. V. 11. N 6. P. 1649–1652. <https://doi.org/10.1007/BF02831842>
44. Lin H.Y., Yeh Y.S. Dynamic multi-secret sharing scheme // *International Journal of Contemporary Mathematical Sciences*. 2008. V. 3. N 1. P. 37–42.
45. Das A., Adhikari A. An efficient multi-use multi-secret sharing scheme based on hash function // *Applied Mathematics Letters*. 2010. V. 23. N 9. P. 993–996. <https://doi.org/10.1016/j.aml.2010.04.024>
23. Georgescu A. A LWE-based secret sharing scheme. *IJCA Special Issue on Network Security and Cryptography*, NSC, 2011, vol. 3, pp. 27–29.
24. Dehkordi M.H., Farahi S.T., Mashhadi S. LWE-based verifiable essential secret image sharing scheme $((t, s, k, n) \rightarrow (t, s, k, n))$ -VESIS). *IET Image Processing*, 2024, vol. 18, no. 4, pp. 1053–1072. <https://doi.org/10.1049/ipr2.13006>
25. Golub G.H., Van Loan C.F. *Matrix Computations*. JHU press, 2013, 784 p.
26. Shor P.W. Algorithms for quantum computation: discrete logarithms and factoring. *Proc. of the 35th Annual Symposium on Foundations of Computer Science*, 1994, pp. 124–134. <https://doi.org/10.1109/SFCS.1994.365700>
27. Smith B. Pre-and post-quantum Diffie–Hellman from groups, actions, and isogenies. *Lecture Notes in Computer Science*, 2018, vol. 11321, pp. 3–40. https://doi.org/10.1007/978-3-030-05153-2_1
28. Couveignes J.M. Hard homogeneous spaces: Preprint. *HAL science ouverte*. 2006. hal-04538731
29. Rostovtsev A., Stolbunov A. Public-key cryptosystem based on isogenies. *IACR Cryptology ePrint Archive*, 2006.
30. Teske E. An elliptic curve trapdoor system. *Journal of Cryptology*, 2006, vol. 19, pp. 115–133. <https://doi.org/10.1007/s00145-004-0328-3>
31. Charles D.X., Lauter K.E., Goren E.Z. Cryptographic hash functions from expander graphs. *Journal of Cryptology*, 2009, vol. 22, no. 1, pp. 93–113. <https://doi.org/10.1007/s00145-007-9002-x>
32. Jao D., De Feo L. Towards quantum-resistant cryptosystems from supersingular elliptic curve isogenies. *Lecture Notes in Computer Science*, 2011, vol. 7071, pp. 19–34. https://doi.org/10.1007/978-3-642-25405-5_2
33. Castryck W., Lange T., Martindale C., Panny L., Renes J. CSIDH: an efficient post-quantum commutative group action. *Lecture Notes in Computer Science*, 2018, vol. 11274, pp. 395–427. https://doi.org/10.1007/978-3-030-03332-3_15
34. Beullens W., Kleinjung T., Vercauteren F. CSI-FiSh: efficient isogeny based signatures through class group computations. *Lecture Notes in Computer Science*, 2019, vol. 11921, pp. 227–247. https://doi.org/10.1007/978-3-030-34578-5_9
35. Stolbunov A. Constructing public-key cryptographic schemes based on class group action on a set of isogenous elliptic curves. *Advances in Mathematics of Communications*, 2010, vol. 4, no. 2, pp. 215–235. <https://doi.org/10.3934/amc.2010.4.215>
36. De Feo L., Meyer M. Threshold schemes from isogeny assumptions. *Lecture Notes in Computer Science*, 2020, vol. 12111, pp. 187–212. https://doi.org/10.1007/978-3-030-45388-6_7
37. Cozzo D., Smart N.P. Sashimi: cutting up CSI-FiSh secret keys to produce an actively secure distributed signing protocol. *Lecture Notes in Computer Science*, 2020, vol. 12100, pp. 169–186. https://doi.org/10.1007/978-3-030-44223-1_10
38. Sotáková J. *Elliptic curves, isogenies, and endomorphism rings*: Preprint. 2020.
39. Kim T. Security analysis of group action inverse problem with auxiliary inputs with application to CSIDH Parameters. *Lecture Notes in Computer Science*, 2020, vol. 11975, pp. 165–174. https://doi.org/10.1007/978-3-030-40921-0_10
40. Meyer M. *Practical isogeny-based cryptography*: dissertation. Universität Würzburg, 2021. 167 p.
41. Rogaway P., Shrimpton T. Cryptographic hash-function basics: Definitions, implications, and separations for preimage resistance, second-preimage resistance, and collision resistance. *Lecture Notes in Computer Science*, 2004, vol. 3017, pp. 371–388. https://doi.org/10.1007/978-3-540-25937-4_24
42. Zheng Y., Hardjono T., Seberry J. Reusing shares in secret sharing schemes. *The Computer Journal*, 1994, vol. 37, no. 3, pp. 199–205. <https://doi.org/10.1093/comjnl/37.3.199>
43. Liaojun P., Huixian L., Yumin W. An efficient and secure multi-secret sharing scheme with general access structures. *Wuhan University Journal of Natural Sciences*, 2006, vol. 11, no. 6, pp. 1649–1652. <https://doi.org/10.1007/BF02831842>
44. Lin H.Y., Yeh Y.S. Dynamic multi-secret sharing scheme. *International Journal of Contemporary Mathematical Sciences*, 2008, vol. 3, no. 1, pp. 37–42.
45. Das A., Adhikari A. An efficient multi-use multi-secret sharing scheme based on hash function. *Applied Mathematics Letters*, 2010, vol. 23, no. 9, pp. 993–996. <https://doi.org/10.1016/j.aml.2010.04.024>

46. Chum C.S., Zhang X. Hash function-based secret sharing scheme designs // *Security and Communication Networks*. 2013. V. 6. N 5. P. 584–592. <https://doi.org/10.1002/sec.576>
47. Bertoni G., Daemen J., Peeters M., Van Assche G. Keccak // *Lecture Notes in Computer Science*. 2013. V. 7881. P. 313–314. https://doi.org/10.1007/978-3-642-38348-9_19
48. Putra R.A., Effendie A.R., Aisah S.N., Christy J.R., Marpaung E.C.F., Syarkowi M.Z.H., Devi P.K. Analyzing the relationship between capital and risk in Indonesian life insurance companies using the 2SLS and GMM methods // *Jurnal Riset dan Aplikasi Matematika*. 2024. V. 8. N 2. P. 142–157.
49. Karnin E., Greene J., Hellman M. On secret sharing systems // *IEEE Transactions on Information Theory*. 1983. V. 29. N 1. P. 35–41. <https://doi.org/10.1109/TIT.1983.1056621>
50. Ning Y., Miao F., Huang W., Meng K., Xiong Y., Wang X. Constructing ideal secret sharing schemes based on Chinese remainder theorem // *Lecture Notes in Computer Science*. 2018. V. 11274. P. 310–331. https://doi.org/10.1007/978-3-030-03332-3_12
51. Asmuth C., Bloom J. A modular approach to key safeguarding // *IEEE Transactions on Information Theory*. 1983. V. 29. N 2. P. 208–210. <https://doi.org/10.1109/tit.1983.1056651>
52. Galibus T., Matveev G. Finite fields, Gröbner bases and modular secret sharing // *Journal of Discrete Mathematical Sciences and Cryptography*. 2012. V. 15. N 6. P. 339–348. <https://doi.org/10.1080/09720529.2012.10698386>
46. Chum C.S., Zhang X. Hash function-based secret sharing scheme designs. *Security and Communication Networks*, 2013, vol. 6, no. 5, pp. 584–592. <https://doi.org/10.1002/sec.576>
47. Bertoni G., Daemen J., Peeters M., Van Assche G. Keccak. *Lecture Notes in Computer Science*, 2013, vol. 7881, pp. 313–314. https://doi.org/10.1007/978-3-642-38348-9_19
48. Putra R.A., Effendie A.R., Aisah S.N., Christy J.R., Marpaung E.C.F., Syarkowi M.Z.H., Devi P.K. Analyzing the relationship between capital and risk in Indonesian life insurance companies using the 2SLS and GMM methods. *Jurnal Riset dan Aplikasi Matematika*, 2024, vol. 8, no. 2, pp. 142–157.
49. Karnin E., Greene J., Hellman M. On secret sharing systems. *IEEE Transactions on Information Theory*, 1983, vol. 29, no. 1, pp. 35–41. <https://doi.org/10.1109/TIT.1983.1056621>
50. Ning Y., Miao F., Huang W., Meng K., Xiong Y., Wang X. Constructing ideal secret sharing schemes based on Chinese remainder theorem. *Lecture Notes in Computer Science*, 2018, vol. 11274, pp. 310–331. https://doi.org/10.1007/978-3-030-03332-3_12
51. Asmuth C., Bloom J. A modular approach to key safeguarding. *IEEE Transactions on Information Theory*, 1983, vol. 29, no. 2, pp. 208–210. <https://doi.org/10.1109/tit.1983.1056651>
52. Galibus T., Matveev G. Finite fields, Gröbner bases and modular secret sharing. *Journal of Discrete Mathematical Sciences and Cryptography*, 2012, vol. 15, no. 6, pp. 339–348. <https://doi.org/10.1080/09720529.2012.10698386>

Авторы

Кустов Елизар Филаретович — аспирант, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, [sc 57672934000](https://orcid.org/0000-0002-0191-1178), <https://orcid.org/0000-0002-0191-1178>, elizarkustov@mail.ru

Беззатеев Сергей Валентинович — доктор технических наук, доцент, заведующий кафедрой, Санкт-Петербургский государственный университет аэрокосмического приборостроения, Санкт-Петербург, 190000, Российская Федерация; профессор, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, [sc 6602425996](https://orcid.org/0000-0002-0924-6221), <https://orcid.org/0000-0002-0924-6221>, sergey.bezzateev@gmail.com

Статья поступила в редакцию 26.02.2025

Одобрена после рецензирования 18.04.2025

Принята к печати 28.05.2025

Authors

Elizhar F. Kustov — PhD Student, ITMO University, Saint Petersburg, 197101, Russian Federation, [sc 57672934000](https://orcid.org/0000-0002-0191-1178), <https://orcid.org/0000-0002-0191-1178>, elizarkustov@mail.ru

Sergey V. Bezzateev — D.Sc., Associate Professor, Head of Department, Saint-Petersburg State University of Aerospace Instrumentation (SUAI), Saint Petersburg, 190000, Russian Federation; Professor, ITMO University, Saint Petersburg, 197101, Russian Federation, [sc 6602425996](https://orcid.org/0000-0002-0924-6221), <https://orcid.org/0000-0002-0924-6221>, sergey.bezzateev@gmail.com

Received 26.02.2025

Approved after reviewing 18.04.2025

Accepted 28.05.2025



Работа доступна по лицензии
Creative Commons
«Attribution-NonCommercial»

doi: 10.17586/2226-1494-2025-25-3-457-465

Deep learning-enhanced contour interpolation techniques for 3D carotid vessel wall segmentation

Nouar Ismail¹, Alexandra S. Vatian², Tatyana A. Polevaya³, Alexander A. Golubev⁴,
 Dmitry A. Dobrenko⁵, Alexey A. Zubanenko⁶, Natalia F. Gusarova⁷✉, Almaz G. Vanyurkin⁸,
 Mikhail A. Chernyavskiy⁹

^{1,2,3,4,5,6,7} ITMO University, Saint Petersburg, 197101, Russian Federation

⁶ IMV LLC, Saint Petersburg, 191119, Russian Federation

^{8,9} Almazov National Medical Research Centre, Saint Petersburg, 197341, Russian Federation

¹ noauresmail@gmail.com, <https://orcid.org/0009-0003-3545-9719>

² alexvatyan@gmail.com, <https://orcid.org/0000-0002-5483-716X>

³ tx15003@gmail.com, <https://orcid.org/0000-0001-6131-0019>

⁴ 9459539@gmail.com, <https://orcid.org/0000-0001-7417-6947>

⁵ enotpalaskun@gmail.com, <https://orcid.org/0009-0006-1485-1166>

⁶ zubdocmri@gmail.com, <https://orcid.org/0000-0001-6953-5239>

⁷ natfed@list.ru✉, <https://orcid.org/0000-0002-1361-6037>

⁸ almaz.vanyurkin@mail.ru, <https://orcid.org/0000-0002-8209-9993>

⁹ machern@mail.ru, <https://orcid.org/0000-0003-1214-0150>

Abstract

When studying human vessels using the contour interpolation method, there is a problem of insufficient data for training neural networks for automatic segmentation of the carotid artery wall. In this paper, automated methods of contour interpolation are proposed to expand the datasets, which allows for improved segmentation of vessel walls and atherosclerotic plaques. In this study, the performance of various interpolation methods is compared with the traditional nearest neighboring technique. A theoretical description and comparative evaluation of Linear, Polar, and Spline interpolation are presented. Quantitative metrics, including the Dice Similarity Coefficient, area and index differences, and normalized Hausdorff distances, are used to evaluate the performance of the methods. Performance evaluations are performed on various vessel morphologies for both the lumen and the outer wall boundaries. The study showed that Linear interpolation achieves better geometric performance (Cohen's Kappa 0.92) and improved neural network performance (Score 0.86) compared to the State-of-the-Art model. The proposed interpolation methods consistently outperform nearest neighbor interpolation. Polar and spline methods are effective in generating anatomically plausible contours with improved smoothness and continuity, eliminating transition artifacts between slices. Statistical analysis confirmed good agreement and reduced variation of these methods. The results of the study are useful for the development of automated tools for assessing atherosclerotic plaque in carotid arteries, which is important for stroke prevention. Implementation of improved interpolation methods into clinical imaging workflows can significantly improve the reliability, accuracy, and clinical utility of vessel wall segmentation.

Keywords

carotid vessel, wall segmentation, contour interpolation, deep learning, plaques, carotid atherosclerosis

Acknowledgements

This work was supported by the Russian Science Foundation, Grant No. 23-11-00346.

For citation: Ismail N., Vatian A.S., Polevaya T.A., Golubev A.A., Dobrenko D.A., Zubanenko A.A., Gusarova N.F., Vanyurkin A.G., Chernyavskiy M.A. Deep learning-enhanced contour interpolation techniques for 3D carotid vessel wall segmentation. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2025, vol. 25, no. 3, pp. 457–465. doi: 10.17586/2226-1494-2025-25-3-457-465

УДК 004.932

Методы интерполяции контуров с использованием глубокого обучения для трехмерной сегментации стенки сонных артерий

Нуар Исмаил¹, Александра Сергеевна Ватян², Татьяна Андреевна Полевая³,
Александр Андреевич Голубев⁴, Дмитрий Александрович Добренко⁵,
Алексей Александрович Зубаненко⁶, Наталия Федоровна Гусарова⁷✉,
Алмаз Гафурович Ванюркин⁸, Михаил Александрович Чернявский⁹

^{1,2,3,4,5,6,7} Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация

⁶ ООО ИМВИЖН, Санкт-Петербург, 191119, Российская Федерация

^{8,9} Национальный медицинский исследовательский центр имени В.А. Алмазова, Санкт-Петербург, 197341, Российская Федерация

¹ noauresmail@gmail.com, <https://orcid.org/0009-0003-3545-9719>

² alexvatyan@gmail.com, <https://orcid.org/0000-0002-5483-716X>

³ tx15003@gmail.com, <https://orcid.org/0000-0001-6131-0019>

⁴ 9459539@gmail.com, <https://orcid.org/0000-0001-7417-6947>

⁵ enotpalaskun@gmail.com, <https://orcid.org/0009-0006-1485-1166>

⁶ zubdocmri@gmail.com, <https://orcid.org/0000-0001-6953-5239>

⁷ natfed@list.ru✉, <https://orcid.org/0000-0002-1361-6037>

⁸ almaz.vanyurkin@mail.ru, <https://orcid.org/0000-0002-8209-9993>

⁹ machern@mail.ru, <https://orcid.org/0000-0003-1214-0150>

Аннотация

Введение. При исследовании сосудов человека методом интерполяции контуров возникает проблема недостатка данных для обучения нейронных сетей с целью автоматической сегментации стенки сонной артерии. В работе предложены автоматизированные методы интерполяции контуров для расширения наборов данных, что позволяет улучшить сегментацию стенок сосудов и атеросклеротических бляшек. В представленном исследовании оценивается эффективность различных методов интерполяции в сравнении с традиционной техникой ближайшего соседа. **Методы.** Представлены теоретическое описание и сравнительная оценка линейной, полярной и сплайн-интерполяции. Для оценки производительности методов использованы количественные метрики, включая коэффициент сходства Дайса, различия в площади и индексе, а также нормализованные расстояния Хаусдорфа. Оценки производительности выполнены на различных морфологиях сосудов как для просвета, так и для внешних границ стенок. **Основные результаты.** Исследование показало, что линейная интерполяция достигает лучших геометрических показателей (Каппа Коэна 0,92) и улучшенной эффективности нейронной сети (оценка 0,86) по сравнению с передовой моделью. Предложенные методы интерполяции стабильно превосходят интерполяцию ближайшего соседа. Полярные и сплайн-методы эффективны при создании анатомически правдоподобных контуров с улучшенной гладкостью и непрерывностью, устраняют артефакты перехода между срезами. Статистический анализ подтвердил хорошую согласованность и уменьшение вариации этих методов. **Обсуждение.** Результаты исследования полезны для разработки автоматизированных инструментов оценки атеросклеротических бляшек в сонных артериях, что важно для профилактики инсульта. Внедрение улучшенных методов интерполяции в клинические рабочие процессы визуализации может значительно повысить надежность, точность и клиническую полезность сегментации стенок сосудов.

Ключевые слова

сонные артерии, сегментация стенки, интерполяция контура, глубокое обучение, бляшки, атеросклероз сонной артерии

Благодарности

Исследование поддержано грантом Российского научного фонда, № 23-11-00346.

Ссылка для цитирования: Исмаил Н., Ватян А.С., Полевая Т.А., Голубев А.А., Добренко Д.А., Зубаненко А.А., Гусарова Н.Ф., Ванюркин А.Г., Чернявский М.А. Методы интерполяции контуров с использованием глубокого обучения для трехмерной сегментации стенки сонных артерий // Научно-технический вестник информационных технологий, механики и оптики. 2025. Т. 25, № 3. С. 457–465 (на англ. яз.). doi: 10.17586/2226-1494-2025-25-3-457-465

Introduction

Methods of Carotid Stenosis Imaging

Atherosclerosis is a leading cause of cardiovascular diseases, with plaque being built up within the arterial walls. The serious outcomes of this include stroke, heart attack, and peripheral artery disease. Of the various regions that are affected, carotid atherosclerosis refers to hardening and narrowing of carotid arteries, with a very high risk from its potential to cause ischemic strokes. The carotid arteries are the vessels through which oxygenated blood reaches the brain. Over time, plaques along their inner walls can

reduce or block blood flow. In worse conditions, plaques may break off and cause strokes. The prevalence of carotid atherosclerosis is high, especially in aging populations and individuals with comorbid conditions, such as hypertension, diabetes, and hyperlipidemia.

Detection and evaluation of the severity of carotid atherosclerosis can prevent catastrophic cardiovascular events. Traditionally, carotid ultrasound represented the first-line imaging modality for such studies, while Magnetic Resonance Imaging (MRI) represented a second-level method for the study of carotid stenosis [1]. However, in recent years, several 3D black-blood MR sequences

with high isotropic resolution, high signal-to-noise ratio, and large coverage have been developed [2, 3]. Among them, 3D-VISTA (Volume Isotropic Turbo Spin Echo Acquisition) method allows large coverage of carotid arteries with submillimeter isotropic resolution in the coronal acquisition, and can depict atherosclerotic lesion burden, severity, and luminal stenosis. This technique enables submillimeter isotropic resolution, allowing comprehensive visualization of the lumen and outer wall. However, challenges remain, including the complexity of 3D image review, large datasets, and the need for extensive training of radiologists to interpret vessel wall thickness accurately. These limitations underscore the importance of developing automated tools for segmentation and measurement of above-mentioned 3D-VISTA images.

Deep learning has emerged as a revolutionary technology in medical imaging. According to [4], Convolutional Neural Networks (CNNs) outperform traditional methods in a range of image segmentation tasks. Carotid atherosclerosis, requiring precise segmentation of the carotid artery, benefits greatly from these advancements. Accurate segmentation is critical for assessing plaque burden, measuring lumen diameter, and determining stenosis. Fig. 1 illustrates the segmentation process, highlighting lumen and arterial wall separation for measurements.

Due to the high potential of 3D-VISTA approach, a specialized Vessel Wall Segmentation Challenge 2022 was dedicated to it. Its primary task was to segment the carotid vessel wall from 3D-VISTA images, enabling clinically relevant measurements, such as wall thickness, lumen area, and stenosis percentage. Our work builds on these objectives by addressing the interpolation of missing annotations and enhancing the segmentation pipeline for robust and clinically applicable results.

Related Works

Several segmentations schemes have been developed over the years to increase accuracy. A semi-automatic

technique was developed in [5], who introduced an inner pathfinding algorithm with active contours without edges. This approach optimized the image-dependent force to detect the walls of the carotid artery, yielding highly accurate results with a Dice coefficient of 0.949. Another semi-automated methods utilized algorithms such as graph cuts and watershed [6]. However, these methods require significant user involvement, making them unsuitable for large-scale clinical settings.

Among the fully automated techniques, the first was the work [7] presenting a method for segmenting and labeling head and neck vessels from CTA volumes. Now among fully automated methods, deep learning-based ones [8–15] completely dominate. Authors [16] pioneered the use of U-Net for coronary artery segmentation and stenosis classification on CTA images. Their work demonstrated the power of CNNs in addressing anatomical variability, and plaque morphology—challenges also present in carotid artery segmentation. Their approach achieved a Dice score of 0.771 for segmentation and an accuracy of 0.750 for stenosis classification. The work [17] introduced a CNN-based approach reformulated as a multi-task regression in polar coordinates for black-blood MRI carotid artery segmentation, achieving a median Dice similarity coefficient of 0.813. Another notable method, CarotidNet [10], employed a 3D convolutional neural network with residual connections and dilated convolutions to segment the carotid artery bifurcation from CTA images, achieving a Dice similarity coefficient of 0.823.

Hybrid approaches have recently gained traction by integrating deep learning with traditional segmentation models. The work [18] proposed a shape-constrained active contour model initialized using deep learning outputs to segment the carotid artery lumen from MR images. This hybrid framework utilized a probability atlas for outer artery wall detection, achieving segmentation accuracy comparable to manual methods while reducing dependency on large labeled datasets. In [9] the Gated

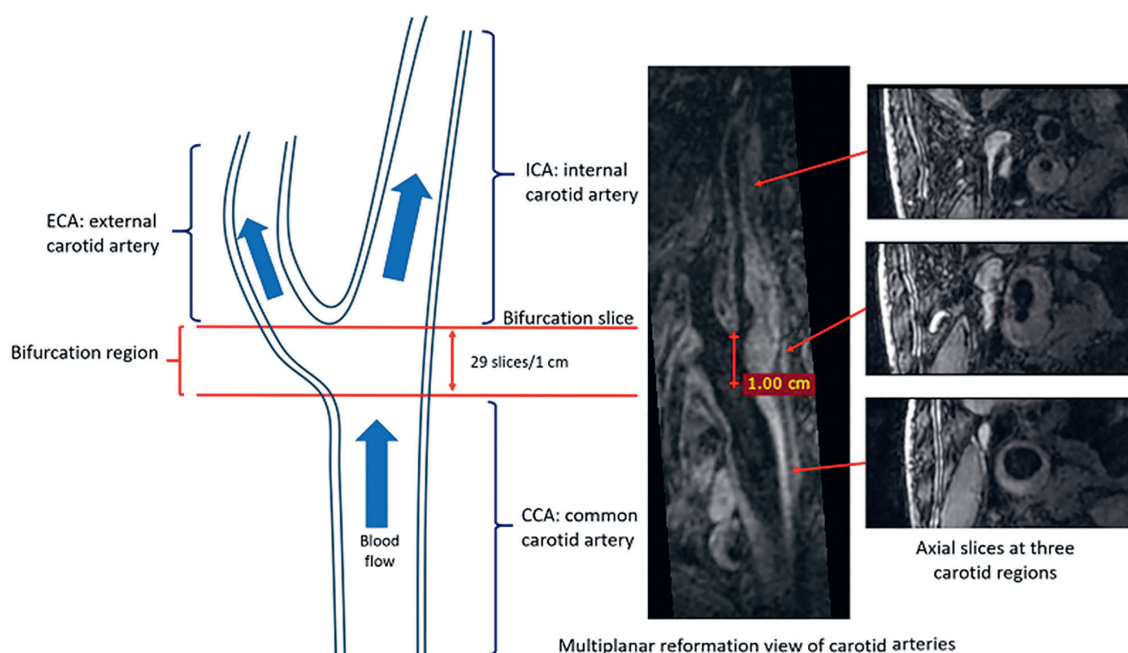


Fig. 1. Illustration of the carotid artery anatomy and corresponding 3D-VISTA images

Multi-Task Network (GMT-Net) is introduced which jointly performs lumen segmentation, outer wall segmentation, and carotid atherosclerosis diagnosis on black-blood MRI. By leveraging multi-task learning and novel gated exchange and fusion layers, the model effectively captured task correlations, achieving Dice scores of 0.9677 and 0.9669 for lumen and outer wall segmentation, respectively, and an AUC of 0.9516 for diagnosis accuracy on the CAREII dataset.

While highly effective, all the above methods rely on a large, annotated dataset, which is challenging in scaling them to smaller datasets. On the contrary, a few works focus on enhancing preprocessing techniques to improve label interpolation before deep learning model training. Specifically, [19] utilized a nearest-neighbor interpolation technique to address the challenge of missing annotations, effectively propagating labels across slices with sparse manual annotations. Notably, their approach achieved the first rank in the Vessel Wall Segmentation 2022 Challenge based on quantitative metrics, so we used it as State-of-the-Art (SOTA) model in our work. However, in our opinion, the potential of interpolation methods is not exhausted by this approach. To test this hypothesis, in our study, we applied three distinct interpolation techniques — Linear, Polar, and Spline interpolation — to generate intermediate contours between annotated slices in a 3D MRI dataset. These methods provided different approaches to interpolate contours based on specific mathematical principles, ensuring smooth transitions between contours in regions lacking direct annotations. The interpolated contours were then used as inputs to our segmentation pipeline to enhance the segmentation of lumen and outer wall structures. Additionally, the performance of these techniques was benchmarked by evaluating their impact on segmentation accuracy, plaque detection, and vessel wall measurement, highlighting their contribution to addressing the challenge of unannotated slices.

Experiments Methodology

Dataset

The dataset used in this study is the Carotid Artery Vessel Wall Segmentation Challenge 2022 dataset consisting images of the carotid artery formed using 3D-VISTA approach. It has specifically been designed for the automation of vessel wall segmentation and has intensive segmentation for both lumen and outer wall in the carotid artery. It provides high resolution images of internal, external, and common carotid arteries, as illustrated in Fig. 1, enabling accurate segmentation, plaque detection, and measurement of vessel wall thickness, all essential for the diagnosis and risk assessment of atherosclerosis.

Each case consists of an axial resliced 3D image volume, with typical dimensions of height 100 px, width 432 px, and depth 432 slices. However, only the middle slices were considered for the assessment of the vessel wall due to the appropriate coronary scan technique. Of these, 80 % of such slices were unannotated and required interpolation techniques to deal with the nonavailability of annotation data. The total dataset includes 50 training cases comprising 2584 manually annotated slices of both

left and right carotids, which outline the contours of the vessel walls: internal common carotid as well as the external carotid. However, due to variable image quality in this dataset, only slices where the vessel wall is clearly visible have been annotated. These annotations originated using the CASCADE software, renowned for its sub-pixel accuracy in vessel wall delineation.

The dataset was annotated in an XML format, using the CASCADE software/algorithm, demonstrate high-resolution visualization of the vessel wall, aiding in the segmentation of lumen and outer wall structures. For the evaluation of participants, key metrics included the Dice Similarity Coefficient (DSC), Hausdorff Distance (HD), and lumen and wall area differences to assess anatomic accuracy and ensure clinically usable results from segmentations. The DSC, defined as

$$DSC = \frac{2|A \cap B|}{|A| + |B|}$$

measures the overlap between the predicted segmentation A and ground truth B . The HD, calculated as

$$HD(A, B) = \max(\sup_{a \in A} \inf_{b \in B} d(a, b), \sup_{b \in B} \inf_{a \in A} d(a, b)),$$

quantifies the largest boundary distance between A and B . Cohen's kappa metric, defined as

$$\kappa = \frac{p_0 - p_e}{1 - p_e},$$

where p_0 represents the observed agreement and p_e the expected agreement by chance, was used to evaluate the inter-rater or algorithmic reliability of segmentation annotations. The normalized area difference was determined using

$$\Delta Area = \frac{|A_{predicted} - A_{GroundTruth}|}{A_{GroundTruth}}$$

providing insights into the clinical relevance of segmentation accuracy.

Preprocessing

Preprocessing is a very significant step in medical image analysis, predominantly when there is incomplete full annotation of the dataset. Only 20 % of slices were manually annotated in the dataset that had been used for this study. To address the lack of annotated slices in the dataset, interpolation techniques were employed to propagate the lumen and wall annotations to unannotated slices. As shown in Fig. 2, the segmentation process involves identifying the lumen and subsequently annotating the outer wall of the carotid artery.

The vessel walls were divided into two important categories to address the discontinuity in the annotations: normal vessel walls and atherosclerotic vessel walls. These differentiation aspects are crucial because variations in the anatomy and pathology can significantly affect the accuracy of segmentation models. Additionally, lumen areas were labeled separately to ensure the accurate representation of vessel structure at a particular point, which is critical for diagnosing stenosis and assessing overall vessel health.

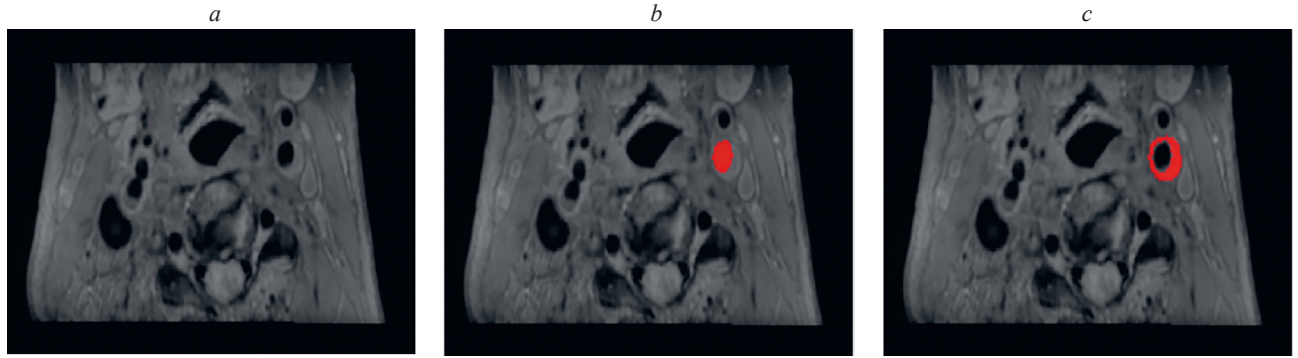


Fig. 2. Visualization of the annotation process for carotid artery segmentation: original axial MRI slice (a); lumen segmentation (b); lumen and outer wall segmentation (c)

Interpolation Methods

Linear Interpolation. Linear interpolation is the most straightforward technique where each point on the interpolated contour is computed as a weighted average between corresponding points on two adjacent contours. Given two contours at slices z_0 and z_1 with points P_0 and P_1 , the linearly interpolated contour at an intermediate slice z_t is given by:

$$P_t = (1 - t)P_0 + tP_1,$$

where t is the normalized interpolation parameter, $0 \leq t \leq 1$. This method provides a direct interpolation of the positions of points but can produce unnatural shapes if the contours are highly irregular, as it does not account for radial transformations or curvature.

Polar Interpolation. Polar interpolation is a modification of Linear interpolation that operates in polar coordinates relative to a centroid, making it more suitable for shapes with rotational symmetry or roughly circular geometry. First, the centroid of each contour is calculated. Then, the contour points are converted to polar coordinates, yielding radius r and angle θ for each point. For corresponding points on two contours, the interpolated radius r_t at an intermediate slice is calculated as:

$$r_t = (1 - t)r_0 + tr_1,$$

while keeping the angle θ constant. The interpolated points are then converted back to Cartesian coordinates. This approach helps maintain the radial structure of the contours, reducing distortions that can arise with linear interpolation when working with circular or elliptical shapes.

Spline Interpolation. To achieve smoother interpolations, we employed Spline interpolation, specifically B-splines, which provide a continuous and smooth fit to the contour points by using piecewise polynomial functions. B-splines are defined by a set of control points and a degree (or order) κ of the polynomial, which determines the smoothness and flexibility of the spline. In our implementation, we selected control points from each original contour and constructed a spline for both x and y coordinates separately.

Let $P_i = (x_i, y_i)_{i=1}^n$ be the set of n points on a given contour. A B-spline $S(t)$ for each coordinate is defined as a weighted sum of basis functions $N_{i,\kappa}(t)$:

$$S_x(t) = \sum_{i=1}^n x_i N_{i,\kappa}(t), \quad S_y(t) = \sum_{i=1}^n y_i N_{i,\kappa}(t),$$

where t is a parameter (typically normalized to range from 0 to 1), and $N_{i,\kappa}(t)$ are the B-spline basis functions of degree κ . The basis functions are recursively defined, starting from $\kappa = 0$ as:

$$N_{i,0}(t) = \begin{cases} 1, & \text{if } t_i \leq (t) < t_{i+1}, \\ 0, & \text{otherwise,} \end{cases}$$

where $\{t_i\}$ is the knot vector that defines the intervals over which each basis function is non-zero.

For a pair of contours on slices z_0 and z_1 , we construct B-splines $S_{x,z_0}(t)$, $S_{y,z_0}(t)$ and $S_{x,z_1}(t)$, $S_{y,z_1}(t)$ for each contour. To interpolate a contour for an intermediate slice z_t (where t is between 0 and 1), we calculate intermediate splines $S_{x,t}(t)$ and $S_{y,t}(t)$ as:

$$S_{x,t}(t) = (1 - t)S_{x,z_0}(t) + tS_{x,z_1}(t),$$

$$S_{y,t}(t) = (1 - t)S_{y,z_0}(t) + tS_{y,z_1}(t).$$

The interpolated contour points for slice z_t are then given by evaluating $S_{x,t}(t)$ and $S_{y,t}(t)$ at uniformly spaced values of t (e.g., from 0 to 1 in equal intervals to ensure uniform sampling).

This approach allows each interpolated contour to inherit the smoothness of the original B-spline functions, providing a natural and continuous shape. The piecewise polynomial nature of B-splines ensures that any sharp changes or high curvature present in the original contours are appropriately captured, while avoiding artifacts like abrupt angles or linear segments that may appear with simpler interpolation methods.

The overall framework implementation for carotid artery segmentation, from preprocessing to final model training, is illustrated in Fig. 3. This pipeline includes preprocessing and label splitting, followed by contours interpolation, machine learning-based label propagation, and the final nnUNet model training.

nnUNet Model Training

After pre-processing, the full image volumes of segmentation results were trained and predicted using the nnUNet model. nnUNet is one of the most important frameworks in medical image segmentation due to its capability in adapting model architecture and hyper

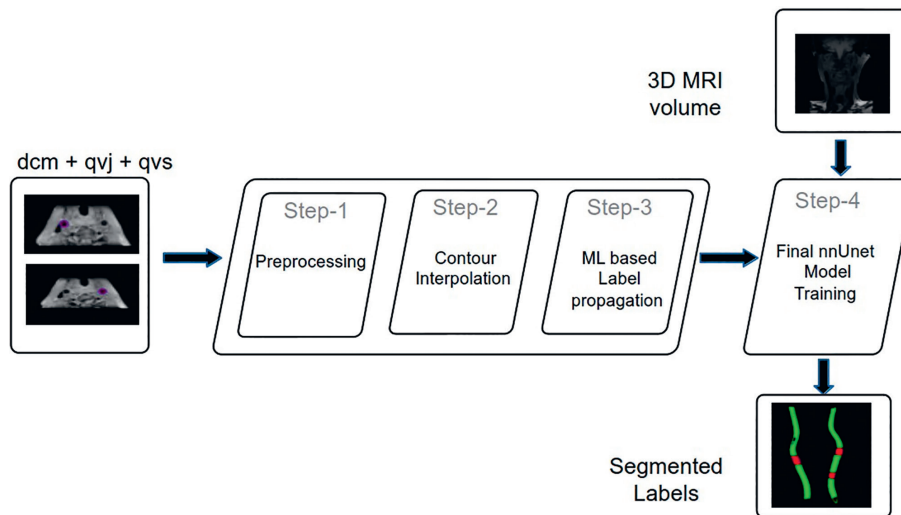


Fig. 3. Implementation of Carotid Artery Segmentation Framework: dcm — input DICOM file, qvj + qvs are XML files holding information about the annotations

parameters to the specifics of a given dataset with minimal manual tuning. Its strength lies in being self-configuring, thereby auto-tuning model settings for image modalities, resolutions, and segmentation tasks across medical datasets.

In our approach, we first trained the nnUNet on the interpolated labels from preprocessing. Training was done on fully interpolated labels throughout the volume of the 3D image, allowing the network to learn from both slices with and without annotations. Segmentation of medical imagery, including this dataset, highlights the strengths of nnUNet. It demonstrates resilience with respect to complex anatomical structures, such as the shapes of the carotid artery wall and the potential presence of atherosclerotic plaques.

The final model used was the nnUNetV2 version which includes several improvements related to efficiency, increasing image sizes, and integrating recent innovations in medical image segmentation. Generally, nnUNet generalizes well to various datasets and medical imaging modalities, making it particularly well-suited for use cases with only a few annotations available in the dataset. By combining both automatic configuration and robust architectural features, the nnUNet model has become the SOTA solution in medical image segmentation, ranking first consecutively in competitions like MICCAI.

The combination of contours interpolation during pre-processing with the powerful architecture of nnUNet created an effective pipeline for segmenting vessel walls of the carotid artery, even with only partial annotations. This approach ensured robust and accurate segmentation, addressing challenges of incomplete labeling and the complex anatomical structures encountered in medical imaging.

Experiments and Results

Implementation Details and Evaluation Metric

The available dataset, consisting of 50 cases, was divided into a training set and a testing set. Specifically, 43 cases (cases 3 to 7, 15, and 17 to 54) were designated for training, while 7 cases (cases 8 to 14) were reserved for testing.

The model was trained for 500 epochs without employing cross-validation to evaluate its performance on the test set.

For this implementation, we utilized an NVIDIA RTX 4060 GPU. It is important to note that all data augmentation techniques, such as random cropping, random rotation, random scaling, random flipping, random Gaussian noise addition, and elastic deformation, were automatically handled by the nnUNet framework. These augmentation techniques were applied by the model to increase the diversity of the training data without any manual intervention.

Results

In this study, we evaluated four interpolation methods — Linear, Nearest Neighbor, Polar, and Spline interpolation — on cases 8, 9, 10, 11, 12, 13, and 14. Our results were then compared with those from the official challenge page of the Vessel Wall Segmentation Challenge 2022, where models were trained on a more extensive dataset (50 cases) and evaluated on a larger test set (25 cases) compared to our limited dataset (43 training cases and 7 test cases). These results were used as SOTA model. Additionally, our requests to obtain annotated test cases from the challenge organizers went unanswered, limiting our ability to directly validate against the challenge data.

The results are presented in Table.

Among the four interpolation methods tested, Polar interpolation generally performed well across most metrics, balancing accuracy and computational stability, particularly in cases with few missing slices. The Spline interpolation method, however, showed high variability, especially in more complex cases with irregular contours, as evidenced by its DSC, lumen area, and wall area differences. For example, case 11 displayed substantial anomalies across all interpolation methods, with Polar and Spline interpolations yielding particularly high errors in Hausdorff Distance and Normalized Wall Index Difference. The inconsistency in case 11 could be attributed to significant structural differences or abnormalities in the vessel wall contours, causing instability in interpolation-based reconstructions. Notably, similar anomalies were detected when applying

Table. Comparison of Contour Interpolation Techniques Results for Carotid Vessel Wall Segmentation Metrics

Metrics	Interpolation Methods			
	Linear	Spline	Polar	SOTA
DSC	0.80	0.72	0.76	0.73
Lumen Area Difference	0.07	0.05	0.06	0.08
Wall Area Difference	0.09	0.18	0.13	0.18
Normalized Wall Index Difference	0.79	3.74	1.01	2.64
Hausdorff Distance (Lumen)	0.17	0.16	0.16	0.17
Hausdorff Distance (Wall)	0.38	0.98	0.58	1.16
Cohen's Kappa	0.92	0.64	0.77	0.85
Quantitative Score	0.86	0.75	0.80	0.79

the Nearest Neighbor interpolation method, indicating that these variances might be inherent to the data or an indication of irregular anatomy in these cases.

Comparing Linear interpolation to Nearest Neighbor interpolation, we found that Linear interpolation yielded consistently higher Quantitative Scores and lower wall area differences, indicating its better ability to smooth transitions between slices. Nearest Neighbor interpolation, however, provided competitive results in lumen area differences and sometimes surpassed Linear in Cohen's Kappa, suggesting that it may better preserve abrupt contour changes in cases with more regular slice structures.

Overall, the Polar interpolation method emerges as a balanced choice, particularly when maintaining smoothness between slices is crucial. Linear interpolation also remains a solid choice for cases with relatively even contours, while Spline interpolation high sensitivity to contour irregularities makes it less reliable in structurally complex cases.

Conclusion

In conclusion, the analysis and comparison of different contour interpolation methods, namely Polar interpolation,

Spline interpolation, and Linear interpolation, demonstrated superior performance over the nearest-neighbor interpolation method across several key evaluation metrics. The suggested interpolation methods consistently yielded higher Dice Similarity Coefficient scores, reduced area and index differences, and smaller normalized Hausdorff distances for both lumen and wall boundaries. These improvements indicate more precise and reliable segmentation, especially in cases where slice alignment is challenging. The enhanced performance of Polar and Spline interpolation methods over nearest-neighbor interpolation is evident in their ability to produce smoother and more anatomically accurate contours, reducing abrupt transitions between slices. This accuracy is particularly beneficial in clinical applications where precise wall and lumen boundaries are essential for diagnostics and treatment planning. The statistical analysis further supports the consistency and robustness of the proposed interpolation techniques, with reduced variance in the results compared to nearest-neighbor interpolation. Therefore, these advanced contour interpolation methods present a valuable improvement for vessel wall segmentation, offering greater accuracy and dependability in medical imaging workflows.

References

- David E., Grazhdani H., Aliotta L., Gavazzi L.M., Foti P.V., Palmucci S., Ini C., Tiralongo F., Castiglione D., Renda M., Pacini P., Di Bella C., Solito C., Gigli S., Fazio A., Bella R., Basile A., Cantisani V. Imaging of carotid stenosis: where are we standing? Comparison of multiparametric ultrasound, CT angiography, and MRI angiography, with recent developments. *Diagnostics*, 2024, vol. 14, no. 16, pp. 1708. <https://doi.org/10.3390/diagnostics14161708>
- Wei H., Zhang M., Li Y., Zhao X., Canton G., Sun J., Xu D., Zhou Z., Chen S., Ferguson M., Hatsukami T., Li R., Yuan C. Evaluation of 3D multi-contrast carotid vessel wall MRI: a comparative study. *Quantitative Imaging in Medicine and Surgery*, 2020, vol. 10, no. 1, pp. 269–282. <https://doi.org/10.21037/qims.2019.09.11>
- Deng F., Le L., Sharma R., et al. *MR vessel wall imaging*. Reference article. Radiopaedia.org. 2019. Available at: <https://doi.org/10.5334/rID-72016> (accessed: 24.01.2025)
- Litjens G., Kooi T., Bejnordi B.E., Setio A.A.A., Ciompi F., Ghafoorian M., van der Laak J.A.W.M., van Ginneken B., Sánchez C.I. A survey on deep learning in medical image analysis. *Medical Image Analysis*, 2017, vol. 42, pp. 60–88. <https://doi.org/10.1016/j.media.2017.07.005>
- Hachaj T., Ogiela M.R. Evaluation of carotid artery segmentation with centerline detection and active contours without edges algorithm. *Lecture Notes in Computer Science*, 2012, vol. 7465, pp. 468–478. https://doi.org/10.1007/978-3-642-32498-7_35

Литература

- David E., Grazhdani H., Aliotta L., Gavazzi L.M., Foti P.V., Palmucci S., Ini C., Tiralongo F., Castiglione D., Renda M., Pacini P., Di Bella C., Solito C., Gigli S., Fazio A., Bella R., Basile A., Cantisani V. Imaging of carotid stenosis: where are we standing? Comparison of multiparametric ultrasound, CT angiography, and MRI angiography, with recent developments // *Diagnostics*. 2024. V. 14. N 16. P. 1708. <https://doi.org/10.3390/diagnostics14161708>
- Wei H., Zhang M., Li Y., Zhao X., Canton G., Sun J., Xu D., Zhou Z., Chen S., Ferguson M., Hatsukami T., Li R., Yuan C. Evaluation of 3D multi-contrast carotid vessel wall MRI: a comparative study // *Quantitative Imaging in Medicine and Surgery*. 2020. V. 10. N 1. P. 269–282. <https://doi.org/10.21037/qims.2019.09.11>
- Deng F., Le L., Sharma R., et al. *MR vessel wall imaging*. Reference article // Radiopaedia.org. 2019 [Электронный ресурс]. URL: <https://doi.org/10.5334/rID-72016> (дата обращения: 24.01.2025)
- Litjens G., Kooi T., Bejnordi B.E., Setio A.A.A., Ciompi F., Ghafoorian M., van der Laak J.A.W.M., van Ginneken B., Sánchez C.I. A survey on deep learning in medical image analysis // *Medical Image Analysis*. 2017. V. 42. P. 60–88. <https://doi.org/10.1016/j.media.2017.07.005>
- Hachaj T., Ogiela M.R. Evaluation of carotid artery segmentation with centerline detection and active contours without edges algorithm // *Lecture Notes in Computer Science*. 2012. V. 7465. P. 468–478. https://doi.org/10.1007/978-3-642-32498-7_35

6. Hemmati H.R., Alizadeh M., Kamali-Asl A., Shirani S. Semi-automated carotid lumen segmentation in computed tomography angiography images. *Journal of Biomedical Research*, 2017, vol. 31, no. 6, pp. 548–558. <https://doi.org/10.7555/JBR.31.20160107>
7. Cuisenaire O. Fully automated carotid artery segmentation from CTA. *The MIDAS Journal*, 2009. <https://doi.org/10.54294/6mmv98>
8. Saba L., Sanagala S.S., Gupta S.K., Koppula V.K., Laird J.R., Viswanathan V., Sanches M.J., Kitis G.D., Johri A.M., Sharma N., Nicolaides A., Suri J.S. A multicenter study on carotid ultrasound plaque tissue characterization and classification using six deep artificial intelligence models: A stroke application. *IEEE Transactions on Instrumentation and Measurement*, 2021, vol. 70, pp. 2505312. <https://doi.org/10.1109/TIM.2021.3052577>
9. Wu J., Xin J., Yang X., Matkovic L., Zhao X., Zheng N., Li R. Segmentation of carotid artery vessel wall and diagnosis of carotid atherosclerosis on black blood magnetic resonance imaging with multi-task learning. *Medical Physics*, 2023, vol. 51, no. 3, pp. 1775–1797. <https://doi.org/10.1002/mp.16728>
10. Zhou T., Tan T., Pan X., Tang H., Li J. Fully automatic deep learning trained on limited data for carotid artery segmentation from large image volumes. *Quantitative Imaging in Medicine and Surgery*, 2021, vol. 11, no. 1, pp. 67–83, 2021. <https://doi.org/10.21037/qims-20-286>
11. Samber D.D., Ramachandran S., Sahota A., Naidu S., Pruzan A., Fayad Z.A., Mani V. Segmentation of carotid arterial walls using neural networks. *World Journal of Radiol.*, 2020, vol. 12, no. 1, pp. 1–9. <https://doi.org/10.4329/wjr.v12.i1.1>
12. Wang J., Yu F., Zhang M., Lu J., Qian Z. A 3D framework for segmentation of carotid artery vessel wall and identification of plaque compositions in multi-sequence MR images. *Computerized Medical Imaging and Graphics*, 2024, vol. 116, pp. 102402. <https://doi.org/10.1016/j.compmedimag.2024.102402>
13. Huang Q., Tian H., Jia L., Li Z., Zhou Z. A review of deep learning segmentation methods for carotid artery ultrasound images. *Neurocomputing*. 2023. V. 545. P. 126298. <https://doi.org/10.1016/j.neucom.2023.126298>
14. Yoo SW., Yang S., Kim J., Huh K., Lee S., Heo M., Yi W. CACSNet for automatic robust classification and segmentation of carotid artery calcification on panoramic radiographs using a cascaded deep learning network. *Scientific Reports*, 2024, vol. 14, no. 1, pp. 13894. <https://doi.org/10.1038/s41598-024-64265-4>
15. Wang Y., Yao Y. Application of artificial intelligence methods in carotid artery segmentation: a review. *IEEE Access*, 2023, vol. 11, pp. 13846–13858. <https://doi.org/10.1109/ACCESS.2023.3243162>
16. Li Y., Wu Y., He J., Jiang W., Wang J., Peng Y., Jia Y., Xiong T., Jia K., Yi Z., Chen M. Automatic coronary artery segmentation and diagnosis of stenosis by deep learning based on computed tomographic coronary angiography. *European Radiology*, 2022, vol. 32, no. 9, pp. 6037–6045. <https://doi.org/10.1007/s00330-022-08761-z>
17. Alblas D., Brune C., Wolterink J.M. Deep learning-based carotid artery vessel wall segmentation in black-blood mri using anatomical priors. *Proceedings of SPIE*, 2022, vol. 12032, pp. 120320Y. <https://doi.org/10.1117/12.2611112>
18. Huang X., Wang J., Li Z. 3D carotid artery segmentation using shape-constrained active contours. *Computers in Biology and Medicine*, 2023, vol. 153, pp. 106530. <https://doi.org/10.1016/j.compbiomed.2022.106530>
19. Hu S., Liao Z., Xia Y. Label propagation for 3D carotid vessel wall segmentation and atherosclerosis diagnosis. *arXiv*, 2022, arXiv:2208.13337. <https://doi.org/10.48550/arXiv.2208.13337>
6. Hemmati H.R., Alizadeh M., Kamali-Asl A., Shirani S. Semi-automated carotid lumen segmentation in computed tomography angiography images // Journal of Biomedical Research. 2017. V. 31. N 6. P. 548–558. <https://doi.org/10.7555/JBR.31.20160107>
7. Cuisenaire O. Fully automated carotid artery segmentation from CTA // The MIDAS Journal. 2009. <https://doi.org/10.54294/6mmv98>
8. Saba L., Sanagala S.S., Gupta S.K., Koppula V.K., Laird J.R., Viswanathan V., Sanches M.J., Kitis G.D., Johri A.M., Sharma N., Nicolaides A., Suri J.S. A multicenter study on carotid ultrasound plaque tissue characterization and classification using six deep artificial intelligence models: A stroke application // IEEE Transactions on Instrumentation and Measurement. 2021. V. 70. P. 2505312. <https://doi.org/10.1109/TIM.2021.3052577>
9. Wu J., Xin J., Yang X., Matkovic L., Zhao X., Zheng N., Li R. Segmentation of carotid artery vessel wall and diagnosis of carotid atherosclerosis on black blood magnetic resonance imaging with multi-task learning // Medical Physics. 2023. V. 51. N 3. P. 1775–1797. <https://doi.org/10.1002/mp.16728>
10. Zhou T., Tan T., Pan X., Tang H., Li J. Fully automatic deep learning trained on limited data for carotid artery segmentation from large image volumes // Quantitative Imaging in Medicine and Surgery. 2021. V. 11. N 1. P. 67–83, 2021. <https://doi.org/10.21037/qims-20-286>
11. Samber D.D., Ramachandran S., Sahota A., Naidu S., Pruzan A., Fayad Z.A., Mani V. Segmentation of carotid arterial walls using neural networks // World Journal of Radiol. 2020. V. 12. N 1. P. 1–9. <https://doi.org/10.4329/wjr.v12.i1.1>
12. Wang J., Yu F., Zhang M., Lu J., Qian Z. A 3D framework for segmentation of carotid artery vessel wall and identification of plaque compositions in multi-sequence MR images // Computerized Medical Imaging and Graphics. 2024. V. 116. P. 102402. <https://doi.org/10.1016/j.compmedimag.2024.102402>
13. Huang Q., Tian H., Jia L., Li Z., Zhou Z. A review of deep learning segmentation methods for carotid artery ultrasound images // Neurocomputing. 2023. V. 545. P. 126298. <https://doi.org/10.1016/j.neucom.2023.126298>
14. Yoo SW., Yang S., Kim J., Huh K., Lee S., Heo M., Yi W. CACSNet for automatic robust classification and segmentation of carotid artery calcification on panoramic radiographs using a cascaded deep learning network // Scientific Reports. 2024. V. 14. N 1. P. 13894. <https://doi.org/10.1038/s41598-024-64265-4>
15. Wang Y., Yao Y. Application of artificial intelligence methods in carotid artery segmentation: a review // IEEE Access. 2023. V. 11. P. 13846–13858. <https://doi.org/10.1109/ACCESS.2023.3243162>
16. Li Y., Wu Y., He J., Jiang W., Wang J., Peng Y., Jia Y., Xiong T., Jia K., Yi Z., Chen M. Automatic coronary artery segmentation and diagnosis of stenosis by deep learning based on computed tomographic coronary angiography // European Radiology. 2022. V. 32. N 9. P. 6037–6045. <https://doi.org/10.1007/s00330-022-08761-z>
17. Alblas D., Brune C., Wolterink J.M. Deep learning-based carotid artery vessel wall segmentation in black-blood mri using anatomical priors // Proceedings of SPIE. 2022. V. 12032. P. 120320Y. <https://doi.org/10.1117/12.2611112>
18. Huang X., Wang J., Li Z. 3D carotid artery segmentation using shape-constrained active contours // Computers in Biology and Medicine. 2023. V. 153. P. 106530. <https://doi.org/10.1016/j.compbiomed.2022.106530>
19. Hu S., Liao Z., Xia Y. Label propagation for 3D carotid vessel wall segmentation and atherosclerosis diagnosis // arXiv. 2022. arXiv:2208.13337. <https://doi.org/10.48550/arXiv.2208.13337>

Authors

Nouar Ismail — PhD Student, ITMO University, Saint Petersburg, 197101, Russian Federation, <https://orcid.org/0009-0003-3545-9719>, noauresmail@gmail.com

Alexandra S. Vatian — PhD, Dean, ITMO University, Saint Petersburg, 197101, Russian Federation, <https://orcid.org/0000-0002-5483-716X>, alexvatyan@gmail.com

Tatyana A. Polevaya — Engineer, Software Developer, ITMO University, Saint Petersburg, 197101, Russian Federation, <https://orcid.org/0000-0001-6131-0019>, tx15003@gmail.com

Авторы

Исмаил Нуар — аспирант, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, <https://orcid.org/0009-0003-3545-9719>, noauresmail@gmail.com

Ватян Александра Сергеевна — кандидат технических наук, декан, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, <https://orcid.org/0000-0002-5483-716X>, alexvatyan@gmail.com

Полёвая Татьяна Андреевна — инженер, программист, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, <https://orcid.org/0000-0001-6131-0019>, tx15003@gmail.com

Alexander A. Golubev — PhD Student, ITMO University, Saint Petersburg, 197101, Russian Federation, <https://orcid.org/0000-0001-7417-6947>, 9459539@gmail.com

Dmitry A. Dobrenko — PhD Student, ITMO University, Saint Petersburg, 197101, Russian Federation, [sc 58793748800](https://orcid.org/0009-0006-1485-1166), <https://orcid.org/0009-0006-1485-1166>, enotpalaskun@gmail.com

Alexey A. Zubanenko — CEO, IMV LLC, Saint Petersburg, 191119, Russian Federation; ITMO University, Saint Petersburg, 197101, Russian Federation; PhD Student, [sc 57215436184](https://orcid.org/0000-0001-6953-5239), <https://orcid.org/0000-0001-6953-5239>, zubdocmri@gmail.com

Natalia F. Gusarova — PhD, Senior Researcher, Associate Professor, ITMO University, Saint Petersburg, 197101, Russian Federation, [sc 57162764200](https://orcid.org/0000-0002-1361-6037), <https://orcid.org/0000-0002-1361-6037>, natfed@list.ru

Almaz G. Vanyurkin — Junior Researcher, Almazov National Medical Research Centre, Saint Petersburg, 197341, Russian Federation, [sc 57295278200](https://orcid.org/0000-0002-8209-9993), <https://orcid.org/0000-0002-8209-9993>, almaz.vanyurkin@mail.ru

Mikhail A. Chernyavskiy — D.Sc. (Medicine), Head of the Department of vascular and endovascular surgery, Almazov National Medical Research Centre, Saint Petersburg, 197341, Russian Federation, [sc 55840916800](https://orcid.org/0000-0003-1214-0150), <https://orcid.org/0000-0003-1214-0150>, machern@mail.ru

Голубев Александр Андреевич — аспирант, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, <https://orcid.org/0000-0001-7417-6947>, 9459539@gmail.com

Добренко Дмитрий Александрович — аспирант, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, [sc 58793748800](https://orcid.org/0009-0006-1485-1166), <https://orcid.org/0009-0006-1485-1166>, enotpalaskun@gmail.com

Зубаненко Алексей Александрович — генеральный директор, ООО ИМВИЖН, Санкт-Петербург, 191119, Российская Федерация; аспирант, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, [sc 57215436184](https://orcid.org/0000-0001-6953-5239), <https://orcid.org/0000-0001-6953-5239>, zubdocmri@gmail.com

Гусарова Наталия Федоровна — кандидат технических наук, старший научный сотрудник, доцент, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, [sc 57162764200](https://orcid.org/0000-0002-1361-6037), <https://orcid.org/0000-0002-1361-6037>, natfed@list.ru

Ванюркин Алмаз Гафурович — младший научный сотрудник, Национальный медицинский исследовательский центр имени В.А. Алмазова, Санкт-Петербург, 197341, Российская Федерация, [sc 57295278200](https://orcid.org/0000-0002-8209-9993), <https://orcid.org/0000-0002-8209-9993>, almaz.vanyurkin@mail.ru

Чернявский Михаил Александрович — доктор медицинских наук, заведующий НИО сосудистой и интервенционной хирургии, Национальный медицинский исследовательский центр имени В.А. Алмазова, Санкт-Петербург, 197341, Российская Федерация, [sc 55840916800](https://orcid.org/0000-0003-1214-0150), <https://orcid.org/0000-0003-1214-0150>, machern@mail.ru

Received 12.12.2024

Approved after reviewing 16.04.2025

Accepted 23.05.2025

Статья поступила в редакцию 12.12.2024

Одобрена после рецензирования 16.04.2025

Принята к печати 23.05.2025



Работа доступна по лицензии
Creative Commons
«Attribution-NonCommercial»

doi: 10.17586/2226-1494-2025-25-3-466-474

Detecting fraud activities in financial transactions using SMOTENN model

Irfan Syamsuddin¹, Sirajuddin Omsa², Andi Rustam³, Dahsan Hasan⁴

^{1,2,4} Politeknik Negeri Ujung Pandang, Makassar, 90245, Indonesia

³ Universitas Muhammadiyah Makassar, Makassar 90221, Indonesia

¹ irfans@poliupg.ac.id, <https://orcid.org/0000-0002-6017-7364>

² sirajud_om@poliupg.ac.id, <https://orcid.org/0000-0003-2776-1531>

³ a.rustam@unismuh.ac.id, <https://orcid.org/0009-0000-2082-4339>

⁴ dahsan.hasan@poliupg.ac.id, <https://orcid.org/0009-0002-6612-0104>

Abstract

The financial industry plays an important role in national economic growth. Because of their critical function, banks have become prime targets for numerous financial crimes. Among these, fraudulent financial transactions are regarded as a severe issue in the financial industry. Conventional approaches are frequently criticized for being ineffective in dealing with fraud in finance; therefore, machine learning approaches have a potential answer to deal with this problem. The goal of this research is to introduce a novel SMOTENN model to establish early detection of cyber fraudulent activities in financial transactions accurately. Two methods are used in this study: first, the Neural Network algorithm is applied to a dataset that contains unbalanced classes; second, the dataset is balanced using the SMOTE (Synthetic Minority Over-sampling Technique) algorithm first, followed by the Neural Network algorithm which we refer to as SMOTENN. The both models are assessed using evaluation metrics of Area Under the Curve, F1-score, precision, recall, specificity, accuracy, and processing time. The comparative analysis shows that the performance of the new SMOTENN model with a balanced dataset is significantly better than that of the neural network approach with an imbalanced dataset, implying that the new SMOTENN model is effective in detecting fraud activities in financial transactions.

Keywords

financial industry, bank, financial fraud, imbalanced dataset, SMOTE, neural network

Acknowledgements

The author would like to thank supports from Politeknik Negeri Ujung Pandang, Indonesia and Universitas Muhammadiyah Makassar, Indonesia.

For citation: Syamsuddin I., Omsa S., Rustam A., Hasan D. Detecting fraud activities in financial transactions using SMOTENN model. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2025, vol. 25, no. 3, pp. 466–474. doi: 10.17586/2226-1494-2025-25-3-466-474

УДК 004.89

Обнаружение мошенничества при финансовых транзакциях с использованием модели SMOTENN

Ирфан Сямсуддин¹, Сираджуддин Омса², Анди Рустам³, Дахсан Хасан⁴

^{1,2,4} Государственный политехнический институт Уджунг Панданга, Макаassar, 90245, Индонезия

³ Университет Мухаммадии в Макассаре, Макаassar, 90221, Индонезия

¹ irfans@poliupg.ac.id, <https://orcid.org/0000-0002-6017-7364>

² sirajud_om@poliupg.ac.id, <https://orcid.org/0000-0003-2776-1531>

³ a.rustam@unismuh.ac.id, <https://orcid.org/0009-0000-2082-4339>

⁴ dahsan.hasan@poliupg.ac.id, <https://orcid.org/0009-0002-6612-0104>

Аннотация

Финансовая индустрия играет важную роль в национальном экономическом росте. Из-за своей критической функции банки стали главными целями для многочисленных финансовых преступлений. Среди них

© Syamsuddin I., Omsa S., Rustam A., Hasan D., 2025

мошеннические финансовые транзакции считаются серьезной проблемой в финансовой индустрии. Традиционные подходы часто критикуются за неэффективность в борьбе с мошенничеством в сфере финансов; поэтому подходы машинного обучения имеют потенциальный ответ для решения этой проблемы. Целью данного исследования является внедрение новой модели Synthetic Minority Over-sampling Technique с Neural Network (SMOTENN) для точного раннего обнаружения кибермошеннических действий в финансовых транзакциях. В работе используются два метода: алгоритм нейронной сети применяется к набору данных, содержащему несбалансированные классы; набор данных первоначально балансируется с помощью алгоритма SMOTE, а затем алгоритма нейронной сети SMOTENN. Обе модели оцениваются с использованием метрик Area Under the Curve, F1-score, точность, полнота, специфичность, достоверность и время обработки. Сравнительный анализ показывает, что эффективность новой модели SMOTENN со сбалансированным набором данных значительно выше, чем у подхода на основе нейронной сети с несбалансированным набором данных. Это свидетельствует об эффективности новой модели SMOTENN в обнаружении мошеннических действий в финансовых транзакциях.

Ключевые слова

финансовая индустрия, банк, финансовое мошенничество, несбалансированный набор данных, SMOTE, нейронная сеть

Благодарности

Автор благодарит за поддержку Политехнический институт Негери Уджунг Панданг, Индонезия, и Университет Мухаммадии Макаassar, Индонезия.

Ссылка для цитирования: Сямсуддин И., Омса С., Рустам А., Хасан Д. Обнаружение мошенничества при финансовых транзакциях с использованием модели SMOTENN // Научно-технический вестник информационных технологий, механики и оптики. 2025. Т. 25, № 3. С. 466–474 (на англ. яз.). doi: 10.17586/2226-1494-2025-25-3-466-474

Introduction

The worldwide financial industry is critical in enabling transactions, managing investments, and assisting firms and individuals with their sophisticated operations. However, the rising sophistication of fraudulent actions has infused the need for powerful fraud detection techniques within the intricacies of this interconnected ecosystem [1]. Financial institutions are continually threatened by a wide range of fraudulent actions, ranging from identity theft to complex cyber attacks.

The potential for fraud to disrupt the financial sector smooth operation has grown significantly in the last decades. In an interconnected world where financial transactions happen at the speed of light, the possibility of fraud wreaking havoc has increased, demanding proactive and advanced fraud detection systems [1, 2].

The significant financial losses experienced by both institutions and their clients are one of the key reasons for the urgency in fraud detection in the financial industry. Fraudulent acts not only deplete financial institutions assets, but also destroy clients' trust in these institutions [3]. The consequences go beyond monetary losses, impacting financial institutions reputation and trustworthiness, and ultimately compromising the stability of the entire financial system. As a result, the requirement for real-time and predictive fraud detection systems has become critical [3–5].

Furthermore, since the global financial system is interrelated, fraud in one region of the world may have repercussions that across international borders. As a result, any solutions that improve the collective resilience of the whole financial ecosystem against the transnational character of financial fraud are urgently needed to identify and trace patterns of suspicious behavior in global financial industries [6].

Machine Learning (ML) emerges as a highly viable method to address the issue. The expanding financial crime scene necessitates a dynamic and adaptive approach to

fraud detection, and traditional methods fall short of modern sophistication [7]. ML algorithms, which are powered by artificial intelligence, can analyze massive volumes of data at previously unheard-of rates, discovering patterns and abnormalities that may signal fraudulent behavior. This adaptive technology keeps financial institutions one step ahead of fraudsters by constantly learning and developing to combat emerging threats [7, 8]. In turn, financial policy makers may obtain correct and quick information regarding any potential fraud in the future.

On one hand, the effectiveness of applying ML algorithms to combat financial fraud is highly depending on the availability of appropriate dataset. The so called imbalanced dataset is prone to results over-fitting or inaccurate results due to very small number of fraud data in comparison to normal data [9].

This phenomenon occurs when the number of transactions is significantly more than the number of fraudulent transactions. This will have a negative impact on the performance of ML models [10]. The so called imbalance data commonly found financial industry in which significant imbalance data between normal and abnormal transactions occurs in the dataset, which is the main concern of our study.

The objective of this study is to present a new ML model based on Synthetic Minority Over-sampling Technique (SMOTE) technique in combination with Neural Network algorithm. An experimental study was also given to compare imbalance dataset and balance dataset after applying SMOTE technique and followed by measuring their performance based on evaluation metrics. It is also important to note that, the study specifically uses Orange data mining software. It is an open source data mining software with many ML algorithms features and it has intuitive graphical user interface [11, 12] suitable for whole analysis in the study in order to showcase the advantages of visual data mining software for non-technical or science audiences.

The whole processes are conducted within Orange data mining software, and open source data mining software with intuitive graphical user interface [11, 12].

Literature Review

The banking sector is a fundamental pillar in supporting global economic dynamics, providing a vital contribution to the circulation of funds and the growth of the business sector. However, the sustainability of this role is faced with serious challenges, especially related to the increase in incidents of fraud in financial transactions. According to Indonesia Fraud Report in 2019 [13], the level of financial crime in Indonesia has increased significantly, posing a serious threat to the stability of the financial sector.

One of the most detrimental forms of threat is financial fraud which includes various activities, such as identity theft, credit card fraud, and transaction manipulation. According to study by Ikbali et al. [14], Indonesia experienced a jump in financial fraud cases in the past year, highlighting the need for a new and effective approach to fighting financial crime.

The importance of involving cutting-edge technology in efforts to prevent banking fraud is becoming increasingly urgent. Reference [15] shows that the application of ML can be a leading solution in mitigating financial risks. ML is able to process and analyze large data volumes quickly, recognize suspicious transaction patterns, and provide early detection of fraudulent activities. Thus, this research puts forward the hypothesis that the integration of ML in banking security systems will pave the way towards more efficient protection of customer finances [16].

The importance of applying ML approach is emphasized in [17] considering bank fraud case is rising significantly, which ends in many damages for the banks. In addition, understanding the underlying patterns in the dataset is important for detecting fraud effectively.

Real time application of ML is studied in [18]. They also realize that datasets in such domain are mostly imbalance. Therefore, techniques to deal with imbalance problems, such as SMOTE, might improve the performance of the classification model after making balance labels for detecting fraud cases in banking sector [19–21].

Another research on imbalance dataset in financial industry is presented in [22]. There is a problem of overfitting in ML results by using datasets that are not directly proportional. Making all labels in appropriately balanced is important for further ML analysis [23]. Using the SMOTE technique, the quality of the dataset processed through ML is very good, so it is recommended to balance data that is still not proportional before further processing.

In [24] it is stated that when the fraud ratio is very small compared to the normal ratio, it is very difficult for classification algorithms to detect fraud cases validly. For this reason, experts must carefully pay attention to this imbalance problem from the beginning of processing the dataset. Apart from that, if imbalanced data is not handled optimally, it is very likely that false positive detection errors will occur, especially if the data is very unbalanced. In fact, in some cases, false warnings are generated, which can also be detrimental to financial institutions.

According to [25], traditional statistical approaches are not suitable to deal with imbalanced data usually occur in bank institutions. Therefore, they propose a model for credit default prediction by employing various credit-related datasets. In addition by considering a significant difference between the minimum and maximum values in different features, they employ Min-Max normalization to scale the features into specific range. Using this approach, imbalance dataset could be handled and solved.

Considering previous studies, the issue of imbalance dataset in financial industry is still a serious problem. Therefore, handling the issue of imbalance dataset has become the objective in this study through the application of oversampling technique and then applying the results to ML algorithm.

Methodology

The research methodology in conducting this study consists of five steps. First of all, it begins with dataset collection. The dataset collected for the study is Bank Marketing Dataset from UCI Machine Learning Repository website¹.

The second step is pre-processing the dataset. Although the dataset has been validated in the study by Moro [26], cleaning the data, solving missing data or other possible errors are still important step to do in any ML analysis in order to ensure readiness of the dataset for further analysis.

Then, once the dataset are clearly free from any errors, we can go through previewing the dataset class or target. In this step, proportion between target classes will be previewed clearly in order to identify the existing of class imbalance as the problem to be tackled in the following steps.

In the fourth step, the dataset will be applied firstly to Neural Network algorithm directly. Then, subsequently the dataset will firstly undergo class balancing using SMOTE technique before Neural Network algorithm. SMOTE technique is a method to handle imbalance dataset by increasing the number of minority class to be equal with majority class [19–21].

Neural Networks is an artificial intelligence algorithms that analyses large datasets by mimicking the functionality of an animal brain such as emulating the synaptic and neural connections found in biological systems [27, 28]. The selection of Neural Network in Orange is based on the requirement of the study to apply algorithm with ability to deal with nonlinear relationships as found in the dataset, in addition to Neural Network powerful capabilities for tasks, such as classification, regression, and pattern recognition. According to the Orange documentation, type of Neural Network used by Orange is a Multi-Layer Perceptron (MLP) algorithm with back propagation². By having the

¹ Bank Marketing Dataset. UCI Repository. Available at: <https://archive.ics.uci.edu/dataset/222/bank+marketing> (accessed: 01.09.2024).

² Neural Network. Orange Documentation. Available at: <https://orange3.readthedocs.io/projects/orange-visual-programming/en/latest/widgets/model/neuralnetwork.html> (accessed: 12.02.2025).

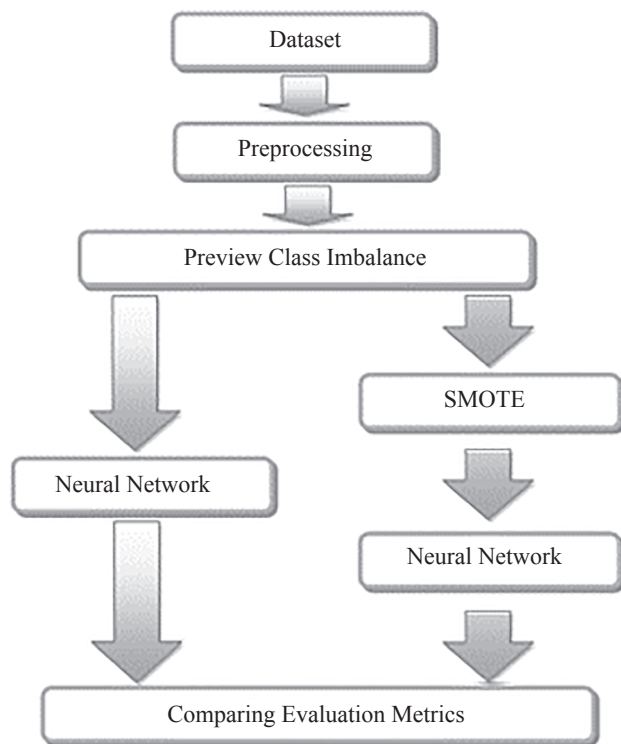


Fig. 1. Research method

basic structure of a Neural Network, the MLP algorithm uses back propagation for increasing the model accuracy which is crucial in such fraud detection case [29].

Finally, both models (Neural Network and SMOTE Neural Network) are evaluated to compare their performance according to accuracy, precision, recall, specificity, Area Under the Curve (AUC), and F1-score. Fig. 1 illustrates the whole steps of research methodology.

Results and Discussion

Dataset collection

This research uses a dataset from an open source or public website (UCI ML Repository: Bank Marketing Data Set, n.d.) with a file size of 3,664 kilobytes. The dataset represents many financial transactions data from a Portuguese banking institution. The data has two labels, namely normal and fraud. Normal label represents normal financial transaction while fraud label means fraudulent financial transactions.

The dataset (Table 1) takes original data obtained from May 2008 to June 2013 with a total of 45,211 rows of data. The dataset has been corrected by undergoing a pre-processing stage to select the data that is really needed so that the data is not too large for research [26].

Data Pre-processing

The study is conducted using Orange Data Mining software installed on Windows 11 machine of AMD A10-9600P RADEON R5, 10 COMPUTE CORES 4C + 6G 2.40 GHz. In Orange, the pre-processing step could be done by applying the Pre-Processing widget which represents the whole pre-processing such as data cleaning and transformation.

The results give a clean and error free dataset as depicted in Fig. 2. It is mentioned that there is no more missing data, 16 features, and a target with 2 values (fraud and normal) and so on. In this clear form, the dataset will be ready for further ML analysis in the next steps.

ML Application with Imbalance Dataset

In the case of fraud detection, we put our concern to the column 17 of the dataset which is Class as the target for analysis with two categories, normal transaction and fraud transaction. Fig. 3 shows the proportion of both class in which there is a significant difference between normal transactions and fraud transactions. This phenomenon is

Table 1. Description of the Dataset

Feature	Type	Description
<i>Age</i>	Numerical	Age
<i>Job</i>	Categorical	Job
<i>Marital</i>	Numerical	Marital status
<i>Education</i>	Categorical	Highest education
<i>Default</i>	Categorical	Credit card status
<i>Balance</i>	Numerical	Amount of transaction
<i>Housing</i>	Categorical	Housing loan status
<i>Loan</i>	Categorical	Bank loan status
<i>Contact</i>	Categorical	Contact
<i>Day</i>	Categorical	Last weekly call log
<i>Month</i>	Categorical	Last monthly call log
<i>Duration</i>	Numerical	Duration of last call
<i>Campaign</i>	Numerical	Number of Bank Campaigns
<i>Pdays</i>	Numerical	Number of Days after Campaigns Campaign
<i>Previous</i>	Numerical	Number of Contact before Campaigns
<i>Poutcome</i>	Categorical	Campaign result
<i>Class</i>	Categorical	Fraud indication (<i>target</i>)

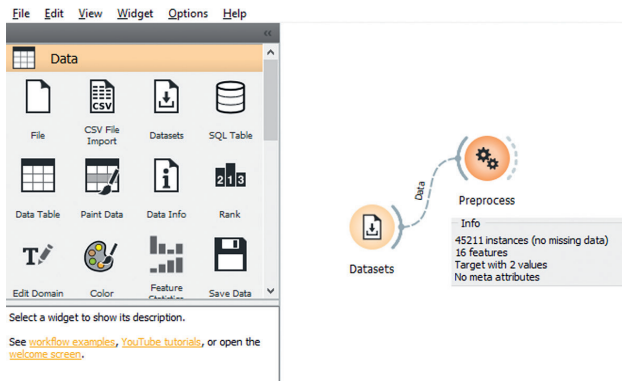


Fig. 2. Pre-processing function in Orange

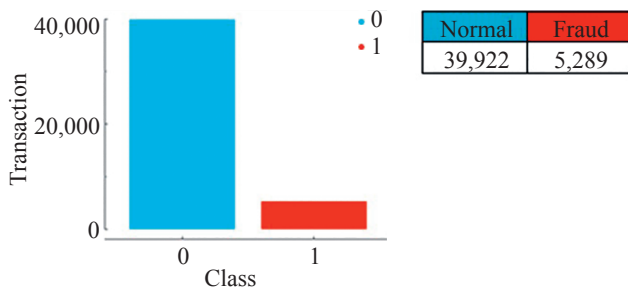


Fig. 3. Imbalanced dataset

commonly found in several cases and is commonly known as imbalance issue of dataset.

Fig. 4 shows that process of ML application in Orange data mining. It begins with loading the dataset, and connect it with ML widget which representing Neural Network algorithm. Furthermore, we perform model evaluation to assess accuracy, precision, recall, specificity, AUC, and F1-score of the applied model.

The Orange Neural Network utilizes sklearn library MLP algorithm by default setting as follows, using Rectified Linear Unit (ReLu) activation, using Adaptive Moment Estimation (Adam) solver, using 100 neurons in

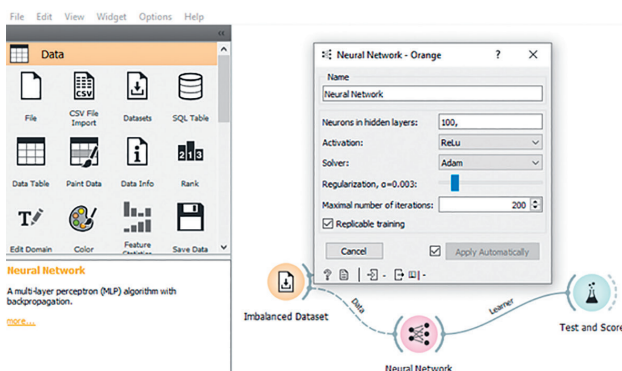


Fig. 4. Implementation of Neural Network on imbalance dataset

hidden layer and 200 iterations. In this step, this default setting is applied to the Imbalanced Dataset as represented in Fig. 4.

Furthermore evaluation is performed using Test and Score widget, we obtain Table 2 which represents the results of evaluation metrics for Neural Network algorithm on the given dataset. These results indicate a robust binary classification model with excellent discrimination (AUC = 90.6 %) and a strong balance between correctly identifying positive instances and avoiding false positives (F1-score = 89.5 %, precision = 89.1 %, recall = 90 %), leading to a high overall accuracy of 90 %. However, only 51.2 % of specificity indicates a notable weakness in correctly identifying negative instances, which is due to class imbalance of the dataset. In addition, the processing time of 1.07 seconds suggests reasonable efficiency of the model.

Implementation of SMOTE

At this step, the dataset that displays imbalance proportion between legitimate and fraudulent financial activities will be corrected by implementing an oversampling technique. As mentioned previously, making the proportional dataset is the objective of oversampling.

The process of oversampling is done in Orange data mining software. The Python code of SMOTE algorithm is supplied to the widget of Python Script in Orange as depicted in Fig. 5.

By using the script feature in Orange software, the SMOTE code in Python could be supplied to perform oversampling process for minority labels. The process is carried out in several steps by adding new fraud data to the dataset repeatedly until a balanced value is obtained. Finally, we obtain a proportional percentage between labels (0 represents normal data, while 1 represents fraud data) as seen in Fig. 6. This new dataset is further called SMOTE Balanced Dataset.

ML Application with Balance Dataset

Previous step has produced a new dataset from the application of SMOTE algorithm to the default dataset called SMOTE Balanced Dataset. In this step, we apply the Neural Network algorithm to the SMOTE Balanced Dataset in Orange data mining.

Similar to the first analysis with imbalance dataset, we apply the default setting of Orange Neural Network as follows by using ReLu activation, Adam solver, 100 neurons in hidden layer and 200 iterations. In this step, this default setting is applied to the SMOTE Balanced Dataset as shown in Fig. 7.

Based on the approach, we produce a new model named SMOTENN which stands for SMOTE Neural Network model and is a combination of SMOTE algorithm and Neural Network. Furthermore, assessment on the SMOTENN model is also conducted to evaluate its accuracy, precision, recall, specificity, AUC, and F1-score for comparison purpose. Table 3 represents the results of

Table 2. Metrics evaluation of Neural Network on imbalance dataset, %

Model	AUC	F1-score	Precision	Recall	Specificity	Accuracy	Processing time, s
Neural Network	90.6	89.5	89.1	90	51.2	90	1.07

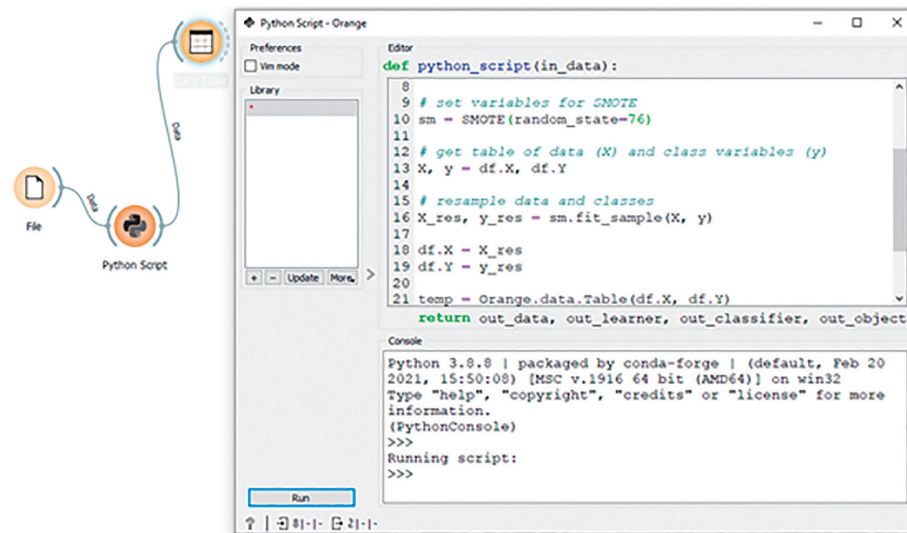


Fig. 5. SMOTE process in Orange

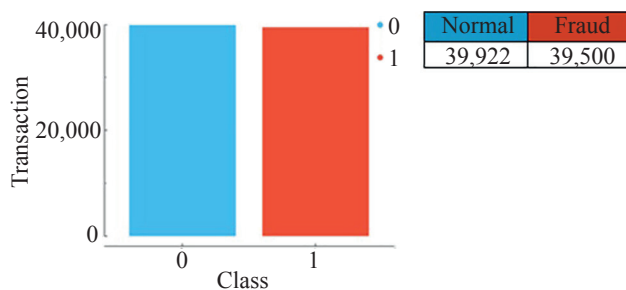


Fig. 6. Balanced dataset after SMOTE

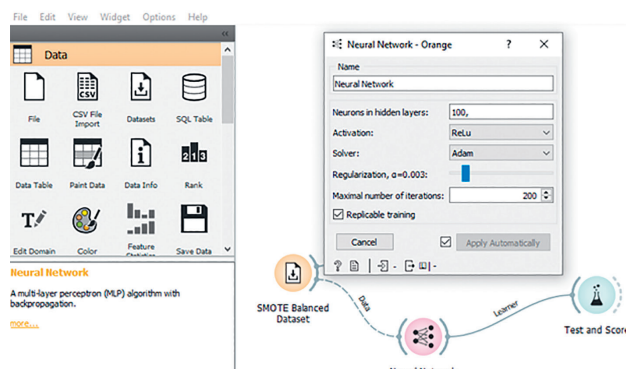


Fig. 7. Implementation of Neural Network on SMOTE balanced dataset

evaluation metrics for SMOTENN model. These results clearly demonstrate a high-performing binary classification model, characterized by excellent discrimination (AUC = 95.7 %) and a strong, well-balanced ability to correctly classify both positive and negative instances, as evidenced by the consistently high F1-score (92.4 %), precision (92.7 %), and recall (92.4 %). A significant

improvement to 92.4 % on specificity is potentially due to class balance impact after applying SMOTE technique. Then, increased accuracy to 92.4 % suggesting effective classification across both groups. Finally, the processing time of 1.44 seconds also imply efficient performance of the model.

Comparative Analysis

In this section, we conduct comparative analysis of the results between the first approach (using imbalance dataset) and the second one using SMOTE based balance dataset.

First of all, we conducted the Kolmogorov–Smirnov (KS) test to compare the distributions of the two datasets, before and after the application of SMOTE (Fig. 8). We obtained KS statistics of 0.3 and a P-value of 0.786. The KS statistics of 0.3 indicates a moderate difference between the two datasets, but the high P-value of 0.7869 shows this difference is not statistically significant as results and suggests the datasets are likely from the same distribution.

So, Fig. 8 illustrates the detail comparison between Neural Network models performance with and without SMOTE across six evaluation metrics, namely AUC, F1-score, Precision, Recall, Specificity, Accuracy. Overall, the SMOTENN model demonstrates consistent improvement in all metrics in comparison to the first model, with detailed comparisons as follows:

From AUC perspective, before applying SMOTE, the AUC was 0.906. After applying SMOTE, it increased to 0.957, representing a relative improvement of approximately 5.62 %. This highlights better discrimination between classes in the new SMOTE model.

Based on F1-score perspective, its F1-score accounts for 0.895, then it rose to 0.924 with SMOTE, reflecting a 3.24 % increase. This indicates enhanced balance between precision and recall in the new SMOTENN model.

Table 3. Metrics evaluation of SMOTE with Neural Network on balance dataset, %

Model	AUC	F1-score	Precision	Recall	Specificity	Accuracy	Processing Time, s
SMOTE Neural Network	95.7	92.4	92.7	92.4	92.4	92.4	1.44

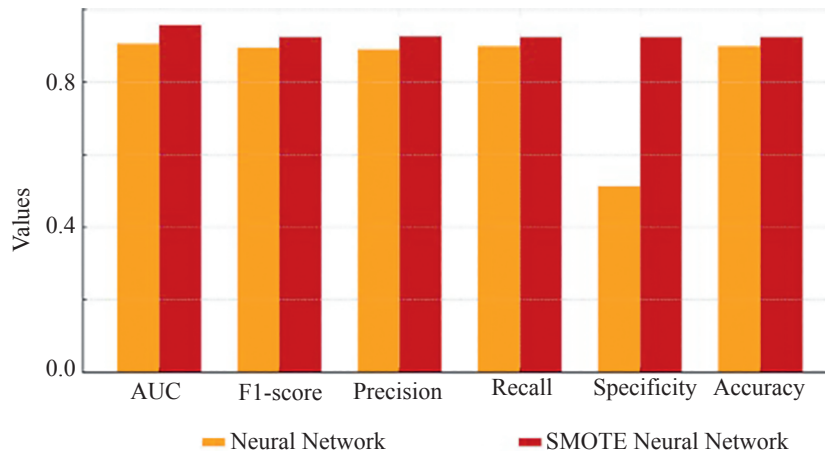


Fig. 8. Overall comparison metrics of Neural Network and SMOTENN

According to Precision metric, it is found that the Precision value climbed from 0.891 to 0.927, marking a 4.04 % improvement. This shows the SMOTENN model increased accuracy in correctly predicting positive instances.

Based on Recall metric, we see the Recall metric before SMOTE was 0.900 and after applying SMOTE it become 0.924. It is a 2.67 % increase for new SMOTENN model enhanced ability to capture positive cases.

Specificity shows the most significant improvement among other metrics. Before applying SMOTE, specificity was only 0.512. After SMOTE, it surged to 0.924, representing a dramatic improvement of approximately 80.47 %. This significant increase underscores SMOTENN model effectiveness in identifying fraudulent activities.

Then the accuracy aspect also shows significant rise. Before, the accuracy was only 0.900, then it jumps to 0.924 or gives a modest 2.67 % increase after applying SMOTE in tackling imbalance dataset. This specific metric result indicates better overall performance for our new SMOTENN model in detecting fraudulent activities in financial transactions.

Lastly, there is a slight increase of processing time before and after SMOTE implementation. The processing time required for SMOTENN model increased slightly, from 1.07 s to 1.44 s, reflecting an expected computational cost due to the augmented dataset size after SMOTE application. However, this difference, a mere 0.37 s, is negligible and does not pose a significant issue, particularly when weighed against the benefits SMOTE offers in addressing class imbalance.

Overall, these results confirm that our new SMOTENN model is effective to be applied in financial fraud detection where minority class identification is crucial.

In the future, next research could explore under-sampling method for balancing the dataset, or applying other hybrid approaches such as ensemble methods or cost-sensitive learning to further enhance these improvements while maintaining robustness across diverse datasets.

Conclusion

Frauds in financial institutions are considered serious problems by the whole financial industry. Sophisticated technologies must be incorporated to deal with the problem since traditional approaches are incapable of catching up with such emerging crimes. Therefore, artificial intelligence approaches such as Machine Learning (ML) are relevant to be proposed in dealing with fraud issues in the recent interconnected financial industry. The availability of historical dataset is prominent to enable ML techniques, showing off their capabilities. However, the proportion of fraud data is often very minimal in comparison to normal data of financial transactions. As a result, an imbalanced dataset issue occurs, which results in poor performance of ML.

In order to accurately identify any possible fraudulent financial activities, this study suggests a new ML model, namely SMOTENN. The model is integration between the Synthetic Minority Over-sampling Technique (SMOTE) dealing with imbalanced dataset and the Neural Network algorithm, a supervised ML algorithm. The new SMOTENN model clarifies that by fixing the imbalanced dataset, the performance could be significantly improved as seen in all evaluation metrics, particularly the specificity metric which accounted for a sharp 80.47 % increase.

References

1. Hidajat T. Rural banks fraud: a story from Indonesia. *Journal of Financial Crime*, 2020, vol. 27, no. 3, pp. 933–943. <https://doi.org/10.1108/jfc-01-2020-0010>
2. Syahria R. Detecting financial statement fraud using fraud diamond (A study on banking companies listed on the Indonesia Stock Exchange period 2012–2016). *Asia Pacific Fraud Journal*, 2019, vol. 4, no. 2, pp. 183–190. <https://doi.org/10.21532/apfjournal.v4i2.114>
3. Nabila P., Omsa S., Bravely I. Financial performance analysis of KSP Berkas Bulukumba in year 2018–2022. *AKUNSIKA: Jurnal Akuntansi dan Keuangan*, 2025, vol. 6, no. 1, pp. 10–20. <https://doi.org/10.31963/akunsika.v6i1.4902>
4. Tjambolang T.A., Radjab E., Hamid A. Financial technology dan gaya hidup dalam perilaku keuangan mahasiswa Politeknik Negeri Ujung Pandang. *Proc. of the Seminar Nasional Terapan Riset Inovatif (SENTRINOV)*, 2023, vol. 9, no. 2, pp. 355–362. (in Indonesian)
5. Omsa S., Hasan D., Bravely I., Suryadi A.S., Ischika A.P. Peningkatan kinerja keuangan “Anisah Catering” melalui peningkatan manajemen usaha mikro. *Proc. of the Seminar Nasional Terapan Riset Inovatif (SENTRINOV)*, 2023, vol. 9, no. 3, pp. 66–73. (in Indonesian)
6. Lim K.S., Lee L.H., Sim Y.W. A review of machine learning algorithms for fraud detection in credit card transaction. *International Journal of Computer Science & Network Security*, 2021, vol. 21, no. 9, pp. 31–40.
7. Ali A., Abd Razak S., Othman S.H., Eisa T.A.E., Al-Dhaqm A., Nasser M., Elhassan T., Elshafie H., Saif A. Financial fraud detection based on machine learning: a systematic literature review. *Applied Sciences*, 2022, vol. 12, no. 19, pp. 9637. <https://doi.org/10.3390/app12199637>
8. Syamsuddin I., Hwang J. The application of AHP model to guide decision makers: a case study of e-banking security. *Proc. of the Fourth International Conference on Computer Sciences and Convergence Information Technology*, 2009, pp. 1469–1473. <https://doi.org/10.1109/iccit.2009.251>
9. Gupta A., Lohani M.C., Manchanda M. Financial fraud detection using naive bayes algorithm in highly imbalance data set. *Journal of Discrete Mathematical Sciences and Cryptography*, 2021, vol. 24, no. 5, pp. 1559–1572. <https://doi.org/10.1080/09720529.2021.1969733>
10. El-Naby A., Hemdan E.E.D., El-Sayed A. An efficient fraud detection framework with credit card imbalanced data in financial services. *Multimedia Tools and Applications*, 2023, vol. 82, no. 3, pp. 4139–4160. <https://doi.org/10.1007/s11042-022-13434-6>
11. Tebala D., Marino D. Companies and artificial intelligence: an example of clustering with orange. *Studies in Systems, Decision and Control*, 2023, vol. 222, pp. 1–12. https://doi.org/10.1007/978-3-031-33461-0_1
12. Raqib F., Dunne M., Gurney J., Harle D.E., Sivapalan T., Sabokbar N., Bhogal-Bhamra G.K. Translational learning with orange data mining. *Proc. of the 11th International Conference on Research Advancement Resilience in the Pandemic Era: A Drive for Innovative Transformation*, 2021.
13. Ariyanto R., Bone H. Fraud awareness in Indonesian governmental sector: Multi-agency responses. *Review of Integrative Business and Economics Research*, 2020, vol. 9, no. 2, pp. 209–222.
14. Ikbal M., Irwansyah I., Paminto A., Ulfah Y., Darma D.C. Financial intelligence: Financial statement fraud in Indonesia. *Journal of Intelligence Studies in Business*, 2020, vol. 10, no. 3, pp. 80–95. <https://doi.org/10.37380/jisib.v10i3.640>
15. Taneja S., Suri B., Kothari C. Application of balancing techniques with ensemble approach for credit card fraud detection. *Proc. of the International Conference on Computing, Power and Communication Technologies (GUCON)*, 2019, pp. 753–758.
16. Saragih M.G., Chin J., Setyawasih R., Nguyen P.T., Shankar K. Machine learning methods for analysis fraud credit card transaction. *International Journal of Engineering and Advanced Technology (IJEAT)*, 2019, vol. 8, no. 6 S, pp. 870–874.
17. Hashemi S.K., Mirtaheeri S.L., Greco S. Fraud detection in banking data by machine learning techniques. *IEEE Access*, 2023, vol. 11, pp. 3034–3043. <https://doi.org/10.1109/ACCESS.2022.3232287>
18. Tadv F., Shinde S., Patil D., Dmello S. Real time credit card fraud detection. *International Research Journal of Engineering and Technology*, 2021, vol. 8, no. 5, pp. 2177–2180.
19. Chawla N.V., Bowyer K.W., Hall L.O., Kegelmeyer W.P. SMOTE: synthetic minority over-sampling technique. *Journal of Artificial*

Литература

1. Hidajat T. Rural banks fraud: a story from Indonesia // *Journal of Financial Crime*. 2020. V. 27. N 3. P. 933–943. <https://doi.org/10.1108/jfc-01-2020-0010>
2. Syahria R. Detecting financial statement fraud using fraud diamond (A study on banking companies listed on the Indonesia Stock Exchange period 2012–2016) // *Asia Pacific Fraud Journal*. 2019. V. 4. N 2. P. 183–190. <https://doi.org/10.21532/apfjournal.v4i2.114>
3. Nabila P., Omsa S., Bravely I. Financial performance analysis of KSP Berkas Bulukumba in year 2018–2022 // *AKUNSIKA: Jurnal Akuntansi dan Keuangan*. 2025. V. 6. N 1. P. 10–20. <https://doi.org/10.31963/akunsika.v6i1.4902>
4. Tjambolang T.A., Radjab E., Hamid A. Financial technology dan gaya hidup dalam perilaku keuangan mahasiswa Politeknik Negeri Ujung Pandang // *Proc. of the Seminar Nasional Terapan Riset Inovatif (SENTRINOV)*. 2023. V. 9. N 2. P. 355–362.
5. Omsa S., Hasan D., Bravely I., Suryadi A.S., Ischika A.P. Peningkatan kinerja keuangan “Anisah Catering” melalui peningkatan manajemen usaha mikro // *Proc. of the Seminar Nasional Terapan Riset Inovatif (SENTRINOV)*. 2023. V. 9. N 3. P. 66–73.
6. Lim K.S., Lee L.H., Sim Y.W. A review of machine learning algorithms for fraud detection in credit card transaction // *International Journal of Computer Science & Network Security*. 2021. V. 21. N 9. P. 31–40.
7. Ali A., Abd Razak S., Othman S.H., Eisa T.A.E., Al-Dhaqm A., Nasser M., Elhassan T., Elshafie H., Saif A. Financial fraud detection based on machine learning: a systematic literature review // *Applied Sciences*. 2022. V. 12. N 19. P. 9637. <https://doi.org/10.3390/app12199637>
8. Syamsuddin I., Hwang J. The application of AHP model to guide decision makers: a case study of e-banking security // *Proc. of the Fourth International Conference on Computer Sciences and Convergence Information Technology*. 2009. P. 1469–1473. <https://doi.org/10.1109/iccit.2009.251>
9. Gupta A., Lohani M.C., Manchanda M. Financial fraud detection using naive bayes algorithm in highly imbalance data set // *Journal of Discrete Mathematical Sciences and Cryptography*. 2021. V. 24. N 5. P. 1559–1572. <https://doi.org/10.1080/09720529.2021.1969733>
10. El-Naby A., Hemdan E.E.D., El-Sayed A. An efficient fraud detection framework with credit card imbalanced data in financial services // *Multimedia Tools and Applications*. 2023. V. 82. N 3. P. 4139–4160. <https://doi.org/10.1007/s11042-022-13434-6>
11. Tebala D., Marino D. Companies and artificial intelligence: an example of clustering with orange // *Studies in Systems, Decision and Control*. 2023. V. 222. P. 1–12. https://doi.org/10.1007/978-3-031-33461-0_1
12. Raqib F., Dunne M., Gurney J., Harle D.E., Sivapalan T., Sabokbar N., Bhogal-Bhamra G.K. Translational learning with orange data mining // *Proc. of the 11th International Conference on Research Advancement Resilience in the Pandemic Era: A Drive for Innovative Transformation*. 2021.
13. Ariyanto R., Bone H. Fraud awareness in Indonesian governmental sector: Multi-agency responses // *Review of Integrative Business and Economics Research*. 2020. V. 9. N 2. P. 209–222.
14. Ikbal M., Irwansyah I., Paminto A., Ulfah Y., Darma D.C. Financial intelligence: Financial statement fraud in Indonesia // *Journal of Intelligence Studies in Business*. 2020. V. 10. N 3. P. 80–95. <https://doi.org/10.37380/jisib.v10i3.640>
15. Taneja S., Suri B., Kothari C. Application of balancing techniques with ensemble approach for credit card fraud detection // *Proc. of the International Conference on Computing, Power and Communication Technologies (GUCON)*. 2019. P. 753–758.
16. Saragih M.G., Chin J., Setyawasih R., Nguyen P.T., Shankar K. Machine learning methods for analysis fraud credit card transaction // *International Journal of Engineering and Advanced Technology (IJEAT)*. 2019. V. 8. N 6 S. P. 870–874.
17. Hashemi S.K., Mirtaheeri S.L., Greco S. Fraud detection in banking data by machine learning techniques // *IEEE Access*. 2023. V. 11. P. 3034–3043. <https://doi.org/10.1109/ACCESS.2022.3232287>
18. Tadv F., Shinde S., Patil D., Dmello S. Real time credit card fraud detection // *International Research Journal of Engineering and Technology*. 2021. V. 8. N 5. P. 2177–2180.
19. Chawla N.V., Bowyer K.W., Hall L.O., Kegelmeyer W.P. SMOTE: synthetic minority over-sampling technique // *Journal of Artificial Intelligence Research*. 2002. V. 16. P. 321–357. <https://doi.org/10.1613/jair.953>

- Intelligence Research*, 2002, vol. 16, pp. 321–357. <https://doi.org/10.1613/jair.953>
20. Fernández A., Garcia S., Herrera F., Chawla N.V. SMOTE for learning from imbalanced data: progress and challenges, marking the 15-year anniversary. *Journal of Artificial Intelligence Research*, 2018, vol. 61, pp. 863–905. <https://doi.org/10.1613/jair.1.11192>
 21. Elreedy D., Atiya A.F., Kamalov F. A theoretical distribution analysis of synthetic minority oversampling technique (SMOTE) for imbalanced learning. *Machine Learning*, 2024, vol. 113, no. 7, pp. 4903–4923. <https://doi.org/10.1007/s10994-022-06296-4>
 22. Veigas K.C., Regulagadda D.S., Kokatnoor S.A. Optimized stacking ensemble (OSE) for credit card fraud detection using synthetic minority oversampling model. *Indian Journal of Science and Technology*, 2021, vol. 14, no. 32, pp. 2607–2615. <https://doi.org/10.17485/ijst/v14i32.807>
 23. Syamsuddin I., Barukab O.M. SUKRY: Suricata IDS with enhanced kNN algorithm on Raspberry Pi for classifying IoT botnet attacks. *Electronics*, 2022, vol. 11, no. 5, pp. 737. <https://doi.org/10.3390/electronics11050737>
 24. Makki S., Assaghir Z., Taher Y., Haque R., Hacid M.S., Zeineddine H. An experimental study with imbalanced classification approaches for credit card fraud detection. *IEEE Access*, 2019, vol. 7, pp. 93010–93022. <https://doi.org/10.1109/ACCESS.2019.2927266>
 25. Alam T.M., Shaukat K., Hameed I.A., Luo S.H., Sarwar M.U., Shabbir S., Li J.M., Khushi M. An investigation of credit card default prediction in the imbalanced datasets. *IEEE Access*, 2020, vol. 8, pp. 201173–201198. <https://doi.org/10.1109/access.2020.3033784>
 26. Moro S., Cortez P., Rita P. A data-driven approach to predict the success of bank telemarketing. *Decision Support Systems*, 2014, vol. 62, pp. 22–31. <https://doi.org/10.1016/j.dss.2014.03.001>
 27. Kawam A.A., Mansour N. Metaheuristic optimization algorithms for training artificial neural networks. *International Journal of Computer and Information Technology*, 2012, vol. 1, no. 2, pp. 156–161.
 28. Quiroga F.M. Invariance and same-equivariance measures for convolutional neural networks. *CLEI Electronic Journal*, 2021, vol. 24, no. 1 SI. <https://doi.org/10.19153/cleiej.24.1.8>
 29. Alla H., Moumoun L., Balouki Y. A multilayer perceptron neural network with selective-data training for flight arrival delay prediction. *Scientific Programming*, 2021, vol. 2021, pp. 5558918. <https://doi.org/10.1155/2021/5558918>
 30. Fernández A., Garcia S., Herrera F., Chawla N.V. SMOTE for learning from imbalanced data: progress and challenges, marking the 15-year anniversary // *Journal of Artificial Intelligence Research*. 2018. V. 61. P. 863–905. <https://doi.org/10.1613/jair.1.11192>
 31. Elreedy D., Atiya A.F., Kamalov F. A theoretical distribution analysis of synthetic minority oversampling technique (SMOTE) for imbalanced learning // *Machine Learning*. 2024. V. 113. N 7. P. 4903–4923. <https://doi.org/10.1007/s10994-022-06296-4>
 32. Veigas K.C., Regulagadda D.S., Kokatnoor S.A. Optimized stacking ensemble (OSE) for credit card fraud detection using synthetic minority oversampling model // *Indian Journal of Science and Technology*. 2021. V. 14. N 32. P. 2607–2615. <https://doi.org/10.17485/ijst/v14i32.807>
 33. Syamsuddin I., Barukab O.M. SUKRY: Suricata IDS with enhanced kNN algorithm on Raspberry Pi for classifying IoT botnet attacks // *Electronics*. 2022. V. 11. N 5. P. 737. <https://doi.org/10.3390/electronics11050737>
 34. Makki S., Assaghir Z., Taher Y., Haque R., Hacid M.S., Zeineddine H. An experimental study with imbalanced classification approaches for credit card fraud detection // *IEEE Access*. 2019. V. 7. P. 93010–93022. <https://doi.org/10.1109/ACCESS.2019.2927266>
 35. Alam T.M., Shaukat K., Hameed I.A., Luo S.H., Sarwar M.U., Shabbir S., Li J.M., Khushi M. An investigation of credit card default prediction in the imbalanced datasets. *IEEE Access*. 2020. V. 8. P. 201173–201198. <https://doi.org/10.1109/access.2020.3033784>
 36. Moro S., Cortez P., Rita P. A data-driven approach to predict the success of bank telemarketing // *Decision Support Systems*. 2014. V. 62. P. 22–31. <https://doi.org/10.1016/j.dss.2014.03.001>
 37. Kawam A.A., Mansour N. Metaheuristic optimization algorithms for training artificial neural networks // *International Journal of Computer and Information Technology*. 2012. V. 1. N 2. P. 156–161.
 38. Quiroga F.M. Invariance and same-equivariance measures for convolutional neural networks // *CLEI Electronic Journal*. 2021. V. 24. N 1 SI. <https://doi.org/10.19153/cleiej.24.1.8>
 39. Alla H., Moumoun L., Balouki Y. A multilayer perceptron neural network with selective-data training for flight arrival delay prediction // *Scientific Programming*. 2021. V. 2021. P. 5558918. <https://doi.org/10.1155/2021/5558918>

Authors

Irfan Syamsuddin — Professor, Politeknik Negeri Ujung Pandang, Makassar, 90245, Indonesia, [sc 25927526500](https://orcid.org/0000-0002-6017-7364), <https://orcid.org/0000-0002-6017-7364>, irfans@poliupg.ac.id

Sirajuddin Omsa — Associate Professor, Politeknik Negeri Ujung Pandang, Makassar, 90245, Indonesia, <https://orcid.org/0000-0003-2776-1531>, sirajud_om@poliupg.ac.id

Andi Rustam — Associate Professor, Universitas Muhammadiyah Makassar, Makassar, 90221, Indonesia, [sc 25927526500](https://orcid.org/0009-0000-2082-4339), <https://orcid.org/0009-0000-2082-4339>, a.rustam@unismuh.ac.id

Dahsan Hasan — Associate Professor, Politeknik Negeri Ujung Pandang, Makassar, 90245, Indonesia, <https://orcid.org/0009-0002-6612-0104>, dahsan.hasan@poliupg.ac.id

Received 12.01.2025

Approved after reviewing 02.05.2025

Accepted 24.05.2025

Авторы

Сямеуддин Ирфан — PhD, профессор, Государственный политехнический институт Уджунг Панданга, Макаassar, 90245, Индонезия, [sc 25927526500](https://orcid.org/0000-0002-6017-7364), <https://orcid.org/0000-0002-6017-7364>, irfans@poliupg.ac.id

Омса Сираджуддин — PhD, доцент, Государственный политехнический институт Уджунг Панданга, Макаassar, 90245, Индонезия, <https://orcid.org/0000-0003-2776-1531>, sirajud_om@poliupg.ac.id

Рустам Анди — доцент, Университет Мухаммадии в Макассаре, Макаassar, 90221, Индонезия, [sc 25927526500](https://orcid.org/0009-0000-2082-4339), <https://orcid.org/0009-0000-2082-4339>, a.rustam@unismuh.ac.id

Хасан Дахсан — PhD, доцент, Государственный политехнический институт Уджунг Панданга, Макаassar, 90245, Индонезия, <https://orcid.org/0009-0002-6612-0104>, dahsan.hasan@poliupg.ac.id

Статья поступила в редакцию 12.01.2025

Одобрена после рецензирования 02.05.2025

Принята к печати 24.05.2025



Работа доступна по лицензии
Creative Commons
«Attribution-NonCommercial»

doi: 10.17586/2226-1494-2025-25-3-475-486

A deep learning approach for adaptive electrocardiogram-based authentication in an internet of things enabled telehealth system

Mohamed Abdalla Elsayed Azab✉

ITMO University, Saint Petersburg, 197101, Russian Federation
 mohamed.a.azab@itmo.ru✉, <https://orcid.org/0009-0000-1748-0029>

Abstract

As telehealth services have become integral to healthcare applications; robust authentication mechanisms are critical for safeguarding sensitive patient data and services. Conventional authentication techniques including passwords and tokens are susceptible to theft and security breaches. This vulnerability highlights the need for alternative methods that offer improved security measures and ease of use. Biometric authentication, which leverages unique physical and behavioral traits, has emerged as a promising alternative. Among various biometric modalities, electrocardiogram (ECG) signals stand out because of their uniqueness, stability, and noninvasive nature. This study introduces an innovative deep-learning-based authentication system that utilizes ECG signals to enhance security in Internet of Things (IoT)-powered telehealth environments. The proposed model employs hybrid architecture, starting with a Siamese Neural Network (SNN) for dynamic verification, followed by a Convolutional Neural Network (CNN) for feature extraction, utilizing an optimized Sequential Beat Aggregation approach for robust ECG-based authentication. The system operates securely and adaptively, and performs real-time authentication without requiring human intervention. The research approach involved the acquisition and processing of electrocardiogram data from the ECG-ID dataset which encompassed 310 ECG individuals obtained from 90 individual subjects. This dataset provided a comprehensive set of samples for training and evaluation. The model achieved high authentication accuracy (98.5 %–99.5 %) and a false acceptance rate of 0.1 % with minimal computational overhead, validating its feasibility for real-time applications. This study integrates ECG-based authentication into telehealth systems, creating a secure foundation for safeguarding patient data. The innovative use of ECG signals advances secure and adaptable for a personalized remote health monitoring system development.

Keywords

biometric authentication, telehealth security, adaptive authentication systems, ECG signal analysis, neural networks, hybrid deep learning

For citation: Azab M.A.E. A deep learning approach for adaptive electrocardiogram-based authentication in an internet of things enabled telehealth system. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2025, vol. 25, no. 3, pp. 475–486. doi: 10.17586/2226-1494-2025-25-3-475-486

УДК 004.056

Глубокое обучение для адаптивной аутентификации на основе электрокардиограммы в системе телемедицины с поддержкой интернета вещей

Мохамед Абдалла Эльсайед Азаб✉

Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация
 mohamed.a.azab@itmo.ru✉, <https://orcid.org/0009-0000-1748-0029>

Аннотация

Телемедицинские услуги стали неотъемлемой частью медицинских приложений, надежные механизмы аутентификации имеют решающее значение для защиты конфиденциальных данных пациентов и предоставляемых сервисов. Традиционные методы аутентификации, такие как пароли и токены, подвержены кражам и нарушениям безопасности. Эта уязвимость подчеркивает необходимость альтернативных методов, которые обеспечивают более высокий уровень безопасности и удобство использования. Биометрическая

© Azab M.A.E., 2025

аутентификация, использующая уникальные физические и поведенческие характеристики, стала перспективной альтернативой. Среди различных биометрических методов сигналы электрокардиограммы (ЭКГ) выделяются своей уникальностью, стабильностью и неинвазивным характером. В данном исследовании представлен инновационный метод аутентификации на основе глубокого обучения, использующий ЭКГ-сигналы для повышения уровня безопасности в телемедицинских системах, работающих на базе интернета вещей (IoT). Предложенная модель использует гибридную архитектуру: сначала применяется сиамская нейронная сеть (SNN) для динамической верификации, затем сверточная нейронная сеть (CNN) для извлечения признаков с использованием оптимизированного метода последовательной агрегации сердечных циклов для надежной аутентификации на основе ЭКГ. Система функционирует безопасно и адаптивно, выполняя аутентификацию в реальном времени без вмешательства человека. В рамках исследования была проведена обработка данных ЭКГ из набора данных ECG-ID, включающего 310 ЭКГ-сигналов от 90 различных участников. Этот набор данных предоставил обширную выборку для обучения и оценки. Модель достигла высокой точности аутентификации (98,5–99,5 %) и показателя ложного допуска на уровне 0,1 % при минимальной вычислительной нагрузке, что подтверждает ее применимость для задач в реальном времени. Настоящее исследование интегрирует аутентификацию на основе ЭКГ в телемедицинские системы, создавая надежную основу для защиты данных пациентов. Инновационное использование ЭКГ-сигналов способствует созданию безопасной, адаптивной и персонализированной системы удаленного мониторинга здоровья.

Ключевые слова

биометрическая аутентификация, безопасность телемедицины, адаптивные системы аутентификации, анализ ЭКГ-сигналов, нейронные сети, гибридное глубокое обучение

Ссылка для цитирования: Азаб М.А.Э. Глубокое обучение для адаптивной аутентификации на основе электрокардиограммы в системе телемедицины с поддержкой интернета вещей // Научно-технический вестник информационных технологий, механики и оптики. 2025. Т. 25, № 3. С. 475–486 (на англ. яз.). doi: 10.17586/2226-1494-2025-25-3-475-486

Introduction

With the growing adoption of telehealth systems, the demand for secure and reliable management of patient health information has significantly increased. IoT-enabled telehealth systems have become integral in enhancing patient care, enabling remote monitoring, and streamlining healthcare delivery [1, 2]. However, the spread of connected devices has also introduced significant security challenges, particularly in safeguarding sensitive patient data and ensuring secure access to telehealth services. Traditional authentication methods, including passwords and tokens, are increasingly inadequate owing to vulnerabilities, such as theft, replication, and brute-force attacks [3]. In the face of the growing threat landscape, there is an urgent need for new, robust, and flexible authentication systems aimed at addressing the unique needs of Internet of Things (IoT) telehealth environments [4].

Biometric authentication, which leverages unique physiological and behavioral traits, has emerged as a promising solution for securing IoT-enabled systems and providing enhanced security and user convenience. Unlike traditional methods, biometric systems leverage unique physiological or behavioral traits, making it difficult to replicate or falsify [3]. Among the various biometric modalities, electrocardiogram (ECG) signals stand out because of their intrinsic liveness, universality, and resistance to spoofing. ECG signals capture the electrical activity of the heart, offering a dynamic and highly individualized biometric pattern [5]. This unique property makes ECG-based systems robust alternatives for secure and reliable user authentication in IoT-enabled telehealth systems.

ECG signals are private and accurate and more reliable than other biometrics [6]. Such P-Q-R-S-T waveform patterns, as shown in Fig. 1 provide proof of the subject liveness and uniqueness. The nature of ECG signals is

complex because of their random, involuntary, and complex characteristics leading to better authentication capabilities [7].

Artificial intelligence improves IoT-enabled telehealth security by enabling reliable ECG-based biometric authentication. Machine learning techniques like K-Nearest Neighbor, Support Vector Machine and Random Forest are used, with deep learning increasingly favored for its pattern analysis and accuracy [8, 9]. Deep Learning techniques, particularly neural networks, have emerged as a possible telehealth ECG signal authentication solution. Using huge datasets of ECG signals, neural networks can be trained to understand the distinctive patterns and properties that distinguish real signals from fraudulent ones [10, 11].

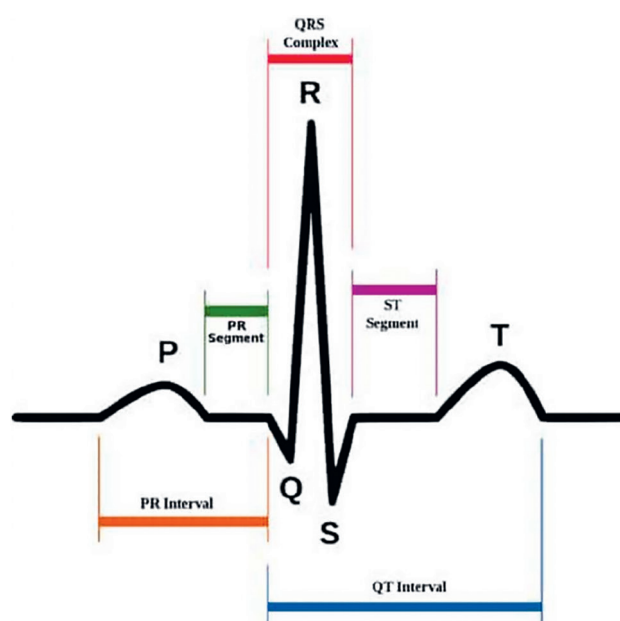


Fig. 1. Segment Representation of ECG

This study addresses the growing demand for robust authentication systems amid rising digital reliance and data security challenges. By leveraging deep learning, it proposes an adaptive authentication framework that combines Siamese Neural Networks (SNNs) for verification and Convolutional Neural Networks (CNNs) for multiclass authentication, utilizing an optimized Sequential Beat Aggregation approach. The system enhances ECG biometrics, improves authentication speed and efficiency, and detects noise-inducing abnormalities, addressing key IoT telehealth challenges, such as evolving threat landscapes, dynamic user behavior, and real-time processing requirements.

Our study presents three key contributions: showing the use of ECG signals for adaptive authentication in IoT telehealth systems, developing advanced deep learning models that merge biometric verification with authentication, and assessing the system resilience under various signal conditions to enhance security and accessibility in IoT-based healthcare environments.

Literature Survey

ECG Authentication Using Machine Learning and Deep Learning Models

Several studies have explored the use of Machine Learning and Deep Learning models for ECG-based authentication. Asadian et al. [12] examined the feasibility of ECG-based authentication systems, while Shdefat et al. [13] discussed opportunities and challenges in their implementation. Lin Li et al. [14] extended the literature by discussing the application of ECGs, Electroencephalography, and Photoplethysmogram in authentication systems. Pereira et al. [15] researched different ways of capturing data to make the authentication process more efficient in terms of its precision.

Hammad et al. [16] proposed using Deep Neural Networks with architectures such as Residual Convolutional Neural Networks (ResNet) and end-to-end CNNs for reliable human authentication. Using datasets like Physikalisch-Technische Bundesanstalt (PTB) and Check Your Bio-Signals Here, their models achieved an average accuracy of 98.5 %. Similarly, Labati et al. [17] introduced Deep-ECG, a biometric recognition method combining signal preprocessing, CNN feature extraction, and SoftMax-based detection. Their work demonstrated improved performance in scenarios such as identity verification and periodic re-authentication, outperforming earlier approaches.

Hybrid and Advanced Architectures

Martin et al. [18] proposed “BioECG” which integrates CNN and Long short-term memory model to enhance authentication precision. Their method emphasized the importance of addressing temporal dependencies in ECG signals, resulting in improved accuracy and robustness. D’angelis et al. [19] employed Vision Transformers for ECG biometric recognition, effectively capturing intricate temporal and spatial signal features. AlDuwaile and Islam [20] utilized single heartbeat analysis with CNNs, simplifying data collection while maintaining high accuracy.

Some of the studies also highlighted novel hybrid system models. For instance, Ivanciu et al. [21] implemented a SNN using ECG signal images, achieving an accuracy of 87.3 %. Albuquerque et al. [22] employed Random Under-Sampling Boosting and Nearest Neighbor Search, achieving accuracy rates of 97.4 % and 99.5 %, respectively, for ECG-based user identification. These works highlight the potential of hybrid models for improving authentication accuracy and adaptability.

Multi-Modal Biometric Systems

In order to overcome the above-mentioned limitations of uni-modal systems, Alkeem et al. [23] introduced a multi-modal biometric system that integrates ECG signals, facial images, and fingerprints. The system utilized multitasking learning and feature fusion, demonstrating superior accuracy and generalization. Multi-modal methods outperformed single-modal approaches, proving highly effective for secure authentication and gender classification.

Emerging Techniques and Applications

Recent advancements have introduced novel techniques such as lightweight multi-factor authentication strategies [23] which incorporate digital signatures and device capabilities to enhance IoT security. Similarly, studies [24] tailored authentication methods for the Internet of Medical Things, utilizing human biometrics to establish secure device communication. Blockchain-based solutions have also emerged, such as the “Bubbles-of-Trust” scheme [25], which employs virtual trust zones and Ethereum blockchain technology to streamline authentication in IoT networks [26–28].

Limitations and Gaps in Existing Research

While the reviewed studies demonstrate significant advancements in ECG biometrics, several limitations persist. Many approaches [16, 17, 29] focus primarily on accuracy, often neglecting computational efficiency and robustness against noisy or heterogeneous datasets. Additionally, while methods integrating CNN and LSTM architectures [18, 20] show promise, they often fail to fully leverage the potential of these models for capturing complex temporal dynamics in ECG signals. Moreover, scalability and real-time processing remain challenges in deploying these systems for practical IoT telehealth applications.

Proposed Methodology

In this paper, we have proposed an innovative approach to use deep learning-based ECG in IoT-based telehealth systems for patient or user authentication and attempted to develop a model. By passing through steps like: data gathering, preparation, model creation, training, testing, tuning, implementation and monitoring it ensures security and flexibility. Fig. 2 shows how this process helps solve authentication problems.

The proposed model

Our Adaptive Authentication System is built on SNN and CNN as foundational models for ECG signal analysis. SNN model is for time related change, and CNN model is to interpret the pattern over the signals. Being trained on thousands of highly processed ECG data, and continuing to show robust performance in test data with accuracy,



Fig. 2. Deep learning process for adaptive authentication

sensitivity, precision, F1-score, Area Under the Curve (AUC), specificity. The system improves incrementally with small changes contributed over time. It also keeps a good balance between being secure and easy to use, offering a safe and practical solution for users.

The core feature of this model is its adaptability to dynamic conditions, making it particularly suitable for IoT-enabled telehealth environments. The deep learning mechanisms within the architecture are designed to adjust to evolving threat landscapes and changing user behaviors, ensuring high levels of dependability and efficiency [30]. The model uses Sequential Beat Aggregation, training with individual heartbeats and aggregating predictions from multiple beats during inference, balancing simplicity with accuracy while reducing complexity and data demands of longer concatenated signals.

Sequential Beat Aggregation approach actually improves robustness by reducing anomalous beats impact, ensuring real-time adaptability essential for IoT telehealth. It can be seamlessly integrated with IoT devices to work efficiently in the real world. The lightweight, adaptive, and fast design of this system makes it an excellent choice for secure and practical telehealth deployments and applications.

To provide a comprehensive overview, Fig. 3 explains the methodology of the proposed model, highlighting the integration of data processing, model architecture, and deployment strategies.

The ECG signal first enters through the SNN which mainly focuses on capturing the changes over time and important details. The result is then sent to the CNN where layers find patterns in the signal and reduce its size for quicker processing. This step-by-step method improves the ability to identify key features. The CNN result is turned into a single line of data and sent to a dense layer, which prepares it for further processing. Lastly, a classifier at the end of the

model turns the scores into a format that can be used for identifying different classes. The model learns from the key points in the ECG signal to tell the difference between real and fake signals, making sure it works well and accurately for user identification. This proposed design effectively combines speed and accuracy in classification tasks.

Convolution neural network layers

It is a feed-forward neural network that is very frequently used for medical analysis, object detection, face recognition, and picture classification [31]. The CNN network has a variety of designs, including VGG-Net [32], Inception [33], ResNet [16] DenseNet [34], and Xception Net [35]. Generally, the layers used for experimental analysis are the same in all CNN models. The first layer typically used to extract features from an ECG trace is the convolution layer. To achieve this, move a kernel or feature detector over the input data or feature vector and compute the dot product of the input and kernel at each location. After extracting the features, the network is made non-linear using activation function which also speeds up total computation. After the activation layer, a rectified feature map is put through the max pooling layer. The down-sampling process of pooling lowers the feature map size. Then, a shared feature map is sent to be flattened into a single, lengthy continuous linear vector. This study used six convolution layers and an identical number of pooling layers. Batch normalization was also performed in various layers to solve the covariate shift problem [23].

Fully connected layers

A dense block comprises interconnected hidden layers where each neuron receives input from all preceding neurons. Using matrix-vector multiplication, it adjusts output dimensionality and applies an activation function for accurate classification or prediction.

$$y = \text{activation}(z), \text{ where } \mathbf{z} = \mathbf{W} \times \mathbf{x} + \mathbf{b},$$

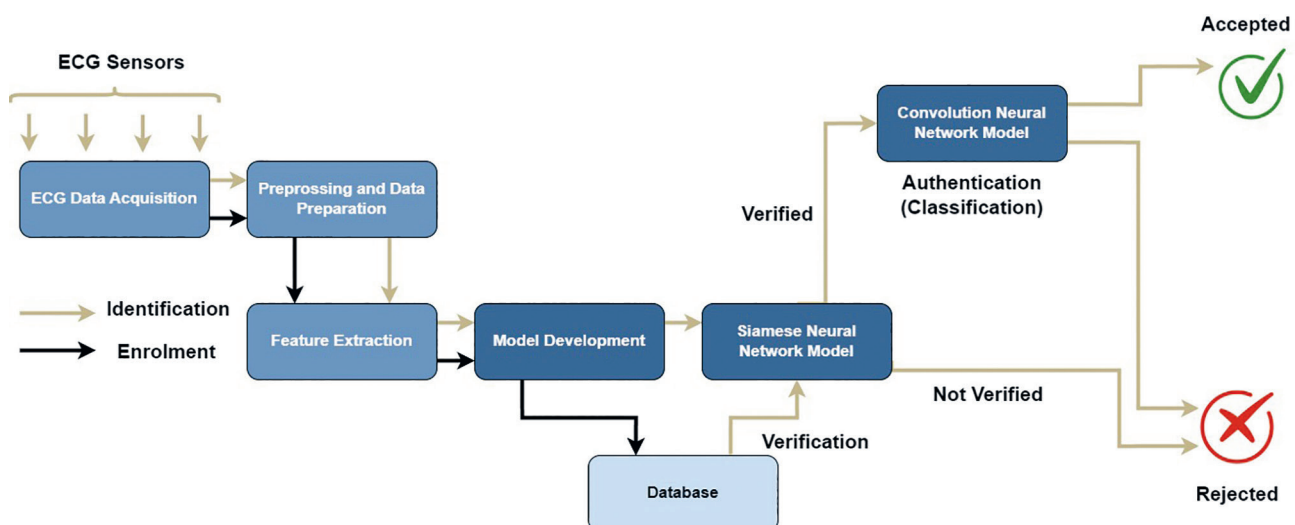


Fig. 3. Adaptive Authentication Biometric System Proposed Model

where \mathbf{z} is the pre-activation vector and activation represents the activation function σ . Dense layers are integral to neural networks, connecting inputs to outputs for precise predictions. For ECG-based authentication, a single dense layer with 286 neurons enables accurate classification. The setup is: Hybrid SNN-CNN output \rightarrow Flatten Layer \rightarrow Dense Layers \rightarrow Classifier. Fig. 4 illustrates the CNN architecture.

Threshold based-authentication algorithm

The probability-based threshold function is the probability of a match between a test signal and a stored reference signal, usually for signal processing applications like a biometric authentication system. It calculates the Probability Density Functions (PDF), of the test and reference signals for likelihoods of different values of the signal. Comparing the PDFs gives the match probability. The threshold value is determined based on the required security level and False Acceptance Rate (FAR). Increasing the threshold makes FAR higher but improves security, whereas decreasing it reduces both.

Loss function

This function measures an algorithm performance by quantifying discrepancies between predicted outputs and target values, often using cross-entropy for assessing classification models accuracy against actual labels. This study used Binary Cross Entropy (BCE) as a loss function for binary classification tasks like ECG data authentication [24]. In a hybrid SNN-CNN model for human authentication, BCE loss can help the general model focus on the essential ECG signal regions while ignoring the areas that are less important to the authentication goal. BCE will be used as an optimization method to discover the best values valid for human authentication based on ECG signals [25].

$$BCE = -\frac{1}{N} \sum_{i=0}^N y_i \log(\hat{y}_i) + (1 - y_i) \log(1 - \hat{y}_i).$$

Here, predictions and true values are represented by y_i and \hat{y}_i , respectively, where N denotes the number of training sets for the i -th slice ($i \in N$).

Hyperparameters settings

Hyperparameters optimize the training success of the hybrid SNN-CNN deep learning model for ECG-based authentication. Key parameters include the loss function, optimizer, batch size, learning rate, and epochs, significantly boosting efficiency while ensuring adaptability if the primary objective encounters challenges. Lists of the hyperparameters chosen for training SNN and CNN: loss functions — Categorical Cross Entropy; learning rate — 0.00001; batch size — 64; epochs — 100; optimizer — Adam.

Similarly, these hyperparameters are frequently adjusted using a grid search methodology to obtain the best possible combination for the training process. For instance, rapid convergence of the model that shoots past the global minimum may be caused by a high learning rate [36]. On the other hand, a slow learning rate may stall training, while too few epochs cause underfitting, and too many lead to overfitting. Proper parameter tuning enhances model accuracy and robustness.

Experiment Results

Dataset description

The ECG-ID database consists of 310 ECG recordings from 90 subjects, aged 13–75, including 44 men and 46 women. Each recording is a 20-second ECG Lead I signal, sampled at 500 Hz with 12-bit resolution and a nominal range of ± 10 mV. The dataset includes 2–20 recordings per individual, collected over a period of up to six months, with intervals ranging from days to months between sessions. This multi-session structure allows for the evaluation of temporal variability in ECG signals, making it suitable for assessing the long-term stability of ECG-based biometric systems.

The dataset provides both noisy raw signals (Signal 0) and filtered signals (Signal 1), enabling detailed analysis of ECG morphology under different noise conditions. The recordings were collected using standard ECG electrodes placed on the wrists, making the acquisition process non-invasive and suitable for real-world applications.

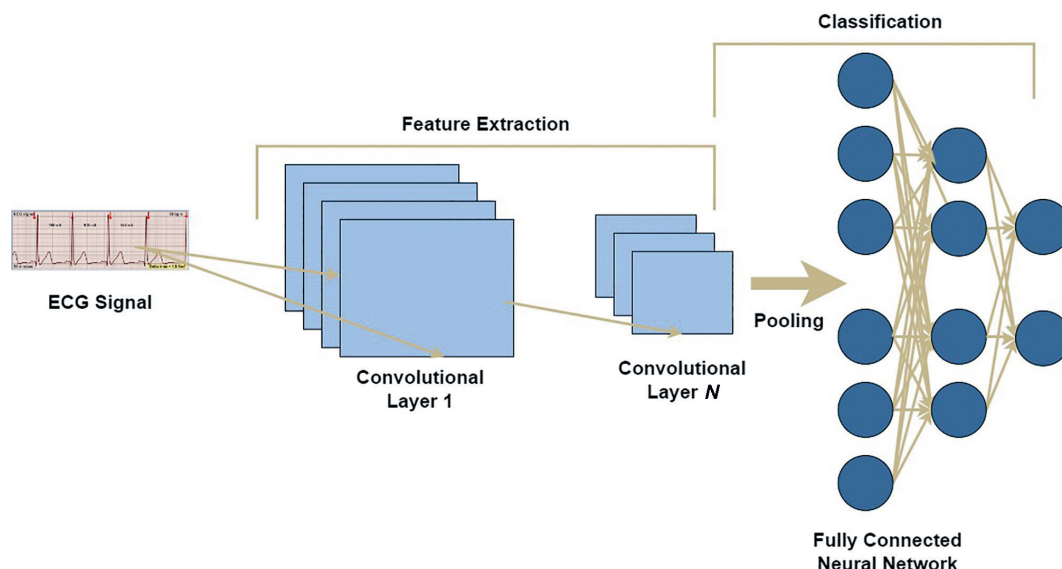


Fig. 4. Diagram of CNN Architecture

Dataset preprocessing and augmentation

Data preparation would enhance quality data analysis on ECG-ID dataset signals by effectively cleaning the initial raw signals. These are the high pass set up to remove, baseline drift and notch filter to remove power line interference along with resampling [37]. Preprocessing also standardizes the ECG signals, and quality control is used in order to filter artifacts also forms part of it. Data augmentation shown in Fig. 5 enlarges the dataset, increases the stability and optimizes performance through the introduction of variability. Scaling, rotation, flipping, and time-shifting are the techniques of diversification of the dataset. These approaches mainly address issues related with overfitting, boosting model accuracy and credibility, scalability, and, therefore, boosting the general performance of ECG-based biometrics. The figure displays the original signal (blue) alongside signals modified by three augmentation methods: noise addition (orange), time-shifting (green), and pitch-shifting (red).

Results

This section presents a detailed analysis and conclusions drawn from experimental evaluations, offering insights into the system performance across various scenarios. Both quantitative and qualitative data are thoroughly examined to highlight the system advantages, identify areas for improvement, and evaluate overall efficacy. The experiment utilized 1D ECG signals of length 256, with SNN and CNN models trained using a batch size of 64, a learning rate of 0.00001, and the Adam optimizer with decay.

Training Phase

During the training phase of SNN model, the ECG dataset underwent a preprocessing phase wherein the ECG segments were truncated to 700 per individual, resulting in a well-balanced distribution of classes. This step ensures an equal representation of different classes, providing a more robust foundation for subsequent analysis.

Following the truncation, the pre-processed data underwent normalization, transforming the values to a standardized range between 0 and 1. This normalization process is crucial for mitigating potential anomalies that could complicate signal analysis. By scaling the data to a uniform range, the impact of variations in magnitude is minimized, facilitating a more consistent and effective analysis of ECG signals.

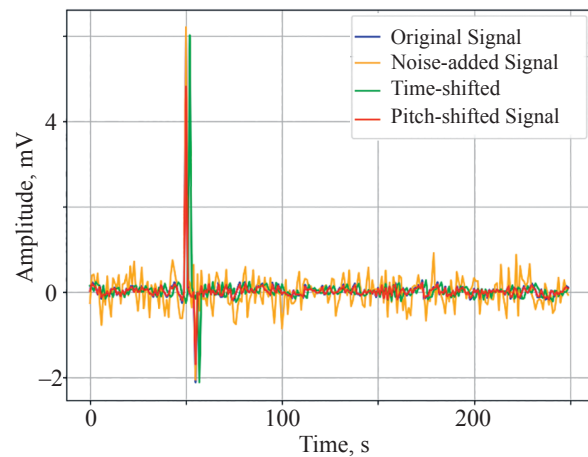


Fig. 5. Data augmentation techniques

The dataset includes signals from 90 individuals, resulting in a total of 63,000 segments (90 individuals \times 700 segments), ensuring a balanced distribution across all individuals. This balance is particularly important for the SNN which compares pairs of ECG signals to verify if they belong to the same individual for human authentication, as it allows the model to learn a robust similarity metric by training on an equal number of positive pairs (same individual) and negative pairs (different individuals) for each class.

Subsequently, the normalized data is shuffled and organized into pairs. Each pair is categorized as either similar (Signal 1) if the segments belong to the same class, or dissimilar (Signal 0) if they pertain to different classes. This pairing strategy, illustrated in Fig. 6, sets the stage for training a classifier to distinguish between similar and dissimilar ECG segments.

Specifically, similar pairs (positive pairs) help the model learn the characteristics of matching signals within the same patient class, while dissimilar pairs (negative pairs) enable the model to identify clear differences between unrelated ECG segments. Fig. 6, *a* shows an example of a positive pair where both signals closely match, and Fig. 6, *b* displays a negative pair where the signals differ significantly. This approach improves the model ability to generalize and enhances classification performance.

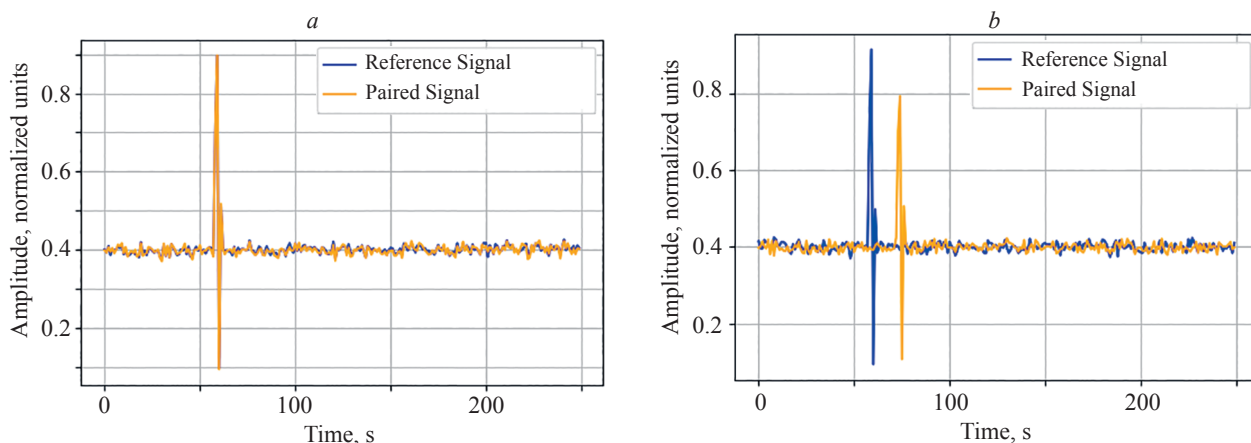


Fig. 6. Examples of pairs: positive (*a*) and negative (*b*)

To calculate the classifier's performance, the dataset is divided into training, validation, and test sets. The Sklearn selection model is utilized for this task, randomly assigning 60 % of the data to training, 20 % to validation, and 20 % to testing. The training set is used to design and fine-tune the classifier, while the test set is held out to assess the classifier's performance and accuracy.

This strategic division of the dataset into training, validation, and test sets not only facilitates robust training but also ensures an unbiased assessment of the classifier's generalization to new, unseen data. The allocation percentages are chosen to strike a balance between providing sufficient data for model training and maintaining an adequate portion for independent evaluation. This meticulous approach enhances the reliability and applicability of the classifier in real-world scenarios.

During the training phase of CNN model, the ECG dataset underwent a segmentation process resulting in 10,000 segments per individual, ensuring a balanced class distribution. This segmentation strategy aims to capture a comprehensive representation of each class, laying the groundwork for a more robust analysis.

Following segmentation, the pre-processed data underwent normalization, scaling the values to a standardized range between 0 and 1. This normalization step is pivotal in mitigating anomalies that might complicate signal analysis. By standardizing the data, variations in magnitude are minimized providing a more consistent foundation for subsequent analysis of ECG signals.

Subsequently, the normalized data is shuffled and partitioned into training and validation sets. This shuffling process is crucial for preventing the model from learning sequence patterns inherent in the data, ensuring a more unbiased evaluation. The division into training and validation sets enables a thorough assessment of the classifier's performance during development.

The dataset includes signals from 90 individuals, with each individual's 700 original segments augmented to

create 10,000 segments, resulting in a total of 900,000 segments (90 individuals \times 10,000 segments), ensuring a balanced distribution across all individuals. This balance is particularly important for the CNN which classifies ECG signals to identify individuals for human authentication, as it allows the model to learn robust features from an equal number of segments for each class, preventing bias toward overrepresented individuals and improving classification accuracy.

To achieve this partitioning, the Sklearn selection model was utilized, randomly allocating 60 % of the data to the training set, 20 % to the validation set, and 20 % for testing. The training set serves as the basis for designing and fine-tuning the classifier, while the validation set aids in monitoring the model performance and making adjustments to enhance generalization.

The remaining 20 % of the data is reserved for testing, serving as an independent set for the final evaluation of the classifier's performance and accuracy. This separation into distinct training, validation, and test sets ensures a comprehensive assessment of the classifier's ability to generalize to new, unseen data.

Testing Phase

For the verification task, the SNN was trained for 100 epochs, incorporating Model Checkpoint and early stopping mechanisms to monitor validation loss. Using Euclidean Distance, the SNN employed contrastive loss during training. This loss function evaluates the distance between outputs for positive and negative samples, ensuring the network effectively discriminates between these instances for accurate authentication. The training process typically required 20 to 25 minutes, with validation loss stabilizing at 0.3 %. Through experimentation, a decision threshold margin of 0.0009999 was established. If the similarity score equals or exceeds this threshold, the ECG segment is classified as a match with an enrolled template, resulting in successful verification. Some test samples from our experiment are illustrated in Fig. 7.

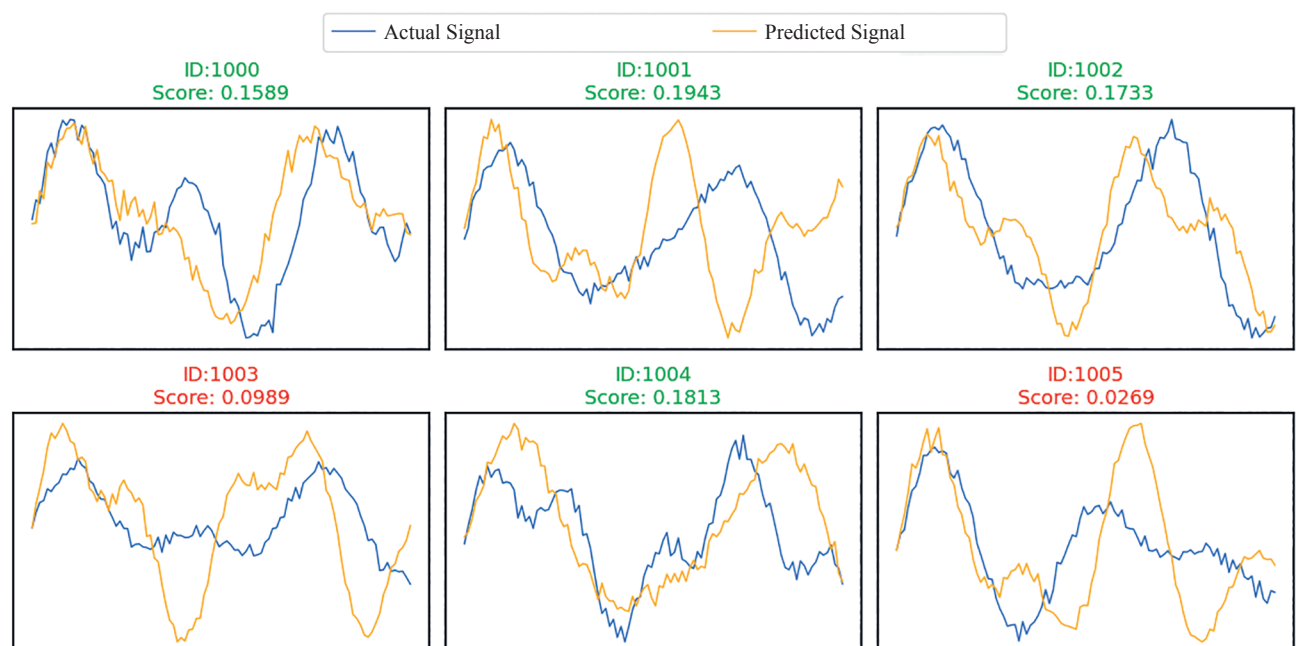


Fig. 7. Predicted samples from verification (SNN) model

For the authentication task, the CNN model underwent training for an average duration of 4 hours across 100 epochs. Early stopping and Model Checkpoint prevented the possibility of finishing after around 20–25 epochs on average. With an accuracy of 98.8 % and a validation loss of 0.32 %, as shown in Fig. 8, *a*, the rapid convergence of training and validation accuracy within the first 20 epochs are demonstrated, while Fig. 8, *b* shows a steady decrease in both training and validation loss, confirming effective learning with minimal overfitting.

This study consolidates the performance of the proposed SNN-CNN framework for ECG-based authentication in IoT telehealth, highlighting its resource efficiency, accuracy, reliability, and adaptability, with results demonstrated through comprehensive figures.

The confusion matrix in Fig. 9 grounds the precision, recall, and F1-score to nearly 100 %. The confusion matrix shows patient information in a diagonal matrix form and denotes that the total person strength has been estimated correctly. It is used to visually represent the performance

of the classification model by showing how well predicted labels match the actual person classes. A perfect or near-perfect diagonal pattern indicates the model accuracy in correctly identifying each class without confusion. Several samples predicted by the model are given in Fig. 10, which shows the applicability of the model for distinguishing between instances.

The Receiver Operating Characteristic (ROC) curve can then be used to evaluate the model effectiveness in user authentication with ECG data in terms of decision thresholds. The AUC is an estimate of the overall model performance. This indirectly provides a possibility of selecting an optimal decision point depending on the needs of a particular application having both sensitivity and specificity in trade off. If your model ROC curve is located near the top left corner, you have a strong model that is great at both high sensitivity and specificity. Fig. 11 shows performance of the proposed model at threshold value.

The performance of the proposed model is 98.8 % based on main parameters and significantly higher than

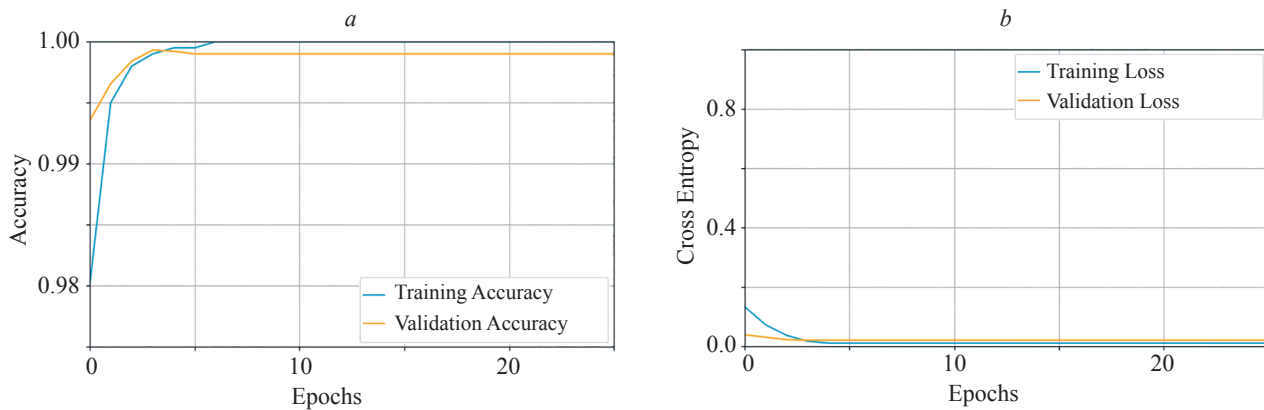


Fig. 8. Accuracy (*a*) and loss (*b*) scores for training and validation

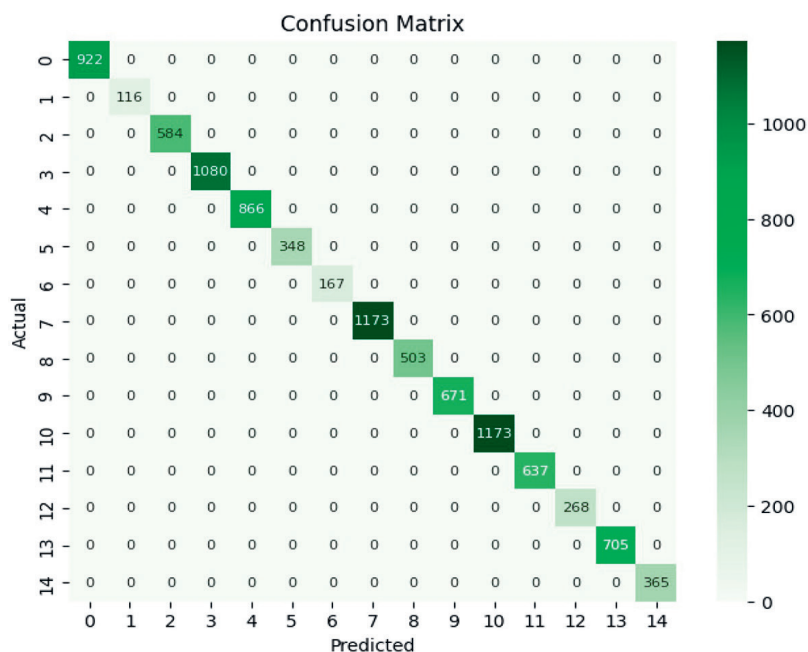


Fig. 9. Confusion Matrix of the proposed authentication model

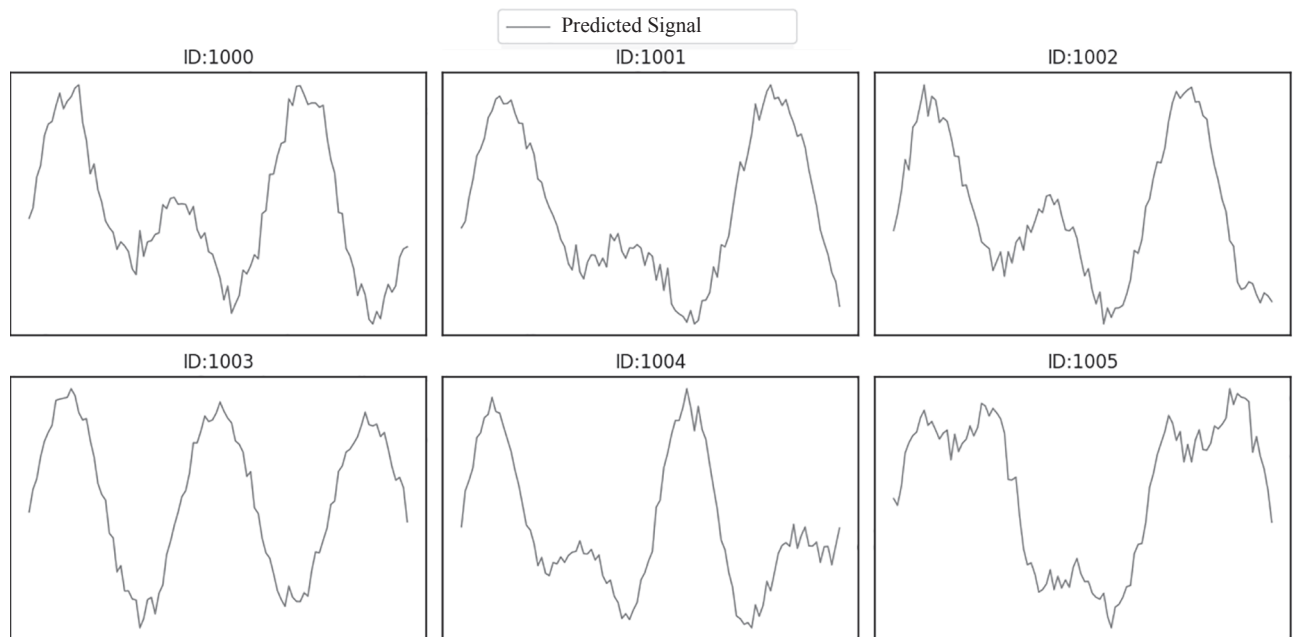


Fig. 10. Predicted samples from authentication (CNN) model

Table. Performance comparison of the proposed model and other models

Publication	Database	Methods	Accuracy, %
[38]	PhysioNet	Random Forest classifier	92.0
[39]	ECG-ID	Decision Curve Analysis	94.0
[40]	ptbdb, mitdb, nsrdb	Matching process	97.6
[20]	MIT-BIH	Feed-Forward Neural Network	95.0
[30]	ECG-ID	Euclidean detector	94.3
[41]	ECG-ID	CNN	96.6
[42]	Low-cost sensors biometrics	One-class classifier density estimation	98.0
[43]	PTB	CNN-LSTM	98.0
Proposed Solution	ECG-ID	SNN-CNN	98.8

the performance of the previous models. It also shows high specificity of 0.99 %, sensitivity of 0.99 %, the AUC of 0.99 % and F1-score of 0.99 %. From the results, we see

that it is efficient for ECG based authentication and could be better than other advanced algorithms. The reliability and accuracy of the model may radically change the existing ECG-based authentication systems, primarily in security fields. Table above indicates a comparison of the proposed model with other architectures.

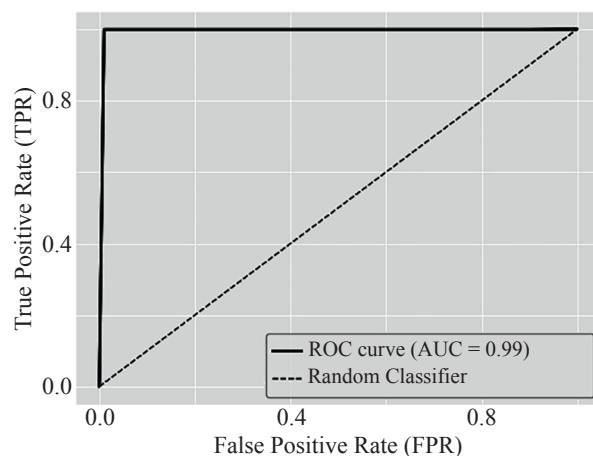


Fig. 11. Predicted samples from authentication (CNN) model

Conclusion

The proposed methodology introduces a significant advancement in Electrocardiogram (ECG)-based biometric systems for real-time authentication in IoT telehealth applications. By integrating a Siamese Neural Network for verification and a Convolutional Neural Networks for authentication, this hybrid approach enhances system reliability, security, and adaptability. Unique biometric signatures, resistant to replication or theft, offer superior security compared to traditional methods.

Separating authentication and verification models provides flexibility, scalability, and optimized performance across diverse scenarios. Lightweight models suit resource-

constrained settings, while advanced models cater to high-security environments. This modularity minimizes errors, supports targeted optimization, and facilitates continuous improvement.

Future research could explore handling noisy ECG signals, integrating multi-modal biometrics, implementing wearable sensors, and applying the model to secure governmental or military domains. Another key direction involves leveraging the two separate models for solving

Open-Set Recognition challenges, ensuring accurate identification of known users while effectively rejecting unfamiliar subjects. Testing on a large ECG dataset demonstrated high accuracy, leveraging preprocessing techniques like filtering and normalization. Advanced architectures and expanded applications to other physiological signals, such as photoplethysmography and electroencephalography, could further enhance the model versatility.

References

1. Suran Melissa. Increased use of medicare telehealth during the pandemic. *JAMA*, 2022, vol. 327, no. 4, pp. 313. <https://doi.org/10.1001/jama.2021.23332>
2. Marquez G., Astudillo H., Taramasco C. Security in telehealth systems from a software engineering viewpoint: a systematic mapping study. *IEEE Access*, 2020, vol. 8, pp. 10933–10950. <https://doi.org/10.1109/access.2020.2964988>
3. Watzlaf V.J.M., Zhou L., DeAlmeida D., Hartman L.M. A systematic review of research studies examining telehealth privacy and security practices used by healthcare providers. *International Journal of Telerehabilitation*, 2017, vol. 9, no. 2, pp. 39–59. <https://doi.org/10.5195/IJT.2017.6231>
4. Yousuf T., Mahmoud R., Aloul F., Zualkernan I. Internet of Things (IoT) Security: current status, challenges and countermeasures. *International Journal for Information Security Research*, 2015, vol. 5, no. 4, pp. 608–616. <https://doi.org/10.20533/ijisr.2042.4639.2015.0070>
5. Zhou L., Thieret R., Watzlaf V., Fahima, Dealmeida D., Parmanto B. A telehealth privacy and security self-assessment questionnaire for telehealth providers: development and validation. *International Journal of Telerehabilitation*, 2019, vol. 11, no. 1, pp. 3–14. <https://doi.org/10.5195/ijt.2019.6276>
6. Sodhro A.H., Sennersten C., Ahmad A. Towards cognitive authentication for smart healthcare applications. *Sensors*, 2022, vol. 22, no. 6, pp. 2101. <https://doi.org/10.3390/s22062101>
7. Khan H., Jan Z.H., Ullah I., Alwabli A., Alharbi F., Habib S., Islam M., Shin B.J., Lee M.Y., Koo J. A deep dive into AI integration and advanced nanobiosensor technologies for enhanced bacterial infection monitoring. *Nanotechnology Reviews*, 2024, vol. 13, no. 1, pp. 20240056. <https://doi.org/10.1515/ntrev-2024-0056>
8. Barros A., Rosário D., Resque P., Cerqueira E. Heart of IoT: ECG as biometric sign for authentication and identification. *Proc. of the 15th International Wireless Communications & Mobile Computing Conference (IWCMC)*, 2019, pp. 307–312. <https://doi.org/10.1109/iwcmc.2019.8766495>
9. Kim S.-K., Yeun C.Y., Damiani E., Lo N.W. A machine learning framework for biometric authentication using electrocardiogram. *IEEE Access*, 2019, vol. 7, pp. 94858–94868. <https://doi.org/10.1109/access.2019.2927079>
10. Rehman I.U., Ullah I., Khan H., Guellil M.S., Koo J., Min J., Habib S., Islam M., Lee M.Y. A comprehensive systematic literature review of ML in nanotechnology for sustainable development. *Nanotechnology Reviews*, 2024, vol. 13, no. 1, pp. 20240069. <https://doi.org/10.1515/ntrev-2024-0069>
11. Ibtihaz N., Chowdhury M.E.H., Khandakar A., Kiranyaz S., Rahman M.S., Tahir A., Qiblawey Y., Rahman T. EDITH: ECG biometrics aided by deep learning for reliable individual authentication. *IEEE Transactions on Emerging Topics in Computational Intelligence*, 2022, vol. 6, no. 4, pp. 928–940. <https://doi.org/10.1109/tetci.2021.3131374>
12. Asadifam S., Talebi M.J., Nikougoftar E. ECG-based authentication systems: a comprehensive and systematic review. *Multimedia Tools and Applications*, 2024, vol. 82, no. 9, pp. 27647–27701. <https://doi.org/10.1007/s11042-023-16506-3>
13. Shdefat, A.Y., Mostafa, N., Saker, I., Topcu, A. A survey study of the current challenges and opportunities of deploying the ECG biometric authentication method in IoT and 5G environments. *Indonesian Journal of Electrical Engineering and Informatics*, 2021, vol. 9, no. 2, pp. 394–416. <https://doi.org/10.52549/ijeei.v9i2.2890>
14. Li L., Chen C., Pan L., Zhang L.Y., Wang Z.F., Zhang J., Xiang Y. A survey of PPG's application in authentication. *Computers & Security*,

Литература

1. Suran Melissa. Increased use of medicare telehealth during the pandemic // *JAMA*. 2022. V. 327. N 4. P. 313. <https://doi.org/10.1001/jama.2021.23332>
2. Marquez G., Astudillo H., Taramasco C. Security in telehealth systems from a software engineering viewpoint: a systematic mapping study // *IEEE Access*. 2020. V. 8. P. 10933–10950. <https://doi.org/10.1109/access.2020.2964988>
3. Watzlaf V.J.M., Zhou L., DeAlmeida D., Hartman L.M. A systematic review of research studies examining telehealth privacy and security practices used by healthcare providers // *International Journal of Telerehabilitation*. 2017. V. 9. N 2. P. 39–59. <https://doi.org/10.5195/IJT.2017.6231>
4. Yousuf T., Mahmoud R., Aloul F., Zualkernan I. Internet of Things (IoT) Security: current status, challenges and countermeasures // *International Journal for Information Security Research*. 2015. V. 5. N 4. P. 608–616. <https://doi.org/10.20533/ijisr.2042.4639.2015.0070>
5. Zhou L., Thieret R., Watzlaf V., Fahima, Dealmeida D., Parmanto B. A telehealth privacy and security self-assessment questionnaire for telehealth providers: development and validation // *International Journal of Telerehabilitation*. 2019. V. 11. N 1. P. 3–14. <https://doi.org/10.5195/ijt.2019.6276>
6. Sodhro A.H., Sennersten C., Ahmad A. Towards cognitive authentication for smart healthcare applications // *Sensors*. 2022. V. 22. N 6. P. 2101. <https://doi.org/10.3390/s22062101>
7. Khan H., Jan Z.H., Ullah I., Alwabli A., Alharbi F., Habib S., Islam M., Shin B.J., Lee M.Y., Koo J. A deep dive into AI integration and advanced nanobiosensor technologies for enhanced bacterial infection monitoring // *Nanotechnology Reviews*. 2024. V. 13. N 1. P. 20240056. <https://doi.org/10.1515/ntrev-2024-0056>
8. Barros A., Rosário D., Resque P., Cerqueira E. Heart of IoT: ECG as biometric sign for authentication and identification // *Proc. of the 15th International Wireless Communications & Mobile Computing Conference (IWCMC)*. 2019. P. 307–312. <https://doi.org/10.1109/iwcmc.2019.8766495>
9. Kim S.-K., Yeun C.Y., Damiani E., Lo N.W. A machine learning framework for biometric authentication using electrocardiogram // *IEEE Access*. 2019. V. 7. P. 94858–94868. <https://doi.org/10.1109/access.2019.2927079>
10. Rehman I.U., Ullah I., Khan H., Guellil M.S., Koo J., Min J., Habib S., Islam M., Lee M.Y. A comprehensive systematic literature review of ML in nanotechnology for sustainable development // *Nanotechnology Reviews*. 2024. V. 13. N 1. P. 20240069. <https://doi.org/10.1515/ntrev-2024-0069>
11. Ibtihaz N., Chowdhury M.E.H., Khandakar A., Kiranyaz S., Rahman M.S., Tahir A., Qiblawey Y., Rahman T. EDITH: ECG biometrics aided by deep learning for reliable individual authentication // *IEEE Transactions on Emerging Topics in Computational Intelligence*. 2022. V. 6. N 4. P. 928–940. <https://doi.org/10.1109/tetci.2021.3131374>
12. Asadifam S., Talebi M.J., Nikougoftar E. ECG-based authentication systems: a comprehensive and systematic review // *Multimedia Tools and Applications*. 2024. V. 82. N 9. P. 27647–27701. <https://doi.org/10.1007/s11042-023-16506-3>
13. Shdefat, A.Y., Mostafa, N., Saker, I., Topcu, A. A survey study of the current challenges and opportunities of deploying the ECG biometric authentication method in IoT and 5G environments // *Indonesian Journal of Electrical Engineering and Informatics*. 2021. V. 9. N 2. P. 394–416. <https://doi.org/10.52549/ijeei.v9i2.2890>
14. Li L., Chen C., Pan L., Zhang L.Y., Wang Z.F., Zhang J., Xiang Y. A survey of PPG's application in authentication // *Computers &*

- 2023, vol. 135, pp. 103488. <https://doi.org/10.1016/j.cose.2023.103488>
15. Pereira T.M.C., Conceição R.C., Sencadas V., Sebastião R. Biometric recognition: a systematic review on electrocardiogram data acquisition methods. *Sensors*, 2023, vol. 23, no. 3, pp. 1507. <https://doi.org/10.3390/s23031507>
 16. Hammad M., Plawiak P., Wang K.Q., Acharya U.R. ResNet-Attention model for human authentication using ECG signals. *Expert Systems*, 2021, vol. 38, no. 6, pp. e12547. <https://doi.org/10.1111/exsy.12547>
 17. Labati R.D., Muñoz E., Piuri V., Sassi R., Scotti F. Deep-ECG: convolutional neural networks for ECG biometric recognition. *Pattern Recognition Letters*, 2019, vol. 126, pp. 78–85. <https://doi.org/10.1016/j.patrec.2018.03.028>
 18. Tirado-Martín P., Sanchez-Rello R. BioEcg: Improving ECG biometrics with deep learning and enhanced datasets. *Applied Sciences*, 2021, vol. 11, no. 13, pp. 5880. <https://doi.org/10.3390/app11135880>
 19. D'angelis O., Bacco L., Vollero L., Merone M. Advancing ECG biometrics through vision transformers: a confidence-driven approach. *IEEE Access*, 2023, vol. 11, pp. 140710–140721. <https://doi.org/10.1109/ACCESS.2023.3338191>
 20. Alduwaile D., Islam M.S. Single heartbeat ECG biometric recognition using convolutional neural network. *Proc. of the International Conference on Advanced Science and Engineering (ICOASE)*, 2020, pp. 145–150. <https://doi.org/10.1109/ICOASE51841.2020.9436592>
 21. Ivanciu L., Ivanciu I.A., Farago P., Roman M., Hintea S. An ECG-based authentication system using siamese neural networks. *Journal of Medical and Biological Engineering*, 2021, vol. 41, no. 4, pp. 558–570. <https://doi.org/10.1007/s40846-021-00637-9>
 22. Albuquerque S.L., Misoso C.J., da Rocha A.F., Gondim P.R.L. Authentication based on electrocardiography signals and machine learning. *Engineering Research Express*, 2021, vol. 3, no. 2, pp. 023504. <https://doi.org/10.1088/2631-8695/abffa6>
 23. Al Alkeem, E., Yeun C.Y., Yun J., Yoo P.D., Chae M., Rahman A., Asyhari A.T. Robust deep identification using ECG and multimodal biometrics for industrial internet of things. *Ad Hoc Networks*, 2021, vol. 121, pp. 102581. <https://doi.org/10.1016/j.adhoc.2021.102581>
 24. Ahmad I., Yao C., Li L., Chen Y., Liu Z., Ullah I., Shabaz M., Wang X., Huang K., Li G., Zhao G., Samuel O.W., Chen S. An efficient feature selection and explainable classification method for EEG-based epileptic seizure detection. *Journal of Information Security and Applications*, 2024, vol. 80, pp. 103654. <https://doi.org/10.1016/j.jisa.2023.103654>
 25. Jamin A., Humeau-Heurtier A. (Multiscale) cross-entropy methods: a review. *Entropy*, 2019, vol. 22, no. 1, pp. 45. <https://doi.org/10.3390/e22010045>
 26. Domínguez-Bolaño T., Campos O., Barral V., Escudero C.J., García-Naya J.A. An overview of IoT architectures, technologies, and existing open-source projects. *Internet of Things*, 2022, vol. 20, pp. 100626. <https://doi.org/10.1016/j.iot.2022.100626>
 27. Pourghebleh B., Wakil K., Navimipour N.J. A comprehensive study on the trust management techniques in the Internet of Things. *IEEE Internet of Things Journal*, 2019, vol. 6, no. 6, pp. 9626–9337. <https://doi.org/10.1109/jiot.2019.2933518>
 28. Sharma P., Jain S., Gupta S., Chamola V. Role of machine learning and deep learning in securing 5G-driven industrial IoT applications. *Ad Hoc Networks*, 2021, vol. 123, pp. 102685. <https://doi.org/10.1016/j.adhoc.2021.102685>
 29. Dogo E.M., Afolabi O.J., Nwulu N.I., Twala B., Aigbavboa C.O. A comparative analysis of gradient descent-based optimization algorithms on convolutional neural networks. *Proc. of the International Conference on Computational Techniques, Electronics and Mechanical Systems (CTEMS)*, 2018, pp. 92–99. <https://doi.org/10.1109/ctems.2018.8769211>
 30. Tyagi P.K., Agrawal D. Automatic detection of sleep apnea from single-lead ECG signal using enhanced-deep belief network model. *Biomedical Signal Processing and Control*, 2023, vol. 80, part 2, pp. 104401. <https://doi.org/10.1016/j.bspc.2022.104401>
 31. Bento N., Belo D., Gamboa H. ECG biometrics using spectrograms and deep neural networks. *International Journal of Machine Learning and Computing*, 2020, vol. 10, no. 2, pp. 259–264. <https://doi.org/10.18178/ijmlc.2020.10.2.929>
 32. Zhou D.-X. Deep distributed convolutional neural networks: Universality. *Analysis and Applications*, 2018, vol. 16, no. 6, pp. 895–919. <https://doi.org/10.1142/s0219530518500124>
 - Security. 2023. V. 135. P. 103488. <https://doi.org/10.1016/j.cose.2023.103488>
 15. Pereira T.M.C., Conceição R.C., Sencadas V., Sebastião R. Biometric recognition: a systematic review on electrocardiogram data acquisition methods // *Sensors*. 2023. V. 23. N 3. P. 1507. <https://doi.org/10.3390/s23031507>
 16. Hammad M., Plawiak P., Wang K.Q., Acharya U.R. ResNet-Attention model for human authentication using ECG signals // *Expert Systems*. 2021. V. 38. N 6. P. e12547. <https://doi.org/10.1111/exsy.12547>
 17. Labati R.D., Muñoz E., Piuri V., Sassi R., Scotti F. Deep-ECG: convolutional neural networks for ECG biometric recognition // *Pattern Recognition Letters*. 2019. V. 126. P. 78–85. <https://doi.org/10.1016/j.patrec.2018.03.028>
 18. Tirado-Martín P., Sanchez-Rello R. BioEcg: Improving ECG biometrics with deep learning and enhanced datasets // *Applied Sciences*. 2021. V. 11. N 13. P. 5880. <https://doi.org/10.3390/app11135880>
 19. D'angelis O., Bacco L., Vollero L., Merone M. Advancing ECG biometrics through vision transformers: a confidence-driven approach // *IEEE Access*. 2023. V. 11. P. 140710–140721. <https://doi.org/10.1109/ACCESS.2023.3338191>
 20. Alduwaile D., Islam M.S. Single heartbeat ECG biometric recognition using convolutional neural network // *Proc. of the International Conference on Advanced Science and Engineering (ICOASE)*. 2020. P. 145–150. <https://doi.org/10.1109/ICOASE51841.2020.9436592>
 21. Ivanciu L., Ivanciu I.A., Farago P., Roman M., Hintea S. An ECG-based authentication system using siamese neural networks // *Journal of Medical and Biological Engineering*. 2021. V. 41. N 4. P. 558–570. <https://doi.org/10.1007/s40846-021-00637-9>
 22. Albuquerque S.L., Misoso C.J., da Rocha A.F., Gondim P.R.L. Authentication based on electrocardiography signals and machine learning // *Engineering Research Express*. 2021. V. 3. N 2. P. 023504. <https://doi.org/10.1088/2631-8695/abffa6>
 23. Al Alkeem, E., Yeun C.Y., Yun J., Yoo P.D., Chae M., Rahman A., Asyhari A.T. Robust deep identification using ECG and multimodal biometrics for industrial internet of things // *Ad Hoc Networks*, 2021, V. 121. P. 102581. <https://doi.org/10.1016/j.adhoc.2021.102581>
 24. Ahmad I., Yao C., Li L., Chen Y., Liu Z., Ullah I., Shabaz M., Wang X., Huang K., Li G., Zhao G., Samuel O.W., Chen S. An efficient feature selection and explainable classification method for EEG-based epileptic seizure detection // *Journal of Information Security and Applications*. 2024. V. 80. P. 103654. <https://doi.org/10.1016/j.jisa.2023.103654>
 25. Jamin A., Humeau-Heurtier A. (Multiscale) cross-entropy methods: a review // *Entropy*. 2019. V. 22. N 1. P. 45. <https://doi.org/10.3390/e22010045>
 26. Domínguez-Bolaño T., Campos O., Barral V., Escudero C.J., García-Naya J.A. An overview of IoT architectures, technologies, and existing open-source projects // *Internet of Things*. 2022. V. 20. P. 100626. <https://doi.org/10.1016/j.iot.2022.100626>
 27. Pourghebleh B., Wakil K., Navimipour N.J. A comprehensive study on the trust management techniques in the Internet of Things // *IEEE Internet of Things Journal*. 2019. V. 6. N 6. P. 9626–9337. <https://doi.org/10.1109/jiot.2019.2933518>
 28. Sharma P., Jain S., Gupta S., Chamola V. Role of machine learning and deep learning in securing 5G-driven industrial IoT applications // *Ad Hoc Networks*. 2021. V. 123. P. 102685. <https://doi.org/10.1016/j.adhoc.2021.102685>
 29. Dogo E.M., Afolabi O.J., Nwulu N.I., Twala B., Aigbavboa C.O. A comparative analysis of gradient descent-based optimization algorithms on convolutional neural networks // *Proc. of the International Conference on Computational Techniques, Electronics and Mechanical Systems (CTEMS)*. 2018. P. 92–99. <https://doi.org/10.1109/ctems.2018.8769211>
 30. Tyagi P.K., Agrawal D. Automatic detection of sleep apnea from single-lead ECG signal using enhanced-deep belief network model // *Biomedical Signal Processing and Control*. 2023. V. 80. Part 2. P. 104401. <https://doi.org/10.1016/j.bspc.2022.104401>
 31. Bento N., Belo D., Gamboa H. ECG biometrics using spectrograms and deep neural networks // *International Journal of Machine Learning and Computing*. 2020. V. 10. N 2. P. 259–264. <https://doi.org/10.18178/ijmlc.2020.10.2.929>
 32. Zhou D.-X. Deep distributed convolutional neural networks: Universality // *Analysis and Applications*. 2018. V. 16. N 6. P. 895–919. <https://doi.org/10.1142/s0219530518500124>

33. Sainath T.N., Kingsbury B., Saon G., Soltau H., Mohamed A.R., Dahl G., Ramabhadran B. Deep convolutional neural networks for Large-scale speech tasks. *Neural Networks*, 2015, vol. 64, pp. 39–48. <https://doi.org/10.1016/j.neunet.2014.08.005>
34. Khan A., Sohail A., Zahoor U., Qureshi A.S. A survey of the recent architectures of deep convolutional neural networks. *Artificial Intelligence Review*, 2020, vol. 53, no. 8, pp. 5455–5516. <https://doi.org/10.1007/s10462-020-09825-6>
35. Lodhi B., Kang J. Multipath-DenseNet: A Supervised ensemble architecture of densely connected convolutional networks. *Information Sciences*, 2019, vol. 482, pp. 63–72. <https://doi.org/10.1016/j.ins.2019.01.012>
36. Kanatov M., Atymtayeva L., Mendes M. Improved Facial Expression Recognition with xception deep net and preprocessed images. *Applied Mathematics & Information Sciences*, 2019, vol. 13, no. 5, pp. 859–865. <https://doi.org/10.18576/amis/130520>
37. Zhang Y., Wu J. Practical human authentication method based on piecewise corrected Electrocardiogram. *Proc. of the 7th IEEE International Conference on Software Engineering and Service Sciences (ICSESS)*, 2016, pp. 300–303. <https://doi.org/10.1109/icseess.2016.7883071>
38. Nwankpa C., Ijomah W., Gachagan A., Marshall S. Activation Functions: Comparison of Trends in Practice and Research for Deep Learning. *arXiv*, 2018, arxiv.org/abs/1811.03378v1. <https://doi.org/10.48550/arXiv.1811.03378>
39. Barros A., Resque P., Almeida J., Mota R., Oliveira H., Rosário D., Cerqueira E. Data improvement model based on ECG biometric for user authentication and identification. *Sensors*, 2020, vol. 20, no. 10, pp. 2920. <https://doi.org/10.3390/s20102920>
40. Su K., Yang G., Wu B., Yang L., Li D., Su P., Yin Y. Human identification using finger vein and ECG signals. *Neurocomputing*, 2019, vol. 332, pp. 111–118. <https://doi.org/10.1016/j.neucom.2018.12.015>
41. Zhao Z., Zhang Y., Deng Y., Zhang X. ECG authentication system design incorporating a convolutional neural network and generalized S-Transformation. *Computers in Biology and Medicine*, 2018, vol. 102, pp. 168–179. <https://doi.org/10.1016/j.combiomed.2018.09.027>
42. Blasco J., Peris-Lopez P. On the feasibility of low-cost wearable Sensors for multi-modal biometric verification. *Sensors*, 2018, vol. 18, no. 9, pp. 2782. <https://doi.org/10.3390/s18092782>
43. Agrawal V., Hazratifard M., Elmiligi H., Gebali F. Electrocardiogram (ECG)-based user authentication using deep learning algorithms. *Diagnostics*, 2023, vol. 13, no. 3, pp. 439. <https://doi.org/10.3390/diagnostics13030439>
33. Sainath T.N., Kingsbury B., Saon G., Soltau H., Mohamed A.R., Dahl G., Ramabhadran B. Deep convolutional neural networks for Large-scale speech tasks // *Neural Networks*. 2015. V. 64. P. 39–48. <https://doi.org/10.1016/j.neunet.2014.08.005>
34. Khan A., Sohail A., Zahoor U., Qureshi A.S. A survey of the recent architectures of deep convolutional neural networks // *Artificial Intelligence Review*. 2020. V. 53. N 8. P. 5455–5516 <https://doi.org/10.1007/s10462-020-09825-6>
35. Lodhi B., Kang J. Multipath-DenseNet: A Supervised ensemble architecture of densely connected convolutional networks // *Information Sciences*. 2019. V. 482. P. 63–72. <https://doi.org/10.1016/j.ins.2019.01.012>
36. Kanatov M., Atymtayeva L., Mendes M. Improved Facial Expression Recognition with xception deep net and preprocessed images // *Applied Mathematics & Information Sciences*. 2019. V. 13. N 5. P. 859–865. <https://doi.org/10.18576/amis/130520>
37. Zhang Y., Wu J. Practical human authentication method based on piecewise corrected Electrocardiogram // *Proc. of the 7th IEEE International Conference on Software Engineering and Service Sciences (ICSESS)*. 2016. P. 300–303. <https://doi.org/10.1109/icseess.2016.7883071>
38. Nwankpa C., Ijomah W., Gachagan A., Marshall S. Activation Functions: Comparison of Trends in Practice and Research for Deep Learning // *arXiv*. 2018. arxiv.org/abs/1811.03378v1. <https://doi.org/10.48550/arXiv.1811.03378>
39. Barros A., Resque P., Almeida J., Mota R., Oliveira H., Rosário D., Cerqueira E. Data improvement model based on ECG biometric for user authentication and identification // *Sensors*. 2020. V. 20. N 10. P. 2920. <https://doi.org/10.3390/s20102920>
40. Su K., Yang G., Wu B., Yang L., Li D., Su P., Yin Y. Human identification using finger vein and ECG signals // *Neurocomputing*. 2019. V. 332. P. 111–118. <https://doi.org/10.1016/j.neucom.2018.12.015>
41. Zhao Z., Zhang Y., Deng Y., Zhang X. ECG authentication system design incorporating a convolutional neural network and generalized S-Transformation // *Computers in Biology and Medicine*. 2018. V. 102. P. 168–179. <https://doi.org/10.1016/j.combiomed.2018.09.027>
42. Blasco J., Peris-Lopez P. On the feasibility of low-cost wearable Sensors for multi-modal biometric verification // *Sensors*. 2018. V. 18. N 9. P. 2782. <https://doi.org/10.3390/s18092782>
43. Agrawal V., Hazratifard M., Elmiligi H., Gebali F. Electrocardiogram (ECG)-based user authentication using deep learning algorithms // *Diagnostics*. 2023. V. 13. N 3. P. 439. <https://doi.org/10.3390/diagnostics13030439>

Author

Mohamed Abdalla Elsayed Azab — PhD Student, ITMO University, Saint Petersburg, 197101, Russian Federation, <https://orcid.org/0009-0000-1748-0029>, mohamed.a.azab@itmo.ru

Автор

Азаб Мохамед Абдалла Эльсейд — аспирант, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, <https://orcid.org/0009-0000-1748-0029>, mohamed.a.azab@itmo.ru

Received 14.12.2024

Approved after reviewing 30.04.2025

Accepted 26.05.2025

Статья поступила в редакцию 14.12.2024

Одобрена после рецензирования 30.04.2025

Принята к печати 26.05.2025



Работа доступна по лицензии
Creative Commons
«Attribution-NonCommercial»

МАТЕМАТИЧЕСКОЕ И КОМПЬЮТЕРНОЕ МОДЕЛИРОВАНИЕ MODELING AND SIMULATION

doi: 10.17586/2226-1494-2025-25-3-487-497

УДК 004.961

Метод определения активного модуля в биологических графах с многокомпонентными весами вершин

Дмитрий Андреевич Усольцев^{1✉}, Иван Игоревич Молотков², Никита Николаевич Артемов³,
Алексей Александрович Сергушичев⁴, Анатолий Абрамович Шальто⁵

^{1,2,3} Институт геномной медицины, Детская больница Нейшенвайд, Колумбус, 43205, США

^{1,5} Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация

^{2,3} Медицинский колледж Университета штата Огайо, Колумбус, 43210, США

⁴ Университет Вашингтона в Сент-Луисе, Сент-Луис, 63110, США

¹ dusoltsev.27@gmail.com✉, <https://orcid.org/0000-0001-8072-310X>

² ivan.molotkov@nationwidechildrens.org, <https://orcid.org/0009-0008-3566-0160>

³ mykyta.artomov@nationwidechildrens.org, <https://orcid.org/0000-0001-5282-8764>

⁴ asergushichev@wustl.edu, <https://orcid.org/0000-0003-1159-7220>

⁵ anatoly.shalyto@gmail.com, <https://orcid.org/0000-0002-2723-2077>

Аннотация

Введение. Активный модуль в биологических графах представляет собой связанный подграф, вершины которого объединены общей биологической функцией. Для определения активного модуля необходимо сначала построить взвешенный биологический граф. Вес каждой вершины вычисляется на основе биологических экспериментов, исследующих искомую биологическую функцию. Однако результаты одного эксперимента могут не полностью описывать искомый активный модуль, а лишь его часть, внося, например, неопределенность в веса вершин. В работе показано, что использование метода Фишера для объединения данных нескольких экспериментов, а затем применение метода Монте-Карло по схеме марковских цепей (МКМЦ) и машинного обучения к результатам метода Фишера, позволяет более эффективно выделять активные модули в биологических графах. **Метод.** В работе используются граф белок-белковых взаимодействий — InWebIM, граф по реконструкции мозга человека из проекта BigBrain и генный граф для вида живых организмов *Caenorhabditis elegans*. Для объединения результатов нескольких экспериментов в одном графе в единый результат применяется метод Фишера. После этого поиск активных модулей выполняется с использованием метода МКМЦ и машинного обучения. Для валидации предлагаемого метода на реальных данных применяются результаты полногеномного ассоциативного исследования по шизофрении и курению, а также матрица экспрессии генов пациентов с кожной меланомой из проекта The Cancer Genome Atlas. **Основные результаты.** Применение метода Фишера позволяет учитывать результаты нескольких биологических экспериментов одновременно. Последующее использование метода МКМЦ и машинного обучения повышает точность определения активных модулей по сравнению с ранжированием вершин графа только на основе метода Фишера. **Обсуждение.** Учет результатов нескольких биологических экспериментов при определении активных модулей играет ключевую роль в повышении точности нахождения вершин активного модуля. Это способствует лучшему пониманию биологических механизмов заболеваний, что может иметь важное значение для разработки новых методов диагностики и терапии.

Ключевые слова

графы, метод Монте-Карло по схеме марковских цепей, метод Фишера, биологические графы, активный модуль

Ссылка для цитирования: Усольцев Д.А., Молотков И.И., Артемов Н.Н., Сергушичев А.А., Шальто А.А. Метод определения активного модуля в биологических графах с многокомпонентными весами вершин // Научно-технический вестник информационных технологий, механики и оптики. 2025. Т. 25, № 3. С. 487–497. doi: 10.17586/2226-1494-2025-25-3-487-497

Method for identifying the active module in biological graphs with multi-component vertex weights

Dmitrii A. Usoltsev¹✉, Ivan I. Molotkov², Mykyta N. Artomov³, Alexey A. Sergushichev⁴, Anatoly A. Shalyto⁵

^{1,2,3} Institute for Genomic Medicine, Nationwide Children's Hospital, Columbus, 43205, USA

^{1,5} ITMO University, Saint Petersburg, 197101, Russian Federation

^{2,3} The Ohio State University College of Medicine, Columbus, 43210, USA

⁴ Washington University School of Medicine in St. Louis, St. Louis, 63110, USA

¹ dusoltsev.27@gmail.com ✉, <https://orcid.org/0000-0001-8072-310X>

² ivan.molotkov@nationwidechildrens.org, <https://orcid.org/0009-0008-3566-0160>

³ mykyta.artomov@nationwidechildrens.org, <https://orcid.org/0000-0001-5282-8764>

⁴ asergushichev@wustl.edu, <https://orcid.org/0000-0003-1159-7220>

⁵ anatoly.shalyto@gmail.com, <https://orcid.org/0000-0002-2723-2077>

Abstract

An active module in biological graphs is a connected subgraph whose vertices share a common biological function. To identify an active module, one must first construct a weighted biological graph. The weight of each vertex is calculated based on biological experiments investigating the target biological function. However, the results of a single experiment may not fully describe the desired active module, covering only part of it and potentially introducing uncertainty into the vertex weights. This work demonstrates that employing Fisher's method to integrate data from multiple experiments followed by applying a Markov chain Monte Carlo (MCMC) and machine learning-based approach to the results of Fisher's method, enables more effective identification of active modules in biological graphs. The study utilizes the InWebIM protein-protein interaction graph, a human brain reconstruction graph from the BigBrain project, and a gene graph for the organism *Caenorhabditis elegans*. To combine the results of several experiments into a single outcome within one graph, Fisher's method is applied. Afterwards, the search for active modules is conducted using an MCMC and machine learning-based method. To validate the proposed method on real data, results from Genome-Wide Association Studies on schizophrenia and smoking are used, along with the gene expression matrix of patients with skin melanoma from the TCGA project. Applying Fisher's method makes it possible to consider the results of multiple biological experiments simultaneously. Subsequent use of the MCMC and machine learning-based method improves the accuracy of identifying active modules compared to ranking graph vertices solely by Fisher's method. Considering the results of multiple biological experiments when determining active modules plays a crucial role in increasing the accuracy of identifying the vertices of the active module. This, in turn, promotes a deeper understanding of the biological mechanisms of diseases, which can be of great significance for the development of new diagnostic and therapeutic methods.

Keywords

graphs, MCMC, Fisher's method, biological graphs, active module

For citation: Usoltsev D.A., Molotkov I.I., Artomov M.N., Sergushichev A.A., Shalyto A.A. Method for identifying the active module in biological graphs with multi-component vertex weights. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2025, vol. 25, no. 3, pp. 487–497 (in Russian). doi: 10.17586/2226-1494-2025-25-3-487-497

Введение

Графы часто используются при анализе сложных биологических взаимодействий, таких как белок-белковые взаимодействия, ко-экспрессия генов или регуляция метаболических путей [1–3]. Внутри таких графов существуют связанные подграфы, вершины которых выполняют общую биологическую функцию. Эти подграфы называют активными модулями [4]. Модули, например, в графе белок-белковых взаимодействий позволяют понять причины возникновения заболеваний и определить возможные терапевтические мишени [5, 6]. Таким образом, определение активных модулей является одной из важных задач при анализе биологических графов.

В работах [7, 8] представлены подходы для определения активных модулей на основе метода Монте-Карло по схеме марковских цепей (МКМЦ) и метода на основе машинного обучения — градиентного бустинга. Однако точность определения активного модуля как с помощью только метода МКМЦ [7], так и с помощью подхода, предложенного в [8], зависит от точности определения весов вершин графа.

Для определения весов в биологическом графе проводятся эксперименты по дифференциальной экспрессии генов или мутационной нагрузке между больными и здоровыми группами людей. Результаты экспериментов анализируются с помощью статистического теста (например, t-тест, линейная регрессия) [9]. В итоге формируется список генов с соответствующими p -значениями, отражающими значимость ассоциации с исследуемым признаком. Распределение p -значений хорошо аппроксимируется бета-равномерным распределением BUM (α, λ) с плотностью, вычисляемой на основе соотношения:

$$f(x) = \lambda + (1 - \lambda)\alpha x^{\alpha-1}, \quad 0 \leq x, \lambda, \alpha \leq 1, \quad (1)$$

где λ — вес равномерной компоненты; α — параметр формы бета-компоненты [10].

В работах [4, 9] предложено, исходя из p -значений, вычислять априорный вес вершины для заданного уровня ложноположительных результатов (False Discovery Rate, FDR) на основе соотношения:

$$w(p) = \left(\frac{p}{p\text{FDR}} \right)^{\alpha-1}, \quad 0 \leq p, \alpha \leq 1, \quad (2)$$

где p — p -значение; $pFDR$ — вероятность того, что p -значение принадлежит равномерному распределению. Эта вероятность вычисляется по формуле [4, 9]:

$$pFDR = \left(\frac{(\lambda + (1 - \lambda)\alpha - FDR \times \lambda) \left(\frac{1}{\alpha - 1} \right)}{FDR(1 - \lambda)} \right), \quad (3)$$

$$0 \leq \lambda, \alpha \leq 1.$$

Традиционно веса вершин рассчитываются на основе результатов одного статистического теста по формулам (2) и (3). Однако в исследованиях часто доступны результаты множества статистических тестов, отражающих разные аспекты данных: дифференциальная экспрессия, эпигенетические модификации, ассоциации с фенотипами. Результаты каждого теста распределены на основе формулы (1). Учет нескольких статистических тестов для определения априорного веса вершин графа может позволить повысить точность анализа. Такой вес является многокомпонентным (рис. 1).

Для совмещения результатов нескольких статистических тестов используем метод Фишера, который включает вычисление статистики хи-квадрат:

$$X^2 = -2 \sum_{i=1}^n \ln(p_i), \quad 0 < p_i \leq 1, \quad (4)$$

где p_i — p -значение i -го эксперимента; n — число экспериментов (статистических тестов) (рис. 1).

Для вычисления вероятности получения статистики X^2 из распределения хи-квадрат (χ^2) со степенью свободы равной удвоенному числу экспериментов [11] используется соотношение:

$$P_{comb} = P(\chi^2 \geq X^2, df = 2 \times n), \quad (5)$$

где P_{comb} — результирующее p -значение; P — вероятность; df — степень свободы.

Активные модули часто достоверно не известны. Кроме того, один эксперимент позволяет определить лишь часть активного модуля. В этих условиях моделирование активных модулей с заданными параметрами в реальных биологических графах является оптимальным вариантом.

В настоящей работе выполнено моделирование активных модулей с заданными параметрами на подграфах с числом вершин $N = 1000$ в белок-белковом графе InWebIM [12]. Предложен метод определения активного модуля для случая, когда известны результаты n экспериментов, каждый из которых покрывает лишь часть активного модуля. Предлагаемый метод основан на последовательном использовании метода Фишера, метода МКМЦ и машинного обучения (используется градиентный бустинг¹). Эффективность предлагаемого метода продемонстрирована сравнением его с простым ранжированием результатов, полученных только методом Фишера для разного числа экспериментов. Показана применимость предлагаемого метода в графе из проекта по реконструкции мозга BigBrain [13, 14] и геном графе (для вида живых организмов *Caenorhabditis elegans* (HS-LC) [13, 15]. Оценено качество метода Фишера и предлагаемого метода на реальных данных полногеномного ассоциативного исследования (Genome-Wide Association Study, GWAS) по шизофрении, курению и экспрессии генов пациентов с кожной меланомой.

¹ Chen T., He T., Benesty M., Khotilovich V., Tang Y., Cho H., Chen K., Mitchell R., Cano I., Zhou T., Li M., Xie J., Lin M., Geng Y., Li Y. xgboost: Extreme Gradient Boosting. R package version 1.5.0.2. [Электронный ресурс]. Режим доступа: <https://CRAN.R-project.org/package=xgboost>, свободный. Яз. англ. (дата обращения: 17.08.2024).

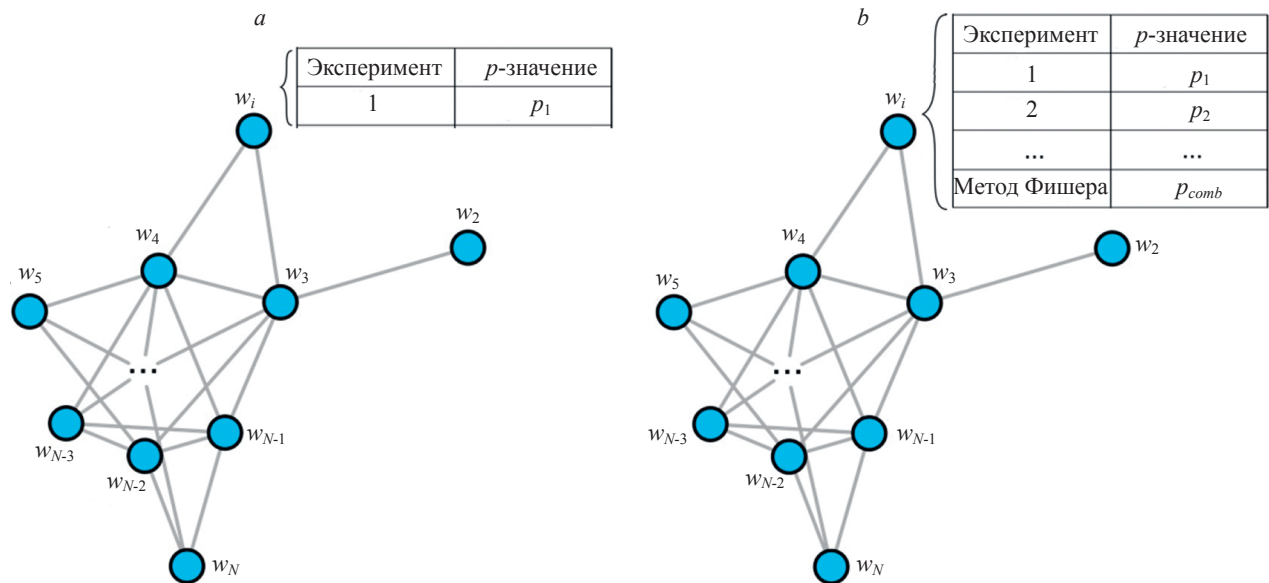


Рис. 1. Сравнение однокомпонентных (a) и многокомпонентных (b) весов вершин.

w_i — вес однокомпонентной i -ой вершины; w_j — вес многокомпонентной j -ой вершины; $i, j \in \{1, N\}$

Fig. 1. Comparison of single-component (a) and multi-component (b) node weights.

w_i — weight of the i -th single-component node; w_j — weight of the j -th multi-component node; $i, j \in \{1, N\}$

Предлагаемый метод

Этапы постановки математической задачи

1. **Определение активного модуля.** В отсутствие достоверно известных активных модулей для большинства биологических состояний проводится симуляция активных модулей с G вершинами в белок-белковом графе InWebIM посредством поиска в ширину из одной случайной вершины. Для повышения скорости вычислений этот граф сэмпляется (разбивается) на связанные подграфы — упрощенные белок-белковые графы. Число вершин в каждом подграфе $N = 1000$.

2. **Определение весов активного модуля.** Для вершин активного модуля генерируются p -значения из бета-распределения, для остальных вершин — из равномерного распределения. Генерация повторяется n раз. Для того чтобы промоделировать направленность биологических экспериментов на исследование только части активного модуля, на каждом этапе генерации p -значений 90 % вершин маскируются случайными значениями из равномерного распределения $X \sim U[0, 1]$. Таким образом, каждой вершине приписывается n независимых p -значений, что соответствует n экспериментам. С помощью метода Фишера для каждой вершины вычисляется одно p -значение. Затем вычисляется априорное значение принадлежности вершины к активному модулю.

Поиск активного модуля и оценка качества поиска

Поиск активного модуля выполняется с помощью метода, предложенного в [8], с учетом 20 % известных белков активного модуля. Качество определения активного модуля оценивается на основе метрик «площадь под кривой ошибок» (Receiver Operating Characteristic Area Under the Curve, ROC AUC; эта величина варьируется в диапазоне от нуля до единицы) и «чувствительность» (Recall@100 — доля правильно определенных белков активного модуля в первых 100 вершинах, ранжированных по уменьшению предсказанной с помощью модели вероятности вхождения в активный модуль; эта величина варьируется в диапазоне от нуля до единицы).

Этапы решения задачи

1. Симуляция 100 различных графов с известными активными модулями

- 1.1. Для получения графов с топологией, встречающейся в реальных данных, выбираются случайные подграфы из графа InWebIM с числом вершин равным 1000.
- 1.2. В каждом подграфе InWebIM из п. 1.1 равномерно среди всех возможных связанных подграфов заданного размера выбирается активный модуль с использованием метода МКМЦ из mcmcRanking (v0.1.0)¹ при условии, что все вершины имеют одинаковый вес. Модули выбираются после 1000 итераций МКМЦ, чтобы гарантировать их

независимость от подграфов, использованных для инициализации МКМЦ.

2. Определение весов активного модуля

- 2.1. Для каждой вершины симулируются p -значения эксперимента, проверяющего принадлежность вершины активному модулю. Для вершин вне активного модуля p -значения выбираются из равномерного распределения, для вершин из активного модуля — из бета-распределения. Для 90 % вершин активного модуля p -значения маскируются случайными значениями из равномерного распределения на интервале $(0, 1]$.
- 2.2. П. 2.1 повторяется n раз, таким образом, чтобы каждой вершине соответствовало n p -значений.
- 2.3. Для каждого p -значения применяется метод Фишера: считается X^2 по формуле (4) и результирующее p -значение по формуле (5).
- 2.4. Полученное p -значение аппроксимацией бета-равномерным распределением преобразуется в априорный вес по формулам (2) и (3).

3. Поиск активного модуля

- 3.1. Вероятность того, что подграф является активным модулем, рассчитывается как произведение вероятностей того, что каждая вершина принадлежит этому модулю. Для получившегося вероятностного пространства на множестве связанных подграфов формируется выборка размера 100 с помощью метода МКМЦ из mcmcRanking (v0.1.0), используя 10 000 симуляций МКМЦ. После этого для каждой вершины определяется эмпирическая вероятность вхождения в активный модуль — доля подграфов из выборки, которые включают в себя эту вершину.
- 3.2. Применяется метод, изложенный в [8]. 20 % вершин активного модуля выбираются с помощью генератора случайных чисел и используются как известные вершины. Вычисляются расстояния от каждой вершины до трех ближайших белков активного модуля. Используя полученные расстояния и вероятность, определенную в п. 3.1, обучается модель градиентного бустинга с параметрами, описанными в [8]. Так как общее число активных модулей равняется 100, то 50 используются для тренировки модели, а оставшиеся 50 — для тестирования. Известные белки в активных модулях исключаются. При тренировке модели белкам, включенным в активный модуль, присваивается единица, а остальным белкам — ноль.

Экспериментальная оценка предлагаемого метода

Симуляции были проведены для $G = 100$ и $n \in \{1, 3, 10, 50\}$. Эффективность метода Фишера и метода, изложенного в [8], оценивалась с помощью метрик ROC AUC и Recall@100. Метрики определялись для каждого подграфа InWebIM, после этого вычислялось среднее значение каждой метрики из выборки в 50 подграфов. При $n = 1$ результаты, получаемые предлагаемым методом и методом, изложенным в [8] совпали, при условии маскировки p -значений для 90 % вершин активного модуля случайными значениями из равномерного распределения на интервале $(0, 1]$.

¹ GitHub — ctlab/mcmcRanking: Tool To Solve The Active Module Problem [Электронный ресурс]. Режим доступа: <https://github.com/ctlab/mcmcRanking>, свободный. Яз. англ. (дата обращения: 15.12.2024).

Таблица. Сравнение предлагаемого метода и метода Фишера в зависимости от параметров α и n Table. Comparison of the proposed method and Fisher's method depending vs. the parameters α and n

Параметры		Предлагаемый метод		Метод Фишера	
α	n	ROC AUC (\pm std)	Recall@100 (\pm std)	ROC AUC (\pm std)	Recall@100 (\pm std)
0,2	1	0,69 \pm 0,11	0,38 \pm 0,19	0,54 \pm 0,03	0,15 \pm 0,03
0,2	3	0,68 \pm 0,09	0,36 \pm 0,17	0,58 \pm 0,03	0,21 \pm 0,04
0,2	10	0,74 \pm 0,07	0,46 \pm 0,13	0,70 \pm 0,03	0,36 \pm 0,04
0,2	50	0,87 \pm 0,05	0,70 \pm 0,11	0,90 \pm 0,02	0,68 \pm 0,06
0,8	1	0,61 \pm 0,12	0,26 \pm 0,19	0,51 \pm 0,04	0,12 \pm 0,03
0,8	3	0,65 \pm 0,11	0,30 \pm 0,21	0,51 \pm 0,04	0,12 \pm 0,05
0,8	10	0,62 \pm 0,10	0,27 \pm 0,16	0,52 \pm 0,03	0,11 \pm 0,03
0,8	50	0,67 \pm 0,14	0,37 \pm 0,24	0,54 \pm 0,03	0,14 \pm 0,04

Примечание. (\pm std) — стандартное отклонение. Жирным шрифтом выделены значения для $n = 1$ и $n = 50$.

В таблице представлены результаты качества определения вершин активного модуля предлагаемым методом и методом Фишера для $\alpha = 0,2$ и $\alpha = 0,8$ и число экспериментов n .

При увеличении n в случае $\alpha = 0,2$, когда бета-равномерное распределение содержит более информативную бета-компоненту, значительно растут как эффективность предлагаемого метода (при увеличении n от 1 до 50 прирост ROC AUC составил 26 %, а Recall@100 — 84 %), так и эффективность метода Фишера (ROC AUC — 67 %, Recall@100 — 353 %). Такой рост значений метрик качества свидетельствует о том, что дополнительная информация (больше число экспериментов) улучшает определение вершин активного модуля. Предлагаемый метод устанавливает активные модули преимущественно эффективнее, чем метод Фишера по обоим метрикам качества, особенно при n от 1 до 10. При $n = 50$ результаты были схожи для обоих методов.

При увеличении n в случае $\alpha = 0,8$, когда бета-равномерное распределение ближе к равномерному, незначительно растет как эффективность предлагаемого метода (при увеличении n от 1 до 50 прирост ROC AUC составил 10 %, Recall@100 — 42 %), так и эффективность метода Фишера (ROC AUC — 6 %, Recall@100 — 17 %). Предлагаемый метод определяет активные модули эффективнее, чем метод Фишера по обоим метрикам качества для всех n .

Для вариантов параметров $\alpha \in \{0,2; 0,8\}$ и $n \in \{1, 3, 10\}$ предлагаемый метод демонстрирует более высокие значения ROC AUC и Recall@100, чем метод Фишера. Однако при $\alpha = 0,2$ и $n = 50$ метод Фишера показал более высокое значение метрики ROC AUC, что указывает на то, что в этих условиях информации, полученной только из экспериментов, уже достаточно, а учет взаимодействий в графе становится избыточным. Дальнейшее повышение числа экспериментов n является нецелесообразным для данного набора параметров G и α в графе InWebIM.

Суммарно, предлагаемый метод демонстрирует более высокую или равную точность определения вершин активного модуля по сравнению с ранжированием вершин на основе метода Фишера, что подтвержда-

ется лучшими значениями ROC AUC (рис. 2, *a–d*) и Recall@100 (рис. 2, *e–h*) в большинстве симуляций с ростом значения α .

Предлагаемый метод может быть применен к различным однородным биологическим графам. В качестве примера были выбраны два независимых графа: граф нейронов в головном мозге человека BigBrain (177 584 вершины и 15 669 037 ребер) и генный граф вида живых организмов *Caenorhabditis elegans* (HS-LC) (4227 вершин и 39 484 ребра).

К обоим графам был применен предлагаемый метод. Приращения метрик ROC AUC и Recall@100 для предлагаемого метода были аналогичны приращениям этих метрик на графе InWebIM по сравнению с методом Фишера за исключением большого числа экспериментов ($n = 50$). При таком числе экспериментов и высокой информативности каждого отдельного эксперимента ($\alpha < 0,6$) метод Фишера несет недостаточную информацию о принадлежности вершины к активному модулю. Информация о топологии графа в таком случае становится избыточной, что приводит к отсутствию прироста ROC AUC предлагаемого метода по сравнению с методом Фишера как в случае графа BigBrain (рис. 3), так и снижению качества ROC AUC в среднем на 6 % для генного графа (HS-LC) при $\alpha < 0,6$ (рис. 4). Recall@100 в среднем ухудшается на 8 % для графа BigBrain и на 10 % для графа (HS-LC) при $\alpha < 0,6$.

Валидация предлагаемого метода с использованием результатов биологических экспериментов

Генетические исследования шизофрении

Один из основных методов определения связи между ДНК и болезнью — GWAS. Результатом GWAS является множество локусов (отрезков ДНК, включающих от одного до нескольких генов) и их статистическую значимость.

1. Возникает задача приоритизации генов внутри локуса по отношению к болезни [16, 17].

2. Систематическое сравнение точности метода Фишера и предлагаемого метода проводилось на примере решения задачи из п. 1. Было использовано

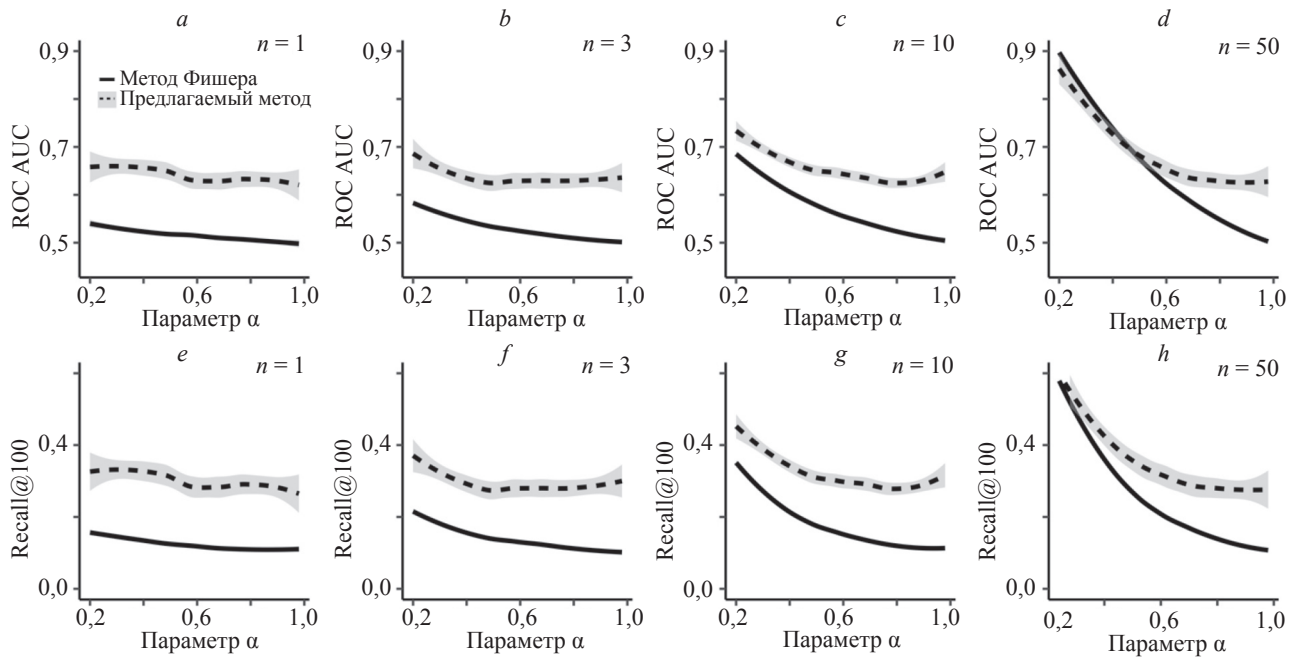


Рис. 2. Результаты симуляций для InWebIM. Сравнение метрик ROC AUC и Recall@100 в зависимости от параметра α бета-равномерного распределения и числа экспериментов. ROC AUC (a–d); Recall@100 (e–h)

Fig. 2. Simulation results for InWebIM. Comparison of ROC AUC and Recall@100 metrics vs. the parameter α of the beta-uniform distribution and the number of experiments. ROC AUC (a–d); Recall@100 (e–h)

GWAS по шизофрении [18]. 49 p -значений для каждого гена были получены методом, описанным в [19, 20] из GWAS по шизофрении. В качестве генов активного модуля были выбраны 105 известных генов шизофрении из работ [21–23]. В качестве генов вне активного модуля выбраны гены, расположенные на расстоя-

нии 200 000 пар оснований от середины каждого гена активного модуля. Каждый локус включал не менее двух генов. Один эксперимент соответствовал одному из 49 полученных p -значений для каждого гена. Эксперимент выбирался случайно с помощью генератора псевдослучайных чисел.

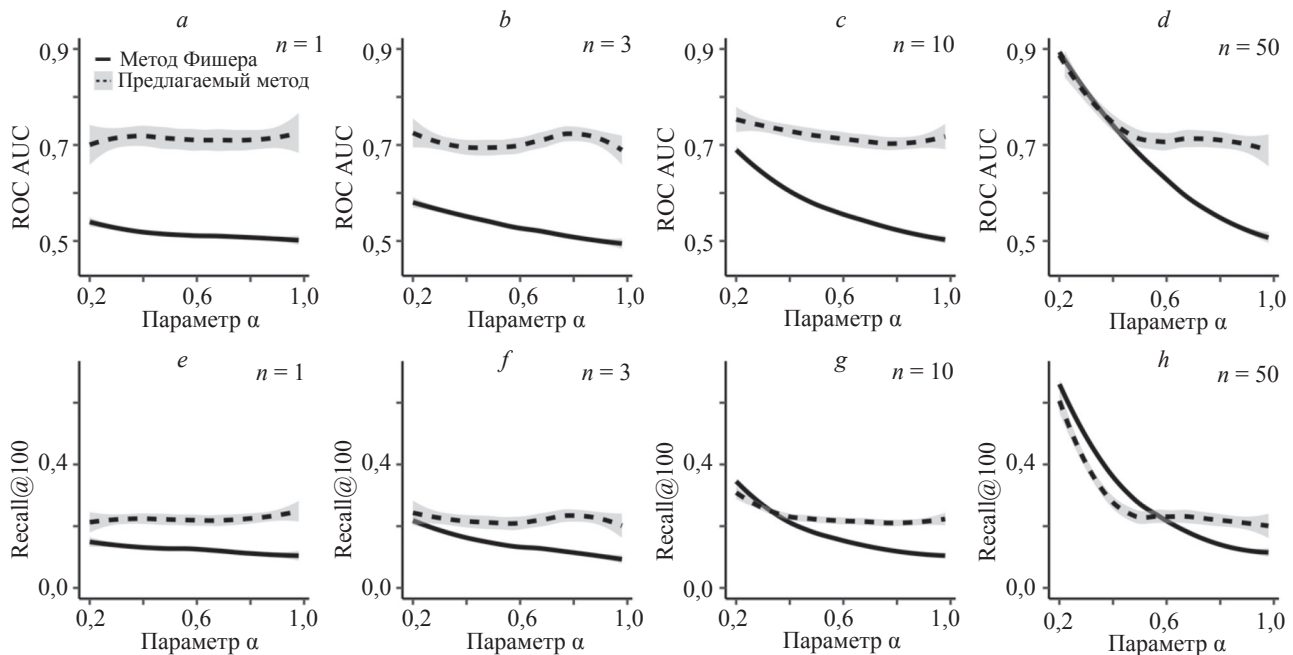


Рис. 3. Результаты симуляций для графа реконструкции мозга человека BigBrain. Сравнение метрик ROC AUC и Recall@100 в зависимости от параметра α бета-равномерного распределения и числа экспериментов.

ROC AUC (a–d); Recall@100 (e–h)

Fig. 3. Simulation results for the BigBrain reconstruction network. Comparison of ROC AUC and Recall@100 metrics vs. the parameter α of the beta-uniform distribution and the number of experiments. ROC AUC (a–d); Recall@100 (e–h)

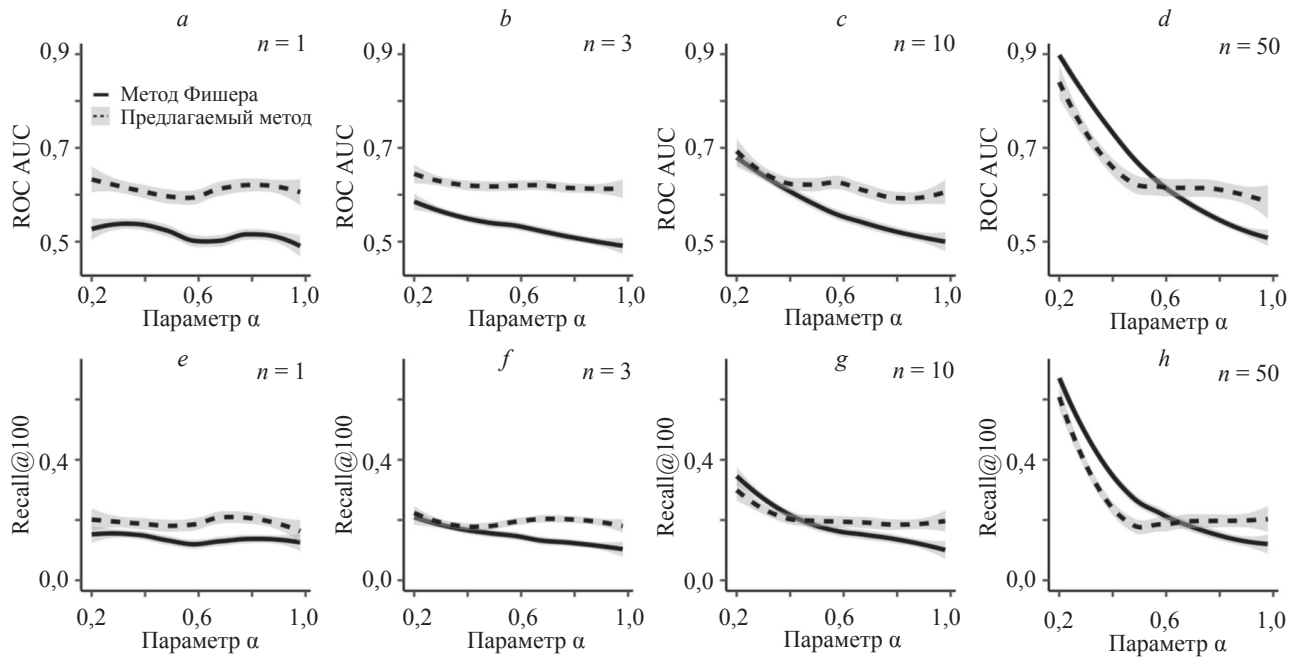


Рис. 4. Результаты симуляций для генного графа вида живых организмов *Caenorhabditis elegans*. Сравнение метрик ROC AUC и Recall@100 в зависимости от параметра α бета-равномерного распределения и числа экспериментов.

ROC AUC (a–d); Recall@100 (e–h)

Fig. 4. Simulation results for the gene network of *Caenorhabditis elegans*. Comparison of ROC AUC and Recall@100 metrics vs. the parameter α of the beta-uniform distribution and the number of experiments. ROC AUC (a–d); Recall@100 (e–h)

3. Применен предлагаемый метод с учетом того, что результаты метода МКМЦ, нормализованы в диапазоне от нуля до единицы для каждого локуса независимо. В качестве биологического графа использован граф белок-белковых взаимодействий InWebIM, где каждому белку соответствовал один ген. Известными генами активного модуля для предлагаемого метода назначались 10 случайных генов из 105 известных. Остальные гены использовались для валидации предлагаемого метода. Этот метод и метод Фишера были применены 100 раз.

4. Точность метода Фишера и предлагаемого метода оценивалась с помощью метрик качества ROC AUC и Recall@200 (доля правильно определенных генов активного модуля в первых 200 вершинах).

5. Из рассмотрения коробчатых диаграмм следует, что предлагаемый метод в среднем показал результаты лучше, чем метод Фишера. ROC AUC предлагаемого метода для одного эксперимента был на 13 % выше, для трех экспериментов на 8 % выше, для 10 экспериментов на 5 % выше, для 49 экспериментов на 6 % выше (рис. 5, a). Recall@200 для предлагаемого метода был

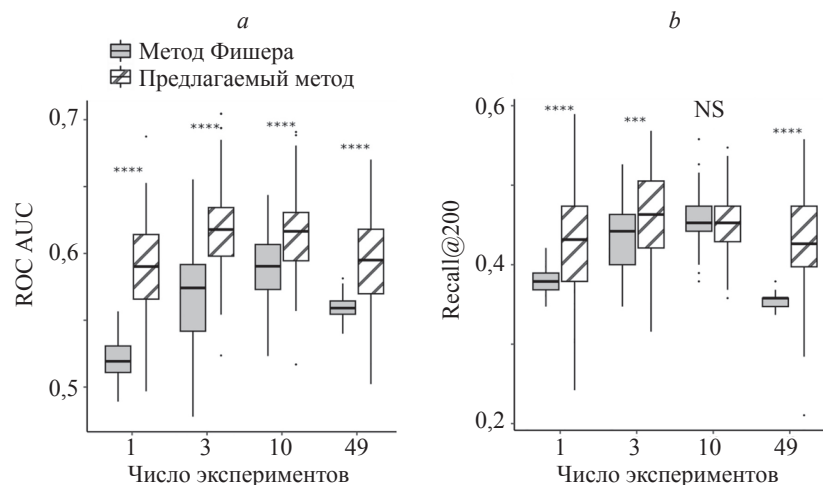


Рис. 5. Коробчатые диаграммы для предлагаемого метода и метода Фишера для GWAS по шизофрении. ROC AUC (a), Recall@200 (b); t-тест использован для сравнения метрик между методами (**** — p -значение $< 0,0001$; *** — p -значение $< 0,001$; NS — незначимо)

Fig. 5. Application of the proposed method and Fisher's method for GWAS on schizophrenia. ROC AUC (a), Recall@200 (b); the t-test was used to compare metrics between methods (**** — p -value < 0.0001 ; *** — p -value < 0.001 ; NS — not significant)

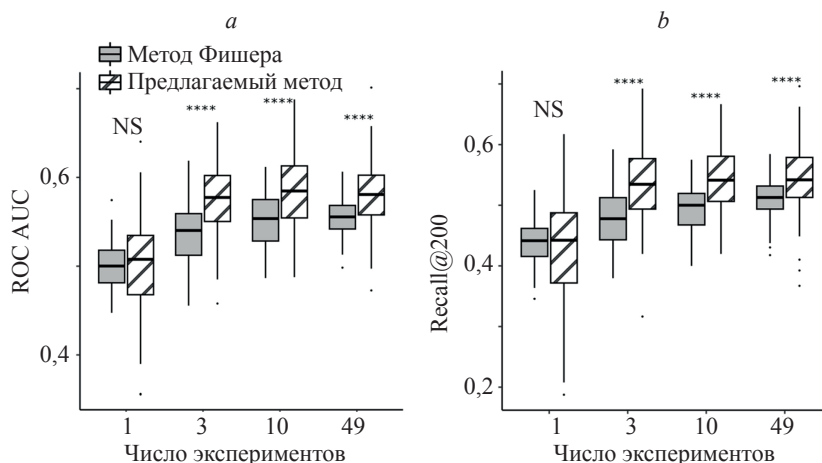


Рис. 6. Коробчатые диаграммы для предлагаемого метода и метода Фишера для данных GWAS по пристрастию к курению. ROC AUC (a), Recall@200 (b); t-тест использован для сравнения метрик между методами (**** — p -значение $< 0,0001$; NS — незначимо)

Fig. 6. Application of the proposed method and Fisher's method to GWAS data on smoking addiction. ROC AUC (a), Recall@200 (b); the t-test was used to compare metrics between methods (**** — p -value < 0.0001 ; NS — not significant)

лучше метода Фишера в среднем на 12 % для одного эксперимента, на 6 % для трех экспериментов и на 20 % для 49 экспериментов. В случае 10 экспериментов Recall@200 предлагаемого метода и метода Фишера не отличался (рис. 5, b).

Генетические исследования в Российской популяции

Состоялся релиз первой базы данных GWAS исследований в России (Биобанк России) [24]. Метод Фишера и предлагаемый метод были применены для такого фенотипа, как генетические риски пристрастия к курению из Биобанка России¹. 175 известных гена для данного фенотипа были взяты из GWAS-каталога². В качестве известных генов активного модуля для предлагаемого метода были выбраны 9 случайных генов из 175 известных. Остальные гены использовались для валидации предлагаемого метода. В качестве биологического графа применен граф белок-белковых взаимодействий InWebIM. ROC AUC для предлагаемого метода не отличался от метода Фишера для одного эксперимента. Recall@200 для предлагаемого метода был лучше метода Фишера в среднем на 11 % для трех экспериментов, на 9 % для 10 экспериментов и на 6 % для 49 экспериментов (рис. 6, b).

Генетические исследования кожной меланомы

В [25] описаны результаты исследования, в котором оценивалось влияние экспрессии каждого отдельного гена на выживаемость пациентов с кожной меланомой в двух независимых группах. К этим результатам были применены метод Фишера и предлагаемый метод. В качестве известных генов использованы 11 иммунных генов из [25]. Биологический граф построен из данных

экспрессии генов методом семплирования байесовских сетей [26].

Для того чтобы оценить правильность полученного активного модуля, гены были ранжированы по предсказанным значениям принадлежности к активному модулю предлагаемым методом и методом Фишера независимо. Затем к первой 1000 генов из каждого списка независимо применен метод обогащения биологических путей [27]. Всего было выделено 62 биологических пути. Результаты статистической значимости для метода обогащения биологических путей на основе результатов предлагаемого метода и метода Фишера показаны на рис. 7. Каждая точка соответствует статистической значимости для метода обогащения биоло-

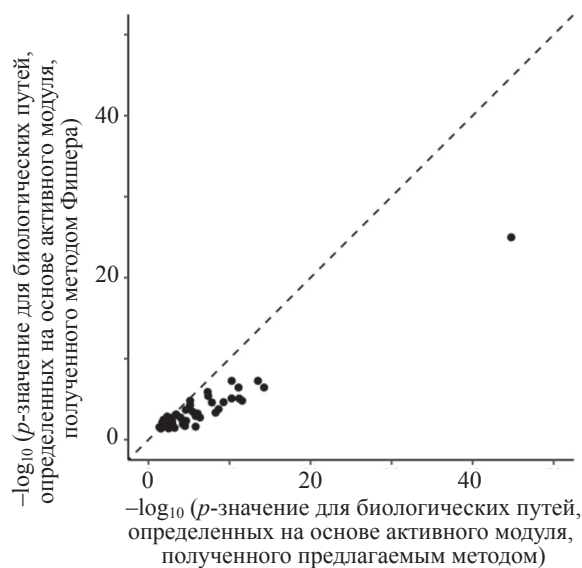


Рис. 7. Оценка качества полученного активного модуля кожной меланомы предлагаемым методом и методом Фишера

Fig. 7. The quality of the obtained active module of skin melanoma by the proposed method and the Fisher method

¹ Биобанк России [Электронный ресурс]. Режим доступа: <https://biobankrus.almazovcentre.ru/pheno/SMNE>, свободный. Яз. рус. (дата обращения: 15.12.2024).

² GWAS-catalog [Электронный ресурс]. Режим доступа: <https://www.ebi.ac.uk/gwas/>, свободный. Яз. англ. (дата обращения: 15.12.2024).

гических путей. Смещение точек вправо относительно диагональной кривой (рис. 7, пунктирная кривая), свидетельствует о том, что результаты метода обогащения биологических путей, полученные на основании результатов предлагаемого метода, статистически надежнее, чем на основании результатов, полученных методом Фишера.

Заключение

В работе предложен метод определения активных модулей в биологических графах с использованием многокомпонентных весов вершин. Этот метод основан на последовательном применении метода Фишера для интеграции данных нескольких биологических экспериментов, метода Монте-Карло по схеме марковских цепей для генерации вероятностных выборок связанных подграфов, и машинного обучения — градиентного бустинга для учета топологии графа.

Экспериментальная оценка, выполненная на симулированных активных модулях графа белок-белковых взаимодействий InWebIM, а также на независимых биологических графах, таких как граф реконструкции мозга BigBrain и генный граф вида живых организ-

мов — *Caenorhabditis elegans*, продемонстрировала значительное повышение точности предлагаемого метода по сравнению с методом Фишера, который не учитывает топологию графа. В частности, было показано, что точность определения активных модулей предлагаемым методом увеличивается с ростом числа учитываемых биологических экспериментов.

Валидация на реальных данных, полученных в генетических исследованиях шизофрении и курения, а также на экспрессионных данных пациентов с меланомой, подтверждает преимущества предлагаемого метода: он демонстрирует более высокое качество идентификации активных модулей по сравнению с методом Фишера.

Важным ограничением предложенного метода является предположение о независимости статистик из разных биологических экспериментов. Если между экспериментами существует сильная корреляция (например, одинаковые образцы или пересекающиеся данные), классический метод Фишера может переоценивать значимость итоговых результатов. В подобных случаях необходимы дополнительные корректировки или расширенные подходы к объединению статистик, учитывающие наличие корреляций между экспериментами.

Литература

1. Wang S., Wu R., Lu J., Jiang Y., Huang T., Cai Y.D. Protein-protein interaction networks as miners of biological discovery // *Proteomics*. 2022. V. 22. N 15-16. P. e2100190. <https://doi.org/10.1002/pmic.202100190>
2. Rao X., Dixon R.A. Co-expression networks for plant biology: why and how // *Acta Biochimica et Biophysica Sinica*. 2019. V. 51. N 10. P. 981–988. <https://doi.org/10.1093/abbs/gmz080>
3. Rawls K., Dougherty B.V., Papin J. Metabolic network reconstructions to predict drug targets and off-target effects // *Methods in Molecular Biology*. 2020. V. 2088. P. 315–330. https://doi.org/10.1007/978-1-0716-0159-4_14
4. Dittrich M.T., Klau G.W., Rosenwald A., Dandekar T., Müller T. Identifying functional modules in protein-protein interaction networks: an integrated exact approach // *Bioinformatics*. 2008. V. 24. N 13. P. i223–i231. <https://doi.org/10.1093/bioinformatics/btn161>
5. Zhu Q.M., Hsu Y.H.H., Lassen F.H., MacDonald B.T., Stead S., Malolepsza E., Kim A., Li T., Mizoguchi T., Schenone M., Guzman G., Tanenbaum B., Fornelos N., Carr S.A., Gupta R.M., Ellinor P.T., Lage K. Protein interaction networks in the vasculature prioritize genes and pathways underlying coronary artery disease // *Communications Biology*. 2024. V. 7. N 1. P. 87. <https://doi.org/10.1038/s42003-023-05705-1>
6. Nehme R., Pietiläinen O., Artomov M., Tegtmeyer M., Valakh V., Lehtonen L., Bell C., Singh T., Trehan A., Sherwood J. et. al. The 22q11.2 region regulates presynaptic gene-products linked to schizophrenia // *Nature Communications*. 2022. V. 13. N 1. P. 3690. <https://doi.org/10.1038/s41467-022-31436-8>
7. Alexeev N., Isomurodov J., Sukhov V., Korotkevich G., Sergushichev A. Markov chain Monte Carlo for active module identification problem // *BMC Bioinformatics*. 2020. V. 21. Suppl. 6. P. 261. <https://doi.org/10.1186/s12859-020-03572-9>
8. Усольцев Д.А., Молотков И.И., Артемов Н.Н., Сергушичев А.А., Шалыто А.А. Применение марковских цепей Монте-Карло и машинного обучения для поиска активного модуля в биологических графах // *Научно-технический вестник информационных технологий, механики и оптики*. 2024. Т. 24. № 6. С. 962–971. <https://doi.org/10.17586/2226-1494-2024-24-6-962-971>
9. Kim T.K., Park J.H. More about the basic assumptions of t-test: normality and sample size // *Korean Journal of Anesthesiology*. 2019. V. 72. N 4. P. 331–335. <https://doi.org/10.4097/kja.d.18.00292>

References

1. Wang S., Wu R., Lu J., Jiang Y., Huang T., Cai Y.D. Protein-protein interaction networks as miners of biological discovery. *Proteomics*, 2022, vol. 22, no. 15-16, P. e2100190. <https://doi.org/10.1002/pmic.202100190>
2. Rao X., Dixon R.A. Co-expression networks for plant biology: why and how. *Acta Biochimica et Biophysica Sinica*, 2019, vol. 51, no. 10, pp. 981–988. <https://doi.org/10.1093/abbs/gmz080>
3. Rawls K., Dougherty B.V., Papin J. Metabolic network reconstructions to predict drug targets and off-target effects. *Methods in Molecular Biology*, 2020, vol. 2088, pp. 315–330. https://doi.org/10.1007/978-1-0716-0159-4_14
4. Dittrich M.T., Klau G.W., Rosenwald A., Dandekar T., Müller T. Identifying functional modules in protein-protein interaction networks: an integrated exact approach. *Bioinformatics*, 2008, vol. 24, no. 13, pp. i223–i231. <https://doi.org/10.1093/bioinformatics/btn161>
5. Zhu Q.M., Hsu Y.H.H., Lassen F.H., MacDonald B.T., Stead S., Malolepsza E., Kim A., Li T., Mizoguchi T., Schenone M., Guzman G., Tanenbaum B., Fornelos N., Carr S.A., Gupta R.M., Ellinor P.T., Lage K. Protein interaction networks in the vasculature prioritize genes and pathways underlying coronary artery disease. *Communications Biology*, 2024, vol. 7, no. 1, pp. 87. <https://doi.org/10.1038/s42003-023-05705-1>
6. Nehme R., Pietiläinen O., Artomov M., Tegtmeyer M., Valakh V., Lehtonen L., Bell C., Singh T., Trehan A., Sherwood J. et. al. The 22q11.2 region regulates presynaptic gene-products linked to schizophrenia. *Nature Communications*, 2022, vol. 13, no. 1, pp. 3690. <https://doi.org/10.1038/s41467-022-31436-8>
7. Alexeev N., Isomurodov J., Sukhov V., Korotkevich G., Sergushichev A. Markov chain Monte Carlo for active module identification problem. *BMC Bioinformatics*, 2020, vol. 21, Suppl. 6, pp. 261. <https://doi.org/10.1186/s12859-020-03572-9>
8. Usoltsev D.A., Molotkov I.I., Artomov M.N., Sergushichev A.A., Shalyto A.A. Application of Markov chain Monte Carlo and machine learning for identifying active modules in biological graphs. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2024, vol. 24, no. 6, pp. 962–971. (in Russian). <https://doi.org/10.17586/2226-1494-2024-24-6-962-971>
9. Kim T.K., Park J.H. More about the basic assumptions of t-test: normality and sample size. *Korean Journal of Anesthesiology*, 2019, vol. 72, no. 4, pp. 331–335. <https://doi.org/10.4097/kja.d.18.00292>

10. Pounds S., Morris S.W. Estimating the occurrence of false positives and false negatives in microarray studies by approximating and partitioning the empirical distribution of p-values // *Bioinformatics*. 2003. V. 19. N 10. P. 1236–1242. <https://doi.org/10.1093/bioinformatics/btg148>
11. Ham H., Park T. Combining p-values from various statistical methods for microbiome data // *Frontiers in Microbiology*. 2022. V. 13. P. 990870. <https://doi.org/10.3389/fmicb.2022.990870>
12. Li T., Wernersson R., Hansen R.B., Horn H., Mercer J., Slodkowitz G., Workman C.T., Rigina O., Rapacki K., Stærfeldt H.H., Brunak S., Jensen T.S., Lage K. A scored human protein-protein interaction network to catalyze genomic interpretation // *Nature Methods*. 2017. V. 14. N 1. P. 61–64. <https://doi.org/10.1038/nmeth.4083>
13. Rossi R., Ahmed N. The network data repository with interactive graph analytics and visualization // *Proc. of the 29th AAAI Conference on Artificial Intelligence*. 2015. V. 29. N 1. <https://doi.org/10.1609/aaai.v29i1.9277>
14. Amunts K., Lepage C., Borgeat L., Mohlberg H., Dickscheid T., Rousseau M.É., Bludau S., Bazin P.L., Lewis L.B., Oros-Peusquens A.M., Shah N.J., Lippert T., Zilles K., Evans A.C. BigBrain: an ultrahigh-resolution 3D human brain model // *Science*. 2013. V. 340. N 6139. P. 1472–1475. <https://doi.org/10.1126/science.1235381>
15. Cho A., Shin J., Hwang S., Kim C., Shim H., Kim H., Kim H., Lee I. WormNet v3: a network-assisted hypothesis-generating server for *Caenorhabditis elegans* // *Nucleic Acids Research*. 2014. V. 42. N W1. P. W76–W82. <https://doi.org/10.1093/nar/gku367>
16. Zhu Z., Zhang F., Hu H., Bakshi A., Robinson M.R., Powell J.E., Montgomery G.W., Goddard M.E., Wray N.R., Visscher P.M., Yang J. Integration of summary data from GWAS and eQTL studies predicts complex trait gene targets // *Nature Genetics*. 2016. V. 48. N 5. P. 481–487. <https://doi.org/10.1038/ng.3538>
17. Usoltsev D., Molotkov I., Artomov M. A meta-predictor for causal gene identification in GWAS overcomes limitations of existing computational approaches // *American Society of Human Genetics (Complex Traits and Polygenic Disorders Poster Friday Session)*. 2024.
18. Pardiñas A.F., Holmans P., Pocklington A.J., Escott-Price V., Ripke S., Carrera N., Legge S.E., Bishop S., Cameron D., Hamshire M.L., et al. Common schizophrenia alleles are enriched in mutation-intolerant genes and in regions under strong background selection // *Nature Genetics*. 2018. V. 50. N 3. P. 381–389. <https://doi.org/10.1038/s41588-018-0059-2>
19. Barbeira A.N., Dickinson S.P., Bonazzola R., Zheng J., Wheeler H.E., Torres J.M., Torstenson E.S., Shah K.P., Garcia T., Edwards T.L., Stahl E.A., Huckins L.M., Nicolae D.L., Cox N.J., Im H.K. Exploring the phenotypic consequences of tissue specific gene expression variation inferred from GWAS summary statistics // *Nature Communications*. 2018. V. 9. N 1. P. 1825. <https://doi.org/10.1038/s41467-018-03621-1>
20. Urbut S.M., Wang G., Carbonetto P., Stephens M. Flexible statistical methods for estimating and testing effects in genomic studies with multiple conditions // *Nature Genetics*. 2019. V. 51. N 1. P. 187–195. <https://doi.org/10.1038/s41588-018-0268-8>
21. Kolosov N., Daly M.J., Artomov M. Prioritization of disease genes from GWAS using ensemble-based positive-unlabeled learning // *European Journal of Human Genetics*. 2021. V. 29. N 10. P. 1527–1535. <https://doi.org/10.1038/s41431-021-00930-w>
22. Lam M., Chen C-Y., Li Z., Martin A.R., Bryois J., Ma X., Gaspar H., Ikeda M., Benyamin B., Brown B.C. et al. Comparative genetic architectures of schizophrenia in East Asian and European populations // *Nature Genetics*. 2019. V. 51. N 12. P. 1670–1678. <https://doi.org/10.1038/s41588-019-0512-x>
23. Singh T., Poterba T., Curtis D., Akil H., Al Eissa M., Barchas J.D., Bass N., Bigdeli T.B., Breen G., Bromet E.J. et al. Rare coding variants in ten genes confer substantial risk for schizophrenia // *Nature*. 2022. V. 604. N 7906. P. 509–516. <https://doi.org/10.1038/s41586-022-04556-w>
24. Usoltsev D., Kolosov N., Rotar O., Loboda A., Boyarinova M., Moguchaya E., Kolesova E., Erina A., Tolkunova K., Rezapova V., Molotkov I. et al. Complex trait susceptibilities and population diversity in a sample of 4,145 Russians // *Nature Communications*. 2024. V. 15. N 1. P. 6212. <https://doi.org/10.1038/s41467-024-50304-1>
25. Usoltsev D., Njauw C.N., Ji Z., Kumar R., Sergushichev A., Zhang S., Shlyakhto E., Daly M.J., Artomov M., Tsao H. Analysis of variants induced by combined ex vivo irradiation and in vivo tumorigenesis
10. Pounds S., Morris S.W. Estimating the occurrence of false positives and false negatives in microarray studies by approximating and partitioning the empirical distribution of p-values. *Bioinformatics*, 2003, vol. 19, no. 10, pp. 1236–1242. <https://doi.org/10.1093/bioinformatics/btg148>
11. Ham H., Park T. Combining p-values from various statistical methods for microbiome data. *Frontiers in Microbiology*, 2022, vol. 13, pp. 990870. <https://doi.org/10.3389/fmicb.2022.990870>
12. Li T., Wernersson R., Hansen R.B., Horn H., Mercer J., Slodkowitz G., Workman C.T., Rigina O., Rapacki K., Stærfeldt H.H., Brunak S., Jensen T.S., Lage K. A scored human protein-protein interaction network to catalyze genomic interpretation. *Nature Methods*, 2017, vol. 14, no. 1, pp. 61–64. <https://doi.org/10.1038/nmeth.4083>
13. Rossi R., Ahmed N. The network data repository with interactive graph analytics and visualization. *Proc. of the 29th AAAI Conference on Artificial Intelligence*, 2015, vol. 29, no. 1. <https://doi.org/10.1609/aaai.v29i1.9277>
14. Amunts K., Lepage C., Borgeat L., Mohlberg H., Dickscheid T., Rousseau M.É., Bludau S., Bazin P.L., Lewis L.B., Oros-Peusquens A.M., Shah N.J., Lippert T., Zilles K., Evans A.C. BigBrain: an ultrahigh-resolution 3D human brain model. *Science*, 2013, vol. 340, no. 6139, pp. 1472–1475. <https://doi.org/10.1126/science.1235381>
15. Cho A., Shin J., Hwang S., Kim C., Shim H., Kim H., Kim H., Lee I. WormNet v3: a network-assisted hypothesis-generating server for *Caenorhabditis elegans*. *Nucleic Acids Research*, 2014, vol. 42, no. W1, pp. W76–W82. <https://doi.org/10.1093/nar/gku367>
16. Zhu Z., Zhang F., Hu H., Bakshi A., Robinson M.R., Powell J.E., Montgomery G.W., Goddard M.E., Wray N.R., Visscher P.M., Yang J. Integration of summary data from GWAS and eQTL studies predicts complex trait gene targets. *Nature Genetics*, 2016, vol. 48, no. 5, pp. 481–487. <https://doi.org/10.1038/ng.3538>
17. Usoltsev D., Molotkov I., Artomov M. A meta-predictor for causal gene identification in GWAS overcomes limitations of existing computational approaches. *American Society of Human Genetics (Complex Traits and Polygenic Disorders Poster Friday Session)*, 2024.
18. Pardiñas A.F., Holmans P., Pocklington A.J., Escott-Price V., Ripke S., Carrera N., Legge S.E., Bishop S., Cameron D., Hamshire M.L., et al. Common schizophrenia alleles are enriched in mutation-intolerant genes and in regions under strong background selection. *Nature Genetics*, 2018, vol. 50, no. 3, pp. 381–389. <https://doi.org/10.1038/s41588-018-0059-2>
19. Barbeira A.N., Dickinson S.P., Bonazzola R., Zheng J., Wheeler H.E., Torres J.M., Torstenson E.S., Shah K.P., Garcia T., Edwards T.L., Stahl E.A., Huckins L.M., Nicolae D.L., Cox N.J., Im H.K. Exploring the phenotypic consequences of tissue specific gene expression variation inferred from GWAS summary statistics. *Nature Communications*, 2018, vol. 9, no. 1, pp. 1825. <https://doi.org/10.1038/s41467-018-03621-1>
20. Urbut S.M., Wang G., Carbonetto P., Stephens M. Flexible statistical methods for estimating and testing effects in genomic studies with multiple conditions. *Nature Genetics*, 2019, vol. 51, no. 1, pp. 187–195. <https://doi.org/10.1038/s41588-018-0268-8>
21. Kolosov N., Daly M.J., Artomov M. Prioritization of disease genes from GWAS using ensemble-based positive-unlabeled learning. *European Journal of Human Genetics*, 2021, vol. 29, no. 10, pp. 1527–1535. <https://doi.org/10.1038/s41431-021-00930-w>
22. Lam M., Chen C-Y., Li Z., Martin A.R., Bryois J., Ma X., Gaspar H., Ikeda M., Benyamin B., Brown B.C. et al. Comparative genetic architectures of schizophrenia in East Asian and European populations. *Nature Genetics*, 2019, vol. 51, no. 12, pp. 1670–1678. <https://doi.org/10.1038/s41588-019-0512-x>
23. Singh T., Poterba T., Curtis D., Akil H., Al Eissa M., Barchas J.D., Bass N., Bigdeli T.B., Breen G., Bromet E.J. et al. Rare coding variants in ten genes confer substantial risk for schizophrenia. *Nature*, 2022, vol. 604, no. 7906, pp. 509–516. <https://doi.org/10.1038/s41586-022-04556-w>
24. Usoltsev D., Kolosov N., Rotar O., Loboda A., Boyarinova M., Moguchaya E., Kolesova E., Erina A., Tolkunova K., Rezapova V., Molotkov I. et al. Complex trait susceptibilities and population diversity in a sample of 4,145 Russians. *Nature Communications*, 2024, vol. 15, no. 1, pp. 6212. <https://doi.org/10.1038/s41467-024-50304-1>
25. Usoltsev D., Njauw C.N., Ji Z., Kumar R., Sergushichev A., Zhang S., Shlyakhto E., Daly M.J., Artomov M., Tsao H. Analysis of variants induced by combined ex vivo irradiation and in vivo tumorigenesis

- suggests a role for the ZNF831 p.R1393Q variant in cutaneous melanoma development // *Journal of Investigative Dermatology*. 2024. In Press, Corrected Proof. <https://doi.org/10.1016/j.jid.2024.08.042>
26. Лобода А.А. Метод графовой кластеризации для совместного анализа данных генотипирования и экспрессии генов: диссертация на соискание ученой степени кандидата технических наук. СПб., 2022, 232 с.
 27. Subramanian A., Tamayo P., Mootha V.K., Mukherjee S., Ebert B.L., Gillette M.A., Paulovich A., Pomeroy S.L., Golub T.R., Lander E.S., Mesirov J.P. Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles // *Proc. of the National Academy of Sciences of the United States of America*. 2005. V. 102. N 43. P. 15545–15550. <https://doi.org/10.1073/pnas.0506580102>
 - suggests a role for the ZNF831 p.R1393Q variant in cutaneous melanoma development. *Journal of Investigative Dermatology*, 2024, in Press, corrected proof. <https://doi.org/10.1016/j.jid.2024.08.042>
 26. Loboda A.A. A method of graphical clustering for joint analysis of genotyping and expression data. *Dissertation for the degree of candidate of technical sciences*. St. Petersburg, 2022, 232 p. (in Russian)
 27. Subramanian A., Tamayo P., Mootha V.K., Mukherjee S., Ebert B.L., Gillette M.A., Paulovich A., Pomeroy S.L., Golub T.R., Lander E.S., Mesirov J.P. Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proc. of the National Academy of Sciences of the United States of America*, 2005, vol. 102, no. 43, pp. 15545–15550. <https://doi.org/10.1073/pnas.0506580102>

Авторы

Усольцев Дмитрий Андреевич — старший научный сотрудник, Институт геномной медицины, Детская больница Нейшенвайд, Колумбус, 43205, США; аспирант, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, [sc 57279360300](https://orcid.org/0000-0001-8072-310X), <https://orcid.org/0000-0001-8072-310X>, dusoltsev.27@gmail.com

Молотков Иван Игоревич — старший научный сотрудник, Институт геномной медицины, Детская больница Нейшенвайд, Колумбус, 43205, США; аспирант, Медицинский колледж Университета штата Огайо, Колумбус, 43210, США, [sc 58651494600](https://orcid.org/0009-0008-3566-0160), <https://orcid.org/0009-0008-3566-0160>, ivan.molotkov@nationwidechildrens.org

Артемов Никита Николаевич — кандидат химических наук, доцент, главный исследователь, Институт геномной медицины, Детская больница Нейшенвайд, Колумбус, 43205, США; профессор педиатрии, Медицинский колледж Университета штата Огайо, Колумбус, 43210, США, [sc 36542095500](https://orcid.org/0000-0001-5282-8764), <https://orcid.org/0000-0001-5282-8764>, mykyta.artomov@nationwidechildrens.org

Сергушичев Алексей Александрович — кандидат технических наук, доцент, Университет Вашингтона в Сент-Луисе, Сент-Луис, 63110, США, [sc 55772694000](https://orcid.org/0000-0003-1159-7220), <https://orcid.org/0000-0003-1159-7220>, asergushichev@wustl.edu

Шалыто Анатолий Абрамович — доктор технических наук, профессор, главный научный сотрудник, профессор, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, [sc 56131789500](https://orcid.org/0000-0002-2723-2077), <https://orcid.org/0000-0002-2723-2077>, anatoly.shalyto@gmail.com

Authors

Dmitrii A. Usoltsev — Senior Researcher, Institute for Genomic Medicine, Nationwide Children's Hospital, Columbus, 43205, USA; PhD Student, ITMO University, Saint Petersburg, 197101, Russian Federation, [sc 57279360300](https://orcid.org/0000-0001-8072-310X), <https://orcid.org/0000-0001-8072-310X>, dusoltsev.27@gmail.com

Ivan I. Molotkov — Senior Researcher, Institute for Genomic Medicine, Nationwide Children's Hospital, Columbus, 43205, USA; PhD Student, The Ohio State University College of Medicine, Columbus, 43210, USA, [sc 58651494600](https://orcid.org/0009-0008-3566-0160), <https://orcid.org/0009-0008-3566-0160>, ivan.molotkov@nationwidechildrens.org

Mykyta N. Artomov — PhD (Chemistry), Associate Professor, Chief Researcher, Institute for Genomic Medicine, Nationwide Children's Hospital, Columbus, 43205, USA; Professor of Pediatrics, The Ohio State University College of Medicine, Columbus, 43210, USA, [sc 36542095500](https://orcid.org/0000-0001-5282-8764), <https://orcid.org/0000-0001-5282-8764>, mykyta.artomov@nationwidechildrens.org

Alexey A. Sergushichev — PhD, Associate Professor, Washington University School of Medicine in St. Louis, St. Louis, 63110, USA, [sc 55772694000](https://orcid.org/0000-0003-1159-7220), <https://orcid.org/0000-0003-1159-7220>, asergushichev@wustl.edu

Anatoly A. Shalyto — D.Sc., Chief Researcher, Full Professor, ITMO University, Saint Petersburg, 197101, Russian Federation, [sc 56131789500](https://orcid.org/0000-0002-2723-2077), <https://orcid.org/0000-0002-2723-2077>, anatoly.shalyto@gmail.com

Статья поступила в редакцию 09.02.2025
Одобрена после рецензирования 14.04.2025
Принята к печати 26.05.2025

Received 09.02.2025
Approved after reviewing 14.04.2025
Accepted 26.05.2025



Работа доступна по лицензии
Creative Commons
«Attribution-NonCommercial»

doi: 10.17586/2226-1494-2025-25-3-498-507

Modeling of nonlocal porous functionally graded nanobeams under moving loads

Ridha A. Ahmed¹, Wael N. Abdullah², Nadhim M. Faleh³✉, Mamoon A. Al-Jaafari⁴

^{1,2,3,4} Engineering College of Al-Mustansiriyah University, Baghdad, 10001, Iraq

¹ ridhalwan@uomustansiriyah.edu.iq, <https://orcid.org/0000-0002-6666-6564>

² wael.najm65@gmail.com, <https://orcid.org/0000-0001-9599-4279>

³ dr.nadhim@uomustansiriyah.edu.iq✉, <https://orcid.org/0000-0003-0003-1643>

⁴ dr.mamoonal-jaafari@uomustansiriyah.edu.iq, <https://orcid.org/0009-0005-9597-9825>

Abstract

This study focuses on the dynamic response of porous functionally graded nanomaterials to moving loads. The analysis was performed using two approaches: the Ritz method with the help of the benefits achieved by employing Chebyshev polynomials in the cosine form and the differential quadrature method with further inverse Laplace transformation. Both approaches utilize the formulation of a nano-thin beam considering an improved higher-order beam model and nonlocal strain gradient theory with two characteristic length scales, referred to as nonlocality and strain gradient length scales. Power-law dependencies steer the constituent designs of pore-graded materials toward pore factors that influence pore volume either with a uniform or non-uniform distribution of pores. Moreover, a variable scale modulus was adopted to further improve accuracy by considering the scale effects for graded nano-thin beams. The first part of the study addresses the equation of motion, which is solved by applying the Ritz technique with Chebyshev polynomials. In the second part, the governing equations for nanobeams are discussed where the differential quadrature method is used to discretise them further, and the inverse Laplace transform is used to obtain the dynamic deflections. The results of the present study elucidate the effects of the moving load speed, nonlocal strain gradient factors, porosity, pore number and distribution, and elastic medium on the dynamic deflection of functionally graded nanobeams.

Keywords

design, material, gradient porous, moving load, nonlocal strain, porous

Acknowledgments

The authors sincerely thank Mustansiriya University (www.uomustansiriyah.edu.iq) in Baghdad, Iraq, for their invaluable support in advancing this work.

For citation: Ahmed R.A., Abdullah W.N., Faleh N.M., Al-Jaafari M.A. Modeling of nonlocal porous functionally graded nanobeams under moving loads. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2025, vol. 25, no. 3, pp. 498–507. doi: 10.17586/2226-1494-2025-25-3-498-507

УДК 004.942

Моделирование нелокальных пористых функционально-градиентных нанобалок под действием движущихся нагрузок

Ридха А. Ахмед¹, Ваэл Н. Абдулла², Надхим М. Фалех³✉, Мамун А. Аль-Джаафари⁴

^{1,2,3,4} Университет Аль-Мустансирия, Багдад, 10001, Ирак

¹ ridhalwan@uomustansiriyah.edu.iq, <https://orcid.org/0000-0002-6666-6564>

² wael.najm65@gmail.com, <https://orcid.org/0000-0001-9599-4279>

³ dr.nadhim@uomustansiriyah.edu.iq✉, <https://orcid.org/0000-0003-0003-1643>

⁴ dr.mamoonal-jaafari@uomustansiriyah.edu.iq, <https://orcid.org/0009-0005-9597-9825>

Аннотация

Исследован динамический отклик пористых функционально-градиентных наноматериалов на движущиеся нагрузки. Анализ проводился с использованием двух подходов: метода Ритца с использованием преимуществ, достигаемых за счет внедрения полиномов Чебышева в форме косинуса, и метода дифференциальных квадратур с последующим обратным преобразованием Лапласа. Оба подхода применяют модель нанотонкой балки с учетом улучшенной модели более высокого порядка и нелокальной теории градиента деформации с двумя характерными шкалами длины, называемыми шкалами нелокальности и градиента деформации. Степенные

© Ahmed R.A., Abdullah W.N., Faleh N.M., Al-Jaafari M.A., 2025

зависимости ориентируются на составных конструкциях пористых материалов с различными факторами: объем пор, равномерное или неравномерное распределение пор. Принят переменный модуль масштаба для дальнейшего повышения точности за счет учета масштабных эффектов для градуированных нанотонких балок. В первой части работы рассмотрено уравнение движения, которое решено путем применения техники Рунге с полиномами Чебышева. Во второй части выполнен анализ уравнения для нанобалок, где применен метод дифференциальных квадратур для их дальнейшей дискретизации, а обратное преобразование Лапласа использовано для получения динамических прогибов. Результаты исследования позволяют объяснить влияние скорости движущейся нагрузки, нелокальных факторов градиента деформации, пористости, числа и распределения пор, а также упругой среды на динамический прогиб функционально-градуированных нанобалок.

Ключевые слова

конструкция, материал, градиентная пористость, подвижная нагрузка, нелокальная деформация, пористость

Благодарности

Авторы благодарят Университет Мустансирия (www.uomustansiriyah.edu.iq) Багдад-Ирак за поддержку в настоящей работе.

Ссылка для цитирования: Ахмед Р.А., Абдулла В.Н., Фалех Н.М., Аль-Джаафари М.А. Моделирование нелокальных пористых функционально-градиентных нанобалок под действием движущихся нагрузок // Научно-технический вестник информационных технологий, механики и оптики. 2025. Т. 25, № 3. С. 498–507 (на англ. яз.). doi: 10.17586/2226-1494-2025-25-3-498-507

Introduction

Mechanical characteristics of nanoscale structures, including nanoscale beams or nanosized plates, based on non-classic elasticity theories have been an important case study in recent decades [1]. This is because of the broad application of nanoscale structures in nanosensors or nanoelectromechanical systems. The most familiar theory for modeling nanoscale structures is nonlocal elasticity [2]. In these theories, scale factors were introduced to interpolate the influence of small sizes. Pursuant to Nonlocal Elasticity Theory (NET), the stress field must be nonlocal because the relationship between the stress and strain at a point depends on the strains at that point and the surrounding points [3]. The nonlocality of the stress field is examined using nonlocal parameters. Any value of a nonlocal parameter can be determined through experiments or numerical simulations. However, deriving the values of these nonlocal parameters using these methods can be complicated and time-consuming. Consequently, considerable academic research on the static and dynamic aspects of nanoscale structures has been conducted as parameter-based studies that rely on specific theoretical values for the nonlocality modulus [3–9].

Recently, several theoretical studies and experiments have reported that small-sized effects must be characterized via stiffness-increasing mechanisms or strain gradient fields [10]. This assertion differs from that of nonlocal elasticity in which stiffness reduction behavior has been stated [11–16]. However, the effects of the reduction and increment on the structural stiffness at the nanoscale can be considered in the context of the nonlocality strain gradient theory. According to the Nonlocal Strain Gradient Theory (NSGT), two criterion factors termed nonlocal and strain gradient factors have been utilized to provide an excellent description of small-size effects. The static and dynamic properties of nanobeams and other nanostructures have been studied extensively using [17, 18]. The impact of different loadings on the vibration behavior of nanobeams has become an essential case study in recent years. Some of these loadings include harmonic forces, impulsive loads, and moving loads from the overhead view of the nanosized beam. Several authors have investigated the

forced vibrations of nanobeams owing to harmonic and impulsive loads in the context of the NET and NSGT. However, causing vibrations of nanosized beams owing to mobile loads has become important because of nano-sensing and nano-probing applications [19–26]. It has been realized that the dynamic deflections of a nanobeam owing to moving loads increase with the inclusion of nonlocal parameters [27].

Method of formulation

As discussed, the influences of the reduction and increment in the structural stiffness at the nanoscale can be considered in the context of the NSGT. According to that theory, two scale criteria termed nonlocal and strain gradient factors have been utilized to provide an excellent description of small-size effects. First, it is essential to define the stress-field components as follows:

$$\sigma_{ij} = \sigma_{ij}^{(0)} - \nabla \sigma_{ij}^{(1)},$$

where σ is stress components; $\sigma_{ij}^{(1)}$ and $\sigma_{ij}^{(0)}$ are stress components; ∇ is gradient; $\nabla \sigma_{ij}$ stress gradients.

ϵ_{kl} and strain gradients $\nabla \epsilon_{kl}$ can be defined as:

$$\sigma_{ij}^{(0)} = \int_V C_{ijkl} \alpha_0(x, x', q_0 u) \epsilon'_{kl}(x') dx',$$

$$\sigma_{ij}^{(1)} = \mathcal{J}^2 \int_V C_{ijkl} \alpha_1(x, x', q_1 u) \nabla \epsilon'_{kl}(x') dx',$$

where the symbol C_{ijkl} is used for the elastic coefficients, $q_0 u$ and $q_1 u$ are used to define the nonlocal effects, and \mathcal{J} is Strain gradient factor which introduces the influences of the strain gradients. If the nonlocal functions $\alpha_0(x, x', q_0 u)$ and $\alpha_1(x, x', q_1 u)$ can satisfy the conditions introduced by [28], the relationship between the stresses and strains in the context of NSGT becomes:

$$[1 - (q_1 a u)^2 \nabla^2][1 - (q_0 u)^2 \nabla^2] \sigma_{ij} = C_{ijkl} [1 - (q_1 u)^2 \nabla^2] \epsilon_{kl} - C_{ijkl} \mathcal{J}^2 [1 - (q_0 u)^2 \nabla^2] \nabla^2 \epsilon_{kl}, \quad (1a)$$

where ∇^2 is the Laplacian operator; qu is combined nonlocal parameter. \mathcal{J}^2 is strain gradient length scale. The above relation can be simplified by assuming $q_1 u = q_0 u = qu$ as:

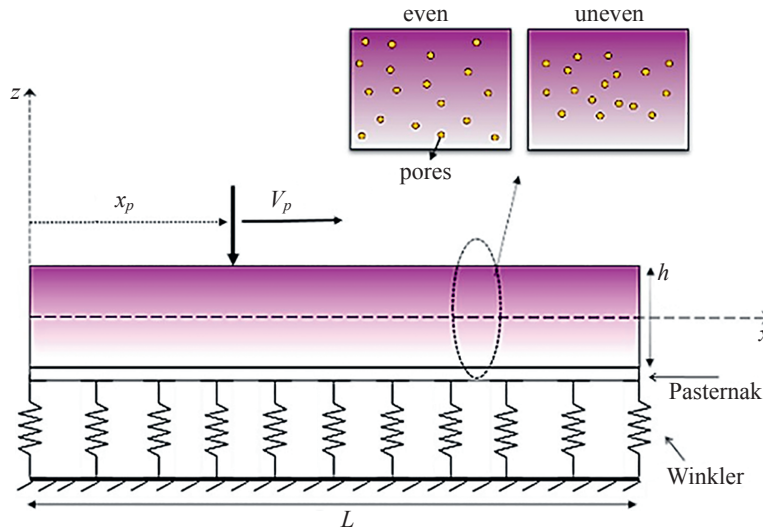


Fig. 1. Gradation nano-thickness beam undersurface of a movable particle

$$[1 - (qu)^2 \nabla^2] \sigma_{ij} = C_{ijkl} [1 - \nabla^2] \varepsilon_{kl}. \quad (1b)$$

The distribution of Functionally Graded (FG) materials in structures can be mathematically modeled using the power law or Mori-Tanaka models. Using the power-law function, it is possible to describe the continuous gradation of material properties with good accuracy. However, the Mori-Tanaka scheme provided more accurate results as reported in some studies. In the first step, we assumed an FG nanobeam of length L and thickness h as shown in Fig. 1, where x_p denotes the instantaneous position of the moving load along the beam length, while V_p indicates the constant velocity of that moving load. Fig. 1 illustrates the configuration of the stepped nanobeam under the influence of the moving load, with x_p and V_p corresponding to the load position and velocity, respectively. According to the Mori-Tanaka FG model, the effective properties of the FG nanobeam, including its bulk modulus K_e and shear modulus G_e , are considered, by the following:

$$\begin{aligned} \frac{K_e - K_m}{K_c - K_m} &= \frac{V_c}{1 + V_m(K_c - K_m)/(K_m + 4G_m/3)}, \\ \frac{G_e - G_m}{G_c - G_m} &= \frac{V_c}{1 + V_m(G_c - G_m)/[(G_m + G_m(9K_m + 8G_m)/6(K_m + 2G_m))]} \end{aligned}$$

Here, K_m and K_c denote the bulk modules of the metal and ceramic parts, respectively; G_m and G_c denote the shear modules of the metal and ceramic parts, respectively; V_c is the volume portion of the ceramic constituent; and V_m is the volume portion of the metal constituent. The following relationship with the volume portion of the metal constituent:

$$\begin{aligned} V_c + V_m &= 1, \\ V_c(z) &= \left(\frac{z}{h} + \frac{1}{2} \right)^p, \end{aligned}$$

where h is the thickness of the nanobeam; z is coordinated along the thickness direction and p is the power-law exponent (material gradient index). The distribution of the FG material along the thickness path depends on the value of the material gradient index (p). According to the above relations, one can express the effectual Young's modulus (E), Poisson's ratio (ν), and mass density (ρ) of such material as:

$$E(z) = \frac{9K_e G_e}{3K_e + G_e},$$

$$\nu(z) = \frac{3K_e - 2G_e}{6K_e + 2G_e},$$

$$\rho(z) = \rho_c V_c + \rho_m V_m.$$

As mentioned, another approach for modeling a foam gradient material is the power-law function. According to the power-law model, each material property can be defined using the following relation:

$$\Psi(z) = (\Psi_c - \Psi_m) \left(\frac{z}{h} + \frac{1}{2} \right)^p + \Psi_m,$$

where Ψ_m and Ψ_c are the material exclusivity of metallic and ceramic.

Based on the above relationship, it is possible to designate Young's modulus (E), Poisson's ratio (ν), and strain gradient factor (J_2). Thus, the effect of graded nonlocality was considered in [29]. However, the above relationships ignore the effect of the porosity inside the FG materials.

Using modified power-law functions, it is possible to model each material property containing porosity volume (ξ) as:

$$\begin{aligned} \Psi(z) &= (\Psi_c - \Psi_m) \left(\frac{z}{h} + \frac{1}{2} \right)^p + \Psi_m - (\Psi_c + \Psi_m) \frac{\xi}{2} \\ &\text{for even distribution,} \end{aligned}$$

$$\Psi(z) = (\Psi_c - \Psi_m) \left(\frac{z}{h} + \frac{1}{2} \right)^p + \Psi_m - \frac{\xi}{2} (\Psi_c + \Psi_m) \left(1 - \frac{2|z|}{h} \right)$$

for uneven distribution,

where $\Psi(z)$ is generic material property (e.g., Young's modulus, density). The ξ is a constant with $\xi = \pi/h$. Higher-order beam theories are useful in establishing the governing equations of beams by considering the effects of shear deformation. A well-known theory is the refined beam theory which has the following form of a displacement field ($u_x, 0, u_z$):

$$u_x(x, z) = u(x) - (z - s^*) \frac{\partial w_b}{\partial x} - [H(z) - s^{**}] \frac{\partial w_s}{\partial x},$$

$$u_z(x, z) = w_b(x) + w_s(x).$$

The displacement field contains the axial displacement (u), bending displacement (w_b), shear displacement (w_s), s^* is the neutral-axis location parameter, s^{**} is the neutral-axis location parameter, and $H(z)$ is the trigonometric shear strain function. Moreover, to determine the neutral-axis location of the FG beam, it is necessary to calculate

$$s^* = \frac{\int_{-0.5h}^{0.5h} E(z)z dz}{\int_{-0.5h}^{0.5h} E(z) dz},$$

$$s^{**} = \frac{\int_{-0.5h}^{0.5h} E(z)H(z) dz}{\int_{-0.5h}^{0.5h} E(z) dz}.$$

The function $H(z)$ has been selected as:

$$H(z) = z - \sin(\xi z)/\xi.$$

The strains derived based on the displacement factors can be expressed as follows:

$$\epsilon_{xx} = \frac{\partial u}{\partial x} - (z - s^*) \frac{\partial^2 w_b}{\partial x^2},$$

$$\gamma_{xz} = \Omega(z) \frac{\partial w_s}{\partial x}, \text{ where } \Omega(z) = 1 - \frac{dH(z)}{dz},$$

ϵ_{xx} is the normal strain component, and γ_{xz} is the shear strain component.

Note that $\Omega(z) = 1 - dH(z)/dz$. Subsequently, the generalized version of Hamilton's formula implies that:

$$\int_0^t \delta(U + V - K) dt = 0. \quad (2)$$

Hitherto, U may be termed the strain energy, V may be termed the work accomplished by external loading, and K may be termed the kinetic energy. The strain energy variation may be composed of

$$\delta U = \int_V \sigma_{ij} \delta \epsilon_{ij} dV = \int_V \left(\sigma_{xx} \delta \epsilon_{xx} + \sigma_{xx}^{(1)} \delta \nabla \epsilon_{xx} + \sigma_{xz} \delta \gamma_{xz} + \sigma_{xz}^{(1)} \delta \nabla \gamma_{xz} \right) dV. \quad (3)$$

$$\delta U = \int_0^L \left(N \frac{\partial \delta u}{\partial x} - M_b \frac{\partial^2 \delta w_b}{\partial x^2} - M_s \frac{\partial^2 \delta w_s}{\partial x^2} + Q \frac{\partial \delta w_s}{\partial x} \right) dx, \quad (4)$$

where N is the resultant force along the beam axial direction; M_b is the moment due to bending stresses; M_s is the moment associated with shear deformation; Q is the transverse force due to shear stresses. Wherein all forces and moments revealed within the overhead equation might be composed as:

$$N_{xx} = \int_{-h/2}^{h/2} \left(\sigma_{xx}^{(0)} - \nabla \sigma_{xx}^{(1)} \right) dz = N_{xx}^{(0)} - \nabla N_{xx}^{(1)},$$

$$M_{xx}^b = \int_{-h/2}^{h/2} z \left(\sigma_{xx}^{(0)} - \nabla \sigma_{xx}^{(1)} \right) dz = M_{xx}^{b(0)} - \nabla M_{xx}^{b(1)},$$

$$M_{xx}^s = \int_{-h/2}^{h/2} f \left(\sigma_{xx}^{(0)} - \nabla \sigma_{xx}^{(1)} \right) dz = M_{xx}^{s(0)} - \nabla M_{xx}^{s(1)},$$

$$Q_{xz} = \int_{-h/2}^{h/2} g \left(\sigma_{xz}^{(0)} - \nabla \sigma_{xz}^{(1)} \right) dz = Q_{xz}^{(0)} - \nabla Q_{xz}^{(1)},$$

where:

$$N_{xx}^{(0)} = \int_{-h/2}^{h/2} \left(\sigma_{xx}^{(0)} \right) dz, \quad N_{xx}^{(1)} = \int_{-h/2}^{h/2} \left(\sigma_{xx}^{(1)} \right) dz,$$

$$M_{xx}^{b(0)} = \int_{-h/2}^{h/2} z \left(\sigma_{xx}^{b(0)} \right) dz, \quad M_{xx}^{b(1)} = \int_{-h/2}^{h/2} z \left(\sigma_{xx}^{b(1)} \right) dz,$$

$$M_{xx}^{s(0)} = \int_{-h/2}^{h/2} f \left(\sigma_{xx}^{s(0)} \right) dz, \quad M_{xx}^{s(1)} = \int_{-h/2}^{h/2} f \left(\sigma_{xx}^{s(1)} \right) dz,$$

$$Q_{xz}^{(0)} = \int_{-h/2}^{h/2} g \left(\sigma_{xz}^{(0)} \right) dz, \quad Q_{xz}^{(1)} = \int_{-h/2}^{h/2} g \left(\sigma_{xz}^{(1)} \right) dz.$$

The variation in the work accomplished via practical loading can be formulated as follows:

$$\delta V = \int_0^L (q \delta(w_b + w_s)) dx. \quad (7)$$

The outward lateral loading q might be composed as:

$$q = +k_w(w_b + w_s) - k_p \frac{\partial^2 (w_b + w_s)}{\partial x^2} + q_{dynamic}. \quad (8)$$

In the above relation, k_w and k_p , respectively, denote the Winkler and Pasternak parts of elastic substrate; $q_{dynamic} = P_0 \delta(x - V_p t)$ is the applied force owing to the moving load, where V_p denotes the velocity of the moving load, P_0 exhibits the valence of the mobile force and x exhibits the applying status of the mobile force. The variation in kinetic energy is formulated as follows:

$$\delta K = \int_0^L \left(I_0 \left[\frac{\partial u}{\partial t} \frac{\partial \delta u}{\partial t} + \left(\frac{\partial w_b}{\partial t} + \frac{\partial w_s}{\partial t} \right) \left(\frac{\partial \delta w_b}{\partial t} + \frac{\partial \delta w_s}{\partial t} \right) \right] - I_1 \left(\frac{\partial u}{\partial t} \frac{\partial^2 \delta w_b}{\partial x \partial t} + \frac{\partial^2 w_b}{\partial x \partial t} \frac{\partial \delta u}{\partial t} \right) + I_2 \left(\frac{\partial^2 w_b}{\partial x \partial t} \frac{\partial^2 \delta w_b}{\partial x \partial t} \right) - J_1 \left(\frac{\partial u}{\partial t} \frac{\partial^2 \delta w_s}{\partial x \partial t} + \frac{\partial^2 w_s}{\partial x \partial t} \frac{\partial \delta u}{\partial t} \right) + K_2 \left(\frac{\partial^2 w_s}{\partial x \partial t} \frac{\partial^2 \delta w_s}{\partial x \partial t} \right) + J_2 \left(\frac{\partial^2 w_b}{\partial x \partial t} \frac{\partial^2 \delta w_s}{\partial x \partial t} + \frac{\partial^2 w_s}{\partial x \partial t} \frac{\partial^2 \delta w_b}{\partial x \partial t} \right) \right) dx. \quad (9)$$

Where,

$$(I_0, I_1, J_1, I_2, J_2, K_2) = \int_{-h/2}^{h/2} (1, z - s^*, (z - s^*)^2, H - s^{**}, (z - s^*)(H - s^{**}), (H - s^{**})^2) \rho(z) dz. \quad (10)$$

In Eq. (10): I_0 : Mass per unit length of the nanobeam (integral of density $\rho(z)$ over the thickness); I_1 : First-order moment coupling axial and bending displacements; J_1 : Second-order moment linked to shear deformation; I_2 : Second-order moment associated with bending deformation; J_2 : Cross-term moment coupling bending and shear deformations; and K_2 : Higher-order moment linked to shear deformation. The '1' under the integral represents the unity term contributing to the total mass (I_0) by integrating $\rho(z)$ across the thickness. Finally, the governing equations for a pore-dependent and nano-thickness beam under a movable particle can be obtained by substituting Eqs. (3)–(9) into (2) as.

$$\begin{aligned} \frac{\partial N}{\partial x} &= I_0 \frac{\partial^2 u}{\partial t^2} - I_1 \frac{\partial^3 w_b}{\partial x \partial t^2} - J_1 \frac{\partial^3 w_s}{\partial x \partial t^2}, \\ \frac{\partial^2 M_b}{\partial x^2} - q_{dynamic} &= + I_0 \left(\frac{\partial^2 w_b}{\partial t^2} + \frac{\partial^2 w_s}{\partial t^2} \right) + I_1 \frac{\partial^3 u}{\partial x \partial t^2} - \\ &- I_2 \frac{\partial^4 w_b}{\partial x^2 \partial t^2} - J_2 \frac{\partial^4 w_s}{\partial x^2 \partial t^2} + k_w (w_b + w_s) - k_p \frac{\partial^2 (w_b + w_s)}{\partial x^2}, \\ \frac{\partial^2 M_s}{\partial x^2} + \frac{\partial Q}{\partial x} - q_{dynamic} &= + I_0 \left(\frac{\partial^2 w_b}{\partial t^2} + \frac{\partial^2 w_s}{\partial t^2} \right) + \\ &+ J_1 \frac{\partial^3 u}{\partial x \partial t^2} - J_2 \frac{\partial^4 w_b}{\partial x^2 \partial t^2} - K_2 \frac{\partial^4 w_s}{\partial x^2 \partial t^2} + \\ &+ k_w (w_b + w_s) - k_p \frac{\partial^2 (w_b + w_s)}{\partial x^2}. \end{aligned}$$

Using Eq. (1a) and (1b), it is possible to establish the stress-strain relations of a higher-order refined FG nanobeam in the context of the NSGT as

$$\begin{aligned} \sigma_{xx} - (qu)^2 \frac{\partial^2 \sigma_{xx}}{\partial x^2} &= \left(1 - J^2 \frac{\partial^2}{\partial x^2} \right) E(z) \varepsilon_{xx}, \\ \sigma_{xz} - (qu)^2 \frac{\partial^2 \sigma_{xz}}{\partial x^2} &= \left(1 - J^2 \frac{\partial^2}{\partial x^2} \right) G(z) \gamma_{xz}. \end{aligned}$$

Nano-thickness beam contains the above stresses which result in the following forces and moments by integrating Eqs. (4) and (5) over the thickness. Thus, the internal strength can be determined by:

$$\begin{aligned} \{A^{local}, B^{local}, B_s^{local}, D^{local}, D_s^{local}, H_s^{local}\} &= \\ &= \int_{-h/2}^{h/2} E(z) \{1, (z - s^*), (H - s^{**}), (z - s^*)^2 (z - s^*)(H - s^{**}), \\ &\quad (H - s^{**})^2\} dz, \\ A_s^{local} &= \int_{-h/2}^{h/2} \Omega^2 G(z) dz, \end{aligned}$$

$$\begin{aligned} \{A^{SGT}, B^{SGT}, B_s^{SGT}, D^{SGT}, D_s^{SGT}, H_s^{SGT}\} &= \\ &= \int_{-h/2}^{h/2} E(z) J^2(z) \{1, (z - s^*), (H - s^{**}), (z - s^*)^2, \\ &\quad (z - s^*)(H - s^{**}), (H - s^{**})^2\} dz, \\ A_s^{SGT} &= \int_{-h/2}^{h/2} \Omega^2 J^2(z) G(z) dz. \end{aligned}$$

The dynamic response problem of a nanothick beam is solved in this section basing on the Chebyshev–Ritz method. First, the field components can be assumed to have the following form:

$$u(x, t) = R^u(x) \sum_{m=1}^{\infty} U_m P_m(x) e^{i\omega_n t}, \quad (11)$$

$$w_b(x, t) = R^{w_b}(x) \sum_{m=1}^{\infty} W_{bm} P_m(x) e^{i\omega_n t}, \quad (12)$$

$$w_s(x, t) = R^{w_s}(x) \sum_{m=1}^{\infty} W_{sm} P_m(x) e^{i\omega_n t}.$$

Consider the following essential boundary conditions. Simply-supported (SS) edges:

$$w = \frac{\partial^2 w}{\partial x^2} = 0 \text{ at } x = +0.5L, -0.5L.$$

In addition, $P_m(x)$ is the n -th Chebyshev polynomial of the 1st type and is expressed as

$$P_m(x) = \cos \left[(m-1) \arccos \left(\frac{2x}{L} \right) \right].$$

Note that $R^i(\odot = u, w_b, w_s)$ denotes the functions associated with the essential boundary conditions. In addition, functions can be introduced in the following form:

$$R^i(x) = \left(1 + \frac{2x}{L} \right)^{p^*} \left(1 - \frac{2x}{L} \right)^{q^*},$$

where p^* and q^* are affected by the pattern of the edge conditions. For the SS edges, $p^* = q^* = 1$. Representing Eqs. (10)–(12) in weak form, together with their minimization of the indefinite variables U_m , W_{bm} , and W_{sm} , the equation below results in simultaneous algebraic equations with respect to the unknown coefficients.

$$\frac{\partial \Pi}{\partial U_m} = \frac{\partial \Pi}{\partial W_{bm}} = \frac{\partial \Pi}{\partial W_{sm}} = 0.$$

Substituting Eqs. (11) and (12) into the obtained relation with the knowledge that the axial and transverse motions are uncoupled yields

$$\left\{ [K] + \frac{\partial}{\partial t^2} [M] \right\} \begin{Bmatrix} W_{bn} \\ W_{sn} \end{Bmatrix} = \begin{Bmatrix} q_{dynamic} - \mu_c \frac{\partial^2 q_{dynamic}}{\partial x^2} \\ q_{dynamic} - \mu_c \frac{\partial^2 q_{dynamic}}{\partial x^2} \end{Bmatrix}, \quad (13)$$

where $[K]$ and $[M]$ are the toughness and mass matrices of the considered nanosized beam, respectively. In addition,

the non-dimensional substrate modules can be distinguished by

$$K_w = k_w \frac{L^4}{E_m I}, K_p = k_p \frac{L^2}{E_m I}.$$

It can be inferred that the dynamic force is mobile along an upstanding path which yields forcing vibrations, and is distinguished by the following formulation:

$$\begin{aligned} q_{dynamic} &= \sum_{n=1}^{\infty} Q_n \sin \left[\frac{n\pi}{L} x \right], \\ Q_n &= \frac{2}{L} \int_0^L \sin \left[\frac{n\pi}{L} x \right] q(x) dx = \frac{2P_0}{L} \sin \left[\frac{n\pi}{L} x_p \right] = \\ &= \frac{2P_0}{L} \sin \left[\frac{n\pi}{L} V_p t \right], \end{aligned}$$

where Q_n expose the Fourier multipliers and $q(x) = P_0 \delta(x - x_p)$; P_0 exhibits the valence of the mobile force and x_p exhibits the applying status of the mobile force. In addition, V_p represents the celerity of the mobile force.

Finally, based on zero primary status and the Laplace transformation branch of knowledge, Eq. (13) changes to

$$\begin{aligned} \{[\mathbf{K}] + S^2[\mathbf{M}]\} \begin{Bmatrix} \&L[W_{bn}] \\ \&L[W_{sn}] \end{Bmatrix} &= \\ = \begin{Bmatrix} \&L \left[q_{dynamic} - \mu_c \frac{\partial^2 q_{dynamic}}{\partial x^2} \right] \\ \&L \left[q_{dynamic} - \mu_c \frac{\partial^2 q_{dynamic}}{\partial x^2} \right] \end{Bmatrix}. \end{aligned} \quad (14)$$

By solving Eq. (14) using the inverse Laplace transformation approach, the values of the bending (W_{bn}) and shear (W_{sn}) displacements can be derived. However, the total deflection of the nanobeam is the sum of the two displacements, $W = W_{bn} + W_{sn}$. In addition, the normalized deflection and normalized velocity factors may be introduced by

$$\bar{W} = W \frac{100E_m I}{P_0 L^3}, V^* = \frac{V_p}{V_{cr}}, V_{cr} = \frac{\omega_n L}{\pi}.$$

The normalized deflection \bar{W} is introduced to non-dimensionalise the beam dynamic response, enabling comparison across different material and geometric configurations. Similarly, the dimensionless velocity ratio $V^* = V_p/V_{cr}$ characterizes the moving load speed relative to the critical velocity $V_{cr} = \omega_1 L/\pi$, where ω_1 is the fundamental natural frequency of the nanobeam. The critical velocity V_{cr} represents the threshold speed at which the dynamic deflection peaks before sharply decreasing, as observed in Fig. 3. Finally, the normalized time t^* can be expressed as follows:

$$t^* = \frac{V_p t}{L}.$$

Main results

The present section examines the dynamic responses of porous FG nanobeams to moving loads that capture both nonlocality and strain gradient influences. The effects of the load velocity, material gradation, porosity distribution, graded nonlocality, and two-scale factors on the dynamic deflection of the nanobeam were studied in detail. In this study, it was determined that the nonlocality and strain gradient multipliers were not constant for the FG nanobeam. They are variable in the direction of thickness as (z) and $\lambda(z)$ according to Eq. (7). Thus, μ_m and μ_c are the nonlocal multipliers of the metal and ceramic phases, respectively; λ_m and λ_c are the strain gradient multipliers of the metal and ceramic parts, respectively. Accordingly, μ_c/μ_m and λ_c/λ_m , respectively, denote the nonlocal ratio and strain gradient ratio. Further discussion of this issue can be found in the following paragraphs. First, a comparison was made with the work of [19] to validate the vibration frequency of an FG nanobeam based on an NSGT. Therefore, the validation of the first dimensionless vibration frequency $\Omega = \omega L^2 \sqrt{\rho_c/E_c}/h$, at different values of nonlocal and strain gradient multipliers, and an excellent agreement can be obtained between the obtained results and [19].

Fig. 2 shows the time history of normalized dynamic deflection at different values of load velocity factor V^* of 0.1, 0.12, and 0.15. Also, in Fig. 2 the nonlocal parameter ratio (μ_c/μ_m) of 0.5, 1, 1.5, and 2 at the FG index $p = 1$ is shown. It is assumed in this research that the ratio of strain gradient multipliers of the metal and ceramic (λ_c/λ_m) is equal 2. For every value of the load velocity factor, an increment in the nonlocal ratio results in higher values of normalized dynamic deflection. This is owing to the reduced stiffness of FG nanobeam when the nonlocal ratio becomes higher. Such behavior indicates that an FG nanobeam displays stiffness-softening effects when the nonlocal ratio increases. Therefore, the dynamic behavior of an FG nanobeam depends on the completion between the nonlocal values of metal and ceramic phases.

Dynamic deflection of the nanosized beam against the load velocity factor based on NETs, Classic Elasticity Theory (CET), and NSGT is plotted in Fig. 3 at $p = 1$ and $t^* = 0.5$. In the case of CET, it is assumed that $\mu_m = \lambda_m = 0$. Moreover, it is considered for the case of NET that $\mu_m = 0.2$ and $\lambda_m = 0$. In this figure, the nonlocal ratio is assumed to be $\mu_c/\mu_m = 2$. The dynamic deflection is significantly affected by the moving load velocity. The dynamic deflection is augmented with the load velocity factor until it overtakes a summit valence and then drops suddenly after this valence (critical velocity). However, the dynamic deflection and summit point are dependent on the values of the nonlocal and strain gradient multipliers. NET provides the maximum value of dynamic deflection owing to the inclusion of nonlocal effects. However, by incorporating the strain gradient effect, the NSGT yielded smaller deflections than the NET.

A comparison between the dynamic deflection (time history) of the FG nanobeam obtained by the power law and Mori-Tanaka models is shown in Fig. 4, assuming that $p = 1$. In this figure, the load Velocity Factor (V^*) is set to 0.12. It can be observed that the Mori-Tanaka model of the

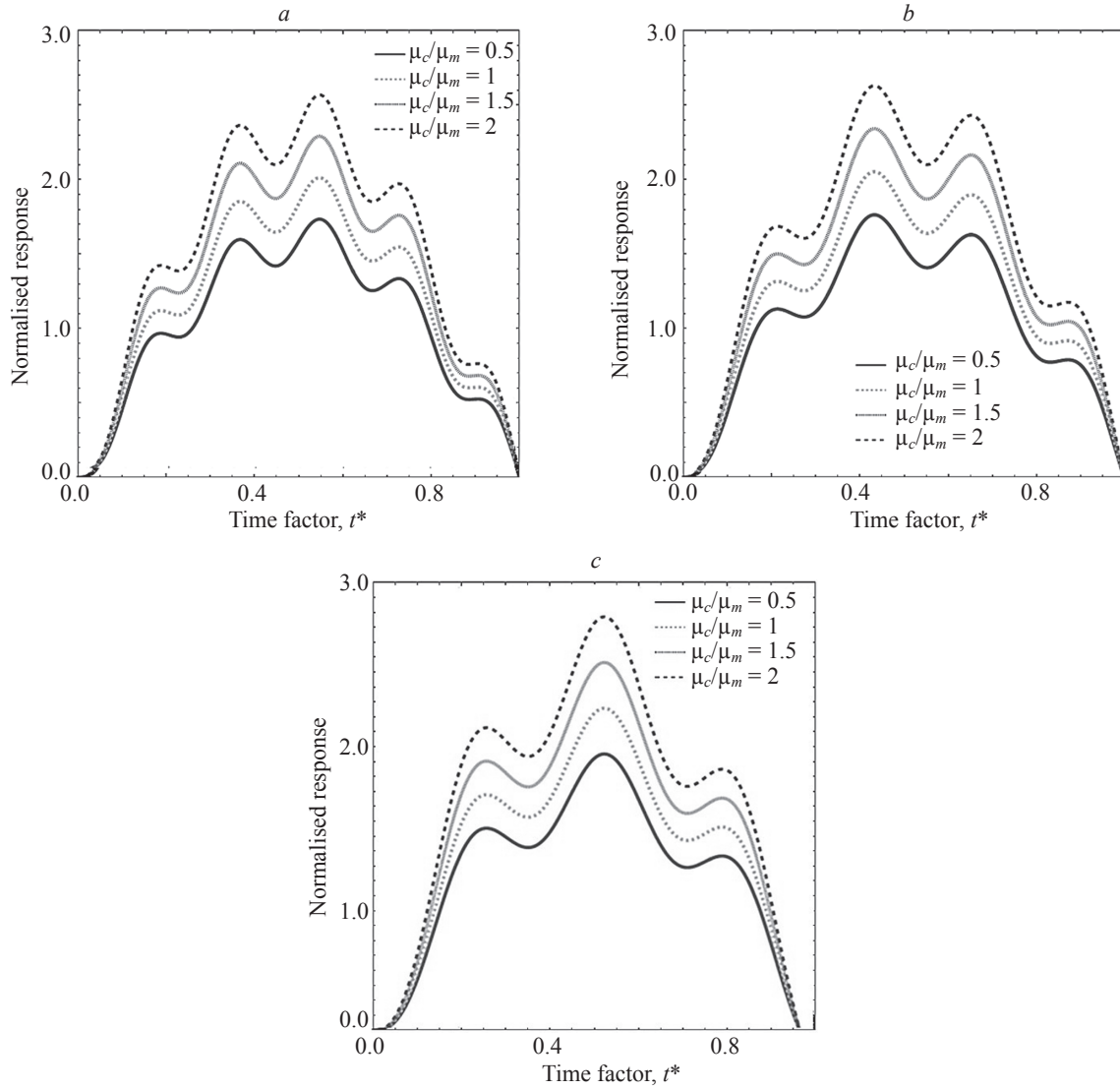


Fig. 2. Time history of normalized dynamic deflection (\bar{W}) for varying nonlocal parameter ratios ($\mu_c/\mu_m = 0.5, 1, 1.5, 2$) at distinct load velocity factors: $V^* = 0.1$ (a); $V^* = 0.12$ (b); $V^* = 0.15$ (c). The FG index is $p = 1$, and the strain gradient ratio is fixed at $\lambda_c/\lambda_m = 2$

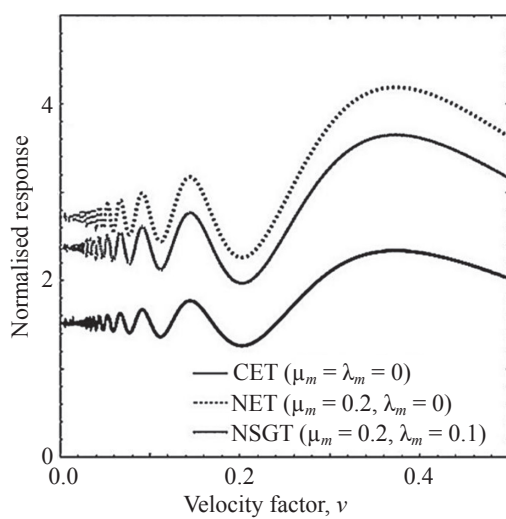


Fig. 3. Nano-thickness unto velocity factor based on distinct elasticity modeling

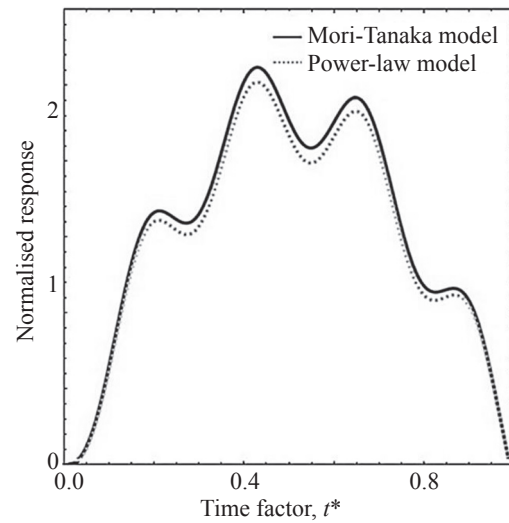


Fig. 4. Thickness beam vs. time factor on distinct gradation modeling

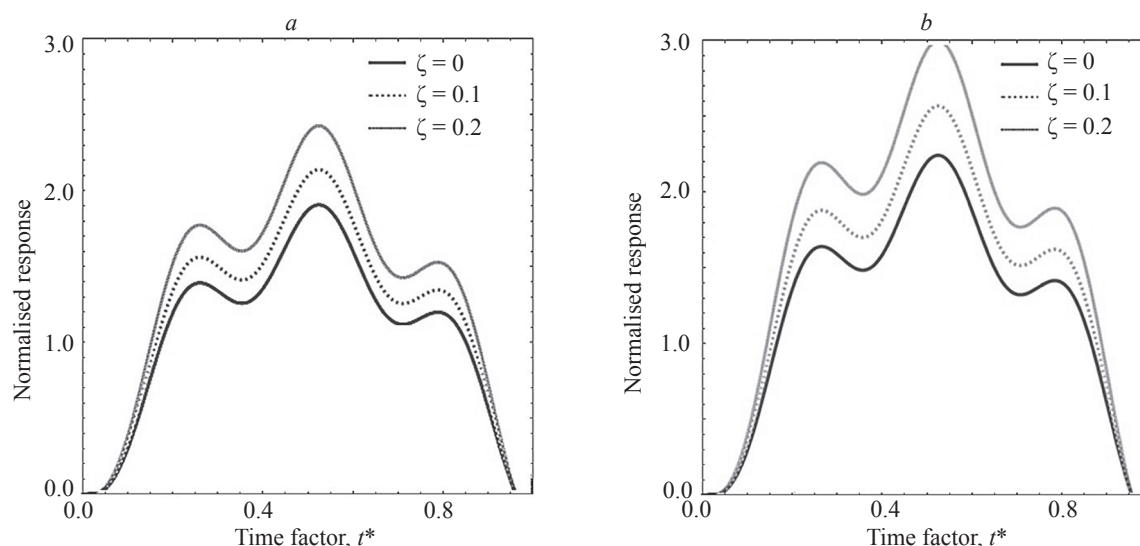


Fig. 5. The nano-thickness beam onto Time factor t^* for $p = 0.5$ (a), 1 (b)

FG materials results in higher values of dynamic deflection than the power-law model. In particular, the FG nanobeam is more flexible because it describes the material properties using the Mori-Tanaka model which is a more reliable scheme for FG materials than the power-law model.

In Fig. 5, the time history of FG nanobeam has been plotted based on various porosity volume fractions ($\zeta = 0, 0.1$, and 0.2) and FG gradient index ($p = 0.5, 1$, and 2) at $V^* = 0.15$. An even porosity dispersion was considered in this figure. The FG nanobeam becomes more flexible at a higher gradient index value because of the higher percentage of the metal constituent compared with the ceramic constituent. Accordingly, the dynamic deflections of the nanobeams increase with increasing gradient index. However, another essential factor in the dynamic response of FG nanobeams is porosity. As the porosity volume (ζ) increases, the value of dynamic deflection becomes higher because porosities inside FG material reduce the structural stiffness.

Conclusions

This article focuses on the dynamic response inspection of a porous Functionally Graded (FG) nanobeam subjected to movable-point loading, considering the effects of graded nonlocality. An altered power-law model was used to investigate the dynamic characteristics of the FG nanobeams, including porosity effects. The nanobeam formulation was based on the higher-order refined beam theory, whereas the side effects were captured according to the Nonlocality Strain Gradient Theory (NSGT). The governing equations were solved using the differential quadrature method and the inverse Laplace transform method. The main findings are summarized as follows:

- An increase in the nonlocal ratio led to higher normalized dynamic deflection values.
- It was reported that NSGT provides smaller deflections than Nonlocal Elasticity Theory.
- The dynamic deflection was augmented with the load velocity until it reached a maximum value, and then dropped after this point, which was called the critical velocity.
- The porosity volume and dynamic deflection increased because the porosity inside the FG material reduced the structural stiffness.

References

1. Khider A.S., Aalsaud A., Faleh N.M., Abd A.K., Al-Jaafari, M.A.A., Fenjan R.M. A review on dynamic characteristics of nonlocal porous FG nanobeams under moving loads. *Steel and Composite Structures*, 2024, vol. 50, no. 1, pp. 15–24. <https://doi.org/10.12989/scs.2024.50.1.015>
2. Raheef K.M., Ahmed R.A., Nayeeif A.A., Fenjan R.M., Faleh N.M. Analyzing dynamic response of nonlocal strain gradient porous beams under moving load and thermal environment. *Geomechanics and Engineering*, 2021, vol. 26, no. 1, pp. 89–99. <https://doi.org/10.12989/gae.2021.26.1.089>
3. Zeighampour H., Beni Y.T. Cylindrical thin-shell model based on modified strain gradient theory. *International Journal of Engineering Science*, 2014, vol. 78, pp. 27–47. <https://doi.org/10.1016/j.ijengsci.2014.01.004>

Литература

1. Khider A.S., Aalsaud A., Faleh N.M., Abd A.K., Al-Jaafari, M.A.A., Fenjan R.M. A review on dynamic characteristics of nonlocal porous FG nanobeams under moving loads // *Steel and Composite Structures*. 2024. V. 50. N 1. P. 15–24. <https://doi.org/10.12989/scs.2024.50.1.015>
2. Raheef K.M., Ahmed R.A., Nayeeif A.A., Fenjan R.M., Faleh N.M. Analyzing dynamic response of nonlocal strain gradient porous beams under moving load and thermal environment // *Geomechanics and Engineering*. 2021. V. 26. N 1. P. 89–99. <https://doi.org/10.12989/gae.2021.26.1.089>
3. Zeighampour H., Beni Y.T. Cylindrical thin-shell model based on modified strain gradient theory // *International Journal of Engineering Science*. 2014. V. 78. P. 27–47. <https://doi.org/10.1016/j.ijengsci.2014.01.004>

4. Li L., Hu Y., Ling L. Flexural wave propagation in small-scaled functionally graded beams via a nonlocal strain gradient theory. *Composite Structures*, 2015, vol. 133, pp. 1079–1092. <https://doi.org/10.1016/j.compstruct.2015.08.014>
5. Zhang B., He Y., Liu D., Shen L., Lei J. Free vibration analysis of four-unknown shear deformable functionally graded cylindrical microshells based on the strain gradient elasticity theory. *Composite Structures*, 2015, vol. 119, pp. 578–597. <https://doi.org/10.1016/j.compstruct.2014.09.032>
6. Arefi M., Zenkour A.M. Free vibration, wave propagation and tension analyses of a sandwich micro/nano rod subjected to electric potential using strain gradient theory. *Materials Research Express*, 2016, vol. 3, no. 11, pp. 115704. <https://doi.org/10.1088/2053-1591/3/11/115704>
7. Lou J., He L., Wu H., Du J. Pre-buckling and buckling analyses of functionally graded microshells under axial and radial loads based on the modified couple stress theory. *Composite Structures*, 2016, vol. 142, pp. 226–237. <https://doi.org/10.1016/j.compstruct.2016.01.083>
8. Zeighampour H., Shojaeian M. Buckling analysis of functionally graded sandwich cylindrical micro/nanoshells based on the couple stress theory. *Journal of Sandwich Structures & Materials*, 2019, vol. 21, no. 3, pp. 917–937. <https://doi.org/10.1177/1099636217703912>
9. Farokhi H., Ghayesh M.H. Nonlinear mechanical behaviour of microshells. *International Journal of Engineering Science*, 2018, vol. 127, pp. 127–144. <https://doi.org/10.1016/j.ijengsci.2018.02.009>
10. Martínez-Criado G. Application of micro-and nanobeams for materials science. *Synchrotron light sources and free-electron lasers: accelerator physics, instrumentation and science applications*, 2016, pp. 1505–1539. https://doi.org/10.1007/978-3-319-14394-1_46
11. Berrabah H.M., Tounsi A., Semmah A., Bedia E.A.A. Comparison of various refined nonlocal beam theories for bending, vibration and buckling analysis of nanobeams. *Structural Engineering and Mechanics*, 2013, vol. 48, no. 3, pp. 351–365. <https://doi.org/10.12989/sem.2013.48.3.351>
12. Aissani K., Bouiadra M.B., Ahouel M., Tounsi A. A new nonlocal hyperbolic shear deformation theory for nanobeams embedded in an elastic medium. *Structural Engineering and Mechanics*, 2015, vol. 55, no. 4, pp. 743–763. <https://doi.org/10.12989/sem.2015.55.4.743>
13. Bouderra B., Houari M.S.A., Tounsi A., Mahmoud S.R. Thermal stability of functionally graded sandwich plates using a simple shear deformation theory. *Structural Engineering and Mechanics*, 2016, vol. 58, no. 3, pp. 397–422. <https://doi.org/10.12989/sem.2016.58.3.397>
14. Chikh A., Bakora A., Heireche H., Houari M.S.A., Tounsi A., Bedia E.A.A. Thermo-mechanical postbuckling of symmetric S-FGM plates resting on Pasternak elastic foundations using hyperbolic shear deformation theory. *Structural Engineering and Mechanics*, 2016, vol. 57, no. 4, pp. 617–639. <https://doi.org/10.12989/sem.2016.57.4.617>
15. Yahiaoui M., Tounsi A., Fahsi B., Bouiadra R.B., Benyoucef S. The role of micromechanical models in the mechanical response of elastic foundation FG sandwich thick beams. *Structural Engineering and Mechanics*, 2018, vol. 68, no. 1, pp. 53–66. <https://doi.org/10.12989/sem.2018.68.1.053>
16. Achouri F., Benyoucef S., Bourada F., Bouiadra R. B., Tounsi A. Robust quasi 3D computational model for mechanical response of FG thick sandwich plate. *Structural Engineering and Mechanics*, 2019, vol. 70, no. 5, pp. 571–589. <https://doi.org/10.12989/sem.2019.70.5.571>
17. Barati M.R. Vibration analysis of porous FG nanoshells with even and uneven porosity distributions using nonlocal strain gradient elasticity. *Acta Mechanica*, 2018, vol. 229, no. 3, pp. 1183–1196. <https://doi.org/10.1007/s00707-017-2032-z>
18. She G.L., Yuan F.G., Ren Y.R., Liu H.B., Xiao W.S. Nonlinear bending and vibration analysis of functionally graded porous tubes via a nonlocal strain gradient theory. *Composite Structures*, 2018, vol. 203, pp. 614–623. <https://doi.org/10.1016/j.compstruct.2018.07.063>
19. Şimşek, M. Dynamic analysis of an embedded microbeam carrying a moving microparticle based on the modified couple stress theory. *International Journal of Engineering Science*, 2010, vol. 48, no. 12, pp. 1721–1732. <https://doi.org/10.1016/j.ijengsci.2010.09.027>
20. Abouelregal A.E., Zenkour A.M. Dynamic response of a nanobeam induced by ramp-type heating and subjected to a moving load. *Materials Research Express*, 2017, V. 4, N 8, P. 085013. <https://doi.org/10.1088/2053-1591/aa7d89>
4. Li L., Hu Y., Ling L. Flexural wave propagation in small-scaled functionally graded beams via a nonlocal strain gradient theory // *Composite Structures*. 2015. V. 133. P. 1079–1092. <https://doi.org/10.1016/j.compstruct.2015.08.014>
5. Zhang B., He Y., Liu D., Shen L., Lei J. Free vibration analysis of four-unknown shear deformable functionally graded cylindrical microshells based on the strain gradient elasticity theory // *Composite Structures*. 2015. V. 119. P. 578–597. <https://doi.org/10.1016/j.compstruct.2014.09.032>
6. Arefi M., Zenkour A.M. Free vibration, wave propagation and tension analyses of a sandwich micro/nano rod subjected to electric potential using strain gradient theory // *Materials Research Express*. 2016. V. 3. N 11. P. 115704. <https://doi.org/10.1088/2053-1591/3/11/115704>
7. Lou J., He L., Wu H., Du J. Pre-buckling and buckling analyses of functionally graded microshells under axial and radial loads based on the modified couple stress theory // *Composite Structures*. 2016. V. 142. P. 226–237. <https://doi.org/10.1016/j.compstruct.2016.01.083>
8. Zeighampour H., Shojaeian M. Buckling analysis of functionally graded sandwich cylindrical micro/nanoshells based on the couple stress theory // *Journal of Sandwich Structures & Materials*. 2019. V. 21. N 3. P. 917–937. <https://doi.org/10.1177/1099636217703912>
9. Farokhi H., Ghayesh M.H. Nonlinear mechanical behaviour of microshells // *International Journal of Engineering Science*. 2018. V. 127. P. 127–144. <https://doi.org/10.1016/j.ijengsci.2018.02.009>
10. Martínez-Criado G. Application of micro-and nanobeams for materials science // *Synchrotron light sources and free-electron lasers: accelerator physics, instrumentation and science applications*. 2016. P. 1505–1539. https://doi.org/10.1007/978-3-319-14394-1_46
11. Berrabah H.M., Tounsi A., Semmah A., Bedia E.A.A. Comparison of various refined nonlocal beam theories for bending, vibration and buckling analysis of nanobeams // *Structural Engineering and Mechanics*. 2013. V. 48. N 3. P. 351–365. <https://doi.org/10.12989/sem.2013.48.3.351>
12. Aissani K., Bouiadra M.B., Ahouel M., Tounsi A. A new nonlocal hyperbolic shear deformation theory for nanobeams embedded in an elastic medium // *Structural Engineering and Mechanics*. 2015. V. 55. N 4. P. 743–763. <https://doi.org/10.12989/sem.2015.55.4.743>
13. Bouderra B., Houari M.S.A., Tounsi A., Mahmoud S.R. Thermal stability of functionally graded sandwich plates using a simple shear deformation theory // *Structural Engineering and Mechanics*. 2016. V. 58. N 3. P. 397–422. <https://doi.org/10.12989/sem.2016.58.3.397>
14. Chikh A., Bakora A., Heireche H., Houari M.S.A., Tounsi A., Bedia E.A.A. Thermo-mechanical postbuckling of symmetric S-FGM plates resting on Pasternak elastic foundations using hyperbolic shear deformation theory // *Structural Engineering and Mechanics*. 2016. V. 57. N 4. P. 617–639. <https://doi.org/10.12989/sem.2016.57.4.617>
15. Yahiaoui M., Tounsi A., Fahsi B., Bouiadra R.B., Benyoucef S. The role of micromechanical models in the mechanical response of elastic foundation FG sandwich thick beams // *Structural Engineering and Mechanics*. 2018. V. 68. N 1. P. 53–66. <https://doi.org/10.12989/sem.2018.68.1.053>
16. Achouri F., Benyoucef S., Bourada F., Bouiadra R. B., Tounsi A. Robust quasi 3D computational model for mechanical response of FG thick sandwich plate // *Structural Engineering and Mechanics*. 2019. V. 70. N 5. P. 571–589. <https://doi.org/10.12989/sem.2019.70.5.571>
17. Barati M.R. Vibration analysis of porous FG nanoshells with even and uneven porosity distributions using nonlocal strain gradient elasticity // *Acta Mechanica*. 2018. V. 229. N 3. P. 1183–1196. <https://doi.org/10.1007/s00707-017-2032-z>
18. She G.L., Yuan F.G., Ren Y.R., Liu H.B., Xiao W.S. Nonlinear bending and vibration analysis of functionally graded porous tubes via a nonlocal strain gradient theory // *Composite Structures*. 2018. V. 203. P. 614–623. <https://doi.org/10.1016/j.compstruct.2018.07.063>
19. Şimşek, M. Dynamic analysis of an embedded microbeam carrying a moving microparticle based on the modified couple stress theory // *International Journal of Engineering Science*. 2010. V. 48. N 12. P. 1721–1732. <https://doi.org/10.1016/j.ijengsci.2010.09.027>
20. Abouelregal A.E., Zenkour A.M. Dynamic response of a nanobeam induced by ramp-type heating and subjected to a moving load // *Microsystem Technologies*. 2017. V. 23. N 12. P. 5911–5920. <https://doi.org/10.1007/s00542-017-3365-1>
21. Shahsavari D., Karami B., Janghorban M., Li L. Dynamic characteristics of viscoelastic nanoplates under moving load embedded within visco-Pasternak substrate and hygrothermal environment // *Materials Research Express*. 2017. V. 4. N 8. P. 085013. <https://doi.org/10.1088/2053-1591/aa7d89>

- Microsystem Technologies*, 2017, vol. 23, no. 12, pp. 5911–5920. <https://doi.org/10.1007/s00542-017-3365-1>
21. Shahsavari D., Karami B., Janghorban M., Li L. Dynamic characteristics of viscoelastic nanoplates under moving load embedded within visco-Pasternak substrate and hygrothermal environment. *Materials Research Express*, 2017, vol. 4, no. 8, pp. 085013. <https://doi.org/10.1088/2053-1591/aa7d89>
 22. Zhang Q., Liu H. On the dynamic response of porous functionally graded microbeam under moving load. *International Journal of Engineering Science*, 2020, vol. 153, pp. 103317. <https://doi.org/10.1016/j.ijengsci.2020.103317>
 23. Liu H., Zhang Q., Ma J. Thermo-mechanical dynamics of two-dimensional FG microbeam subjected to a moving harmonic load. *Acta Astronautica*, 2021, vol. 178, pp. 681–692. <https://doi.org/10.1016/j.actaastro.2020.09.045>
 24. Menasria A., Kaci A., Bousahla A.A., Bourada F., Tounsi A., Benrahou K.H., Tounsi, Bedia E.A.A., Mahmoud S.R. A four-unknown refined plate theory for dynamic analysis of FG-sandwich plates under various boundary conditions. *Steel and Composite Structures*, 2020, vol. 36, no. 3, pp. 355–367. <http://dx.doi.org/10.12989/scs.2020.36.3.355>
 25. Hachemi H., Bousahla A.A., Kaci A., Bourada F., Tounsi A., Benrahou K.H., Tounsi A., Al-Zahrani M.M., Mahmoud S.R. Bending analysis of functionally graded plates using a new refined quasi-3D shear deformation theory and the concept of the neutral surface position. *Steel and Composite Structures*, 2021, vol. 39, no. 1, pp. 51–64. <http://dx.doi.org/10.12989/scs.2021.39.1.051>
 26. Heidari F., Taheri K., Sheybani M., Janghorban M., Tounsi A. On the mechanics of nanocomposites reinforced by wavy/defected/aggregated nanotubes. *Steel and Composite Structures*, 2021, vol. 38, no. 5, pp. 533–545. <http://dx.doi.org/10.12989/scs.2021.38.5.533>
 27. Khaniki H.B., Hosseini-Hashemi S. The size-dependent analysis of multilayered microbridge systems under a moving load/mass based on the modified couple stress theory. *The European Physical Journal Plus*, 2017, vol. 132, no. 5, <https://doi.org/10.1140/epjp/i2017-11466-0>
 28. Eringen A.C. On differential equations of nonlocal elasticity and solutions of screw dislocation and surface waves. *Journal of Applied Physics*, 1983, vol. 54, no. 9, pp. 4703–4710. <https://doi.org/10.1063/1.332803>
 29. Ebrahimi F., Barati M.R., Zenkour A.M. A new nonlocal elasticity theory with graded nonlocality for thermo-mechanical vibration of FG nanobeams via a nonlocal third-order shear deformation theory. *Mechanics of Advanced Materials and Structures*, 2018, vol. 25, no. 6, pp. 512–522. <https://doi.org/10.1080/15376494.2017.1285458>
 22. Zhang Q., Liu H. On the dynamic response of porous functionally graded microbeam under moving load // *International Journal of Engineering Science*. 2020. V. 153. P. 103317. <https://doi.org/10.1016/j.ijengsci.2020.103317>
 23. Liu H., Zhang Q., Ma J. Thermo-mechanical dynamics of two-dimensional FG microbeam subjected to a moving harmonic load // *Acta Astronautica*. 2021. V. 178. P. 681–692. <https://doi.org/10.1016/j.actaastro.2020.09.045>
 24. Menasria A., Kaci A., Bousahla A.A., Bourada F., Tounsi A., Benrahou K.H., Tounsi, Bedia E.A.A., Mahmoud S.R. A four-unknown refined plate theory for dynamic analysis of FG-sandwich plates under various boundary conditions // *Steel and Composite Structures*. 2020. V. 36. N 3. P. 355–367. <http://dx.doi.org/10.12989/scs.2020.36.3.355>
 25. Hachemi H., Bousahla A.A., Kaci A., Bourada F., Tounsi A., Benrahou K.H., Tounsi A., Al-Zahrani M.M., Mahmoud S.R. Bending analysis of functionally graded plates using a new refined quasi-3D shear deformation theory and the concept of the neutral surface position // *Steel and Composite Structures*. 2021. V. 39. N 1. P. 51–64. <http://dx.doi.org/10.12989/scs.2021.39.1.051>
 26. Heidari F., Taheri K., Sheybani M., Janghorban M., Tounsi A. On the mechanics of nanocomposites reinforced by wavy/defected/aggregated nanotubes // *Steel and Composite Structures*. 2021. V. 38. N 5. P. 533–545. <http://dx.doi.org/10.12989/scs.2021.38.5.533>
 27. Khaniki H.B., Hosseini-Hashemi S. The size-dependent analysis of multilayered microbridge systems under a moving load/mass based on the modified couple stress theory // *The European Physical Journal Plus*. 2017. V. 132. N 5. <https://doi.org/10.1140/epjp/i2017-11466-0>
 28. Eringen A.C. On differential equations of nonlocal elasticity and solutions of screw dislocation and surface waves // *Journal of Applied Physics*. 1983. V. 54. N 9. P. 4703–4710. <https://doi.org/10.1063/1.332803>
 29. Ebrahimi F., Barati M.R., Zenkour A.M. A new nonlocal elasticity theory with graded nonlocality for thermo-mechanical vibration of FG nanobeams via a nonlocal third-order shear deformation theory // *Mechanics of Advanced Materials and Structures*. 2018. V. 25. N 6. P. 512–522. <https://doi.org/10.1080/15376494.2017.1285458>

Authors

Ridha A. Ahmed — Magister, Associate Professor, Associate Professor, Engineering College of Al-Mustansiriyah University, Baghdad, 10001, Iraq, [sc 57203618225](https://orcid.org/0000-0002-6666-6564), <https://orcid.org/0000-0002-6666-6564>, ridhalwan@uomustansiriyah.edu.iq

Wael N. Abdullah — Magister, Lecturer, Engineering College of Al-Mustansiriyah University, Baghdad, 10001, Iraq, [sc 58185482000](https://orcid.org/0000-0001-9599-4279), <https://orcid.org/0000-0001-9599-4279>, wael.najm65@gmail.com

Nadhim M. Faleh — D.Sc., Full Professor, Engineering College of Al-Mustansiriyah University, Baghdad, 10001, Iraq, [sc 57201261694](https://orcid.org/0000-0003-0003-1643), <https://orcid.org/0000-0003-0003-1643>, dr.nadhim@uomustansiriyah.edu.iq

Mamoon A. Al-Jaafari — PhD, Lecturer, Engineering College of Al-Mustansiriyah University, Baghdad, 10001, Iraq, [sc 57218543226](https://orcid.org/0009-0005-9597-9825), <https://orcid.org/0009-0005-9597-9825>, dr.mamoonal-jaafare@uomustansiriyah.edu.iq

Авторы

Ахмед Ридха А. — магистр, доцент, доцент, Университет Аль-Мустансирия, Багдад, 10001, Ирак, [sc 57203618225](https://orcid.org/0000-0002-6666-6564), <https://orcid.org/0000-0002-6666-6564>, ridhalwan@uomustansiriyah.edu.iq

Абдулла Вазель Н. — магистр, преподаватель, Университет Аль-Мустансирия, Багдад, 10001, Ирак, [sc 58185482000](https://orcid.org/0000-0001-9599-4279), <https://orcid.org/0000-0001-9599-4279>, wael.najm65@gmail.com

Фалех Надим М. — Dr.Sc., профессор, профессор, Университет Аль-Мустансирия, Багдад, 10001, Ирак, [sc 57201261694](https://orcid.org/0000-0003-0003-1643), <https://orcid.org/0000-0003-0003-1643>, dr.nadhim@uomustansiriyah.edu.iq

Аль-Джафари Мамун Али — PhD, преподаватель, Университет Аль-Мустансирия, Багдад, 10001, Ирак, [sc 57218543226](https://orcid.org/0009-0005-9597-9825), <https://orcid.org/0009-0005-9597-9825>, dr.mamoonal-jaafare@uomustansiriyah.edu.iq

Received 17.10.2024

Approved after reviewing 01.05.2025

Accepted 22.05.2025

Статья поступила в редакцию 17.10.2024

Одобрена после рецензирования 01.05.2025

Принята к печати 22.05.2025



Работа доступна по лицензии
Creative Commons
«Attribution-NonCommercial»

doi: 10.17586/2226-1494-2025-25-3-508-519
УДК 621.3.049.7

Проектирование микроэлектромеханического логического элемента на основе гребенчатого резонатора

Александр Анатольевич Соловьев¹✉, Евгений Филиппович Певцов²,
Владимир Анатольевич Колчужин³

^{1,2} МИРЭА — Российский технологический университет, Москва, 119454, Российская Федерация

³ Qorvo Munich GmbH, Мюнхен, 81829, Германия

¹ solovev_aa@mirea.ru ✉, <https://orcid.org/0009-0008-3578-0963>

² pevtsov@mirea.ru, <https://orcid.org/0000-0001-6264-1231>

³ vladimir.kolchuzhin@ieee.org, <https://orcid.org/0000-0002-6009-1499>

Аннотация

Введение. КМОП-технология практически достигла физического предела уменьшения размеров транзисторов и обладает существенными эксплуатационными ограничениями при экстремальных температурах и ионизирующем излучении. В работе предлагается методика проектирования логических элементов на основе альтернативной технологии, использующей гребенчатые микроэлектромеханические резонаторы, работающие на бесконтактном принципе и перепрограммируемые в процессе эксплуатации. **Метод.** Предложена методика расчета геометрических параметров устройства с помощью аналитических выражений и с учетом технологических норм, необходимых для достижения заданных характеристик: собственной частоты колебаний резонатора (100 кГц) и коэффициента добротности (20) при атмосферном давлении. **Основные результаты.** Определены оптимальные геометрические параметры устройства, характеристики емкостных ячеек, влияющие на чувствительность устройства и добротность с учетом воздушного демпфирования. Точность расчетов достаточна для проектирования фотошаблонов без использования специализированного программного обеспечения. Разработана компактная модель логического микроэлектромеханического элемента, позволяющая проводить системный анализ динамических характеристик и реализовывать функционально полный набор логических операций. **Обсуждение.** Разработанный маршрут проектирования может применяться для создания логических микроэлектромеханических элементов с возможностью перепрограммирования в процессе работы и дальнейшего каскадирования таких устройств для построения сложных цифровых схем. Работа полезна разработчикам микроэлектромеханических акселерометров и гироскопов и предлагает альтернативный подход к созданию трехмерных моделей на основе библиотеки параметрических компонентов и генерации компактных моделей для системного анализа.

Ключевые слова

МЭМС, микроэлектромеханические резонаторы, логические элементы, вентили, системное моделирование, ИЛИ-НЕ, Исключающее ИЛИ

Благодарности

Работа выполнена при поддержке Министерства науки и высшего образования Российской Федерации (государственное задание для университетов № ФГФЗ-2023-0005) и с применением оборудования Центра коллективного пользования РТУ МИРЭА (соглашение от 01.09.2021 № 075–15-2021-689, уникальный идентификационный номер 2296.61321X0010).

Ссылка для цитирования: Соловьев А.А., Певцов Е.Ф., Колчужин В.А. Проектирование микроэлектромеханического логического элемента на основе гребенчатого резонатора // Научно-технический вестник информационных технологий, механики и оптики. 2025. Т. 25, № 3. С. 508–519. doi: 10.17586/2226-1494-2025-25-3-508-519

Design of the microelectromechanical logic element based on a comb-drive resonator

Alexander A. Solov'ev¹, Evgeny F. Pevtsov², Vladimir A. Kolchuzhin³

^{1,2} Russian Technology University MIREA, Moscow, 119454, Russian Federation

³ Qorvo Munich GmbH, Munich, 81829, Germany

¹ solov'ev_aa@mirea.ru, <https://orcid.org/0009-0008-3578-0963>

² pevtsov@mirea.ru, <https://orcid.org/0000-0001-6264-1231>

³ vladimir.kolchuzhin@ieee.org, <https://orcid.org/0000-0002-6009-1499>

Abstract

CMOS technology has nearly reached the physical limits of transistor scaling and exhibits significant operational limitations at extreme temperatures and ionizing radiation. This work proposes a methodology for designing logic elements based on an alternative technology utilizing comb-drive microelectromechanical resonators operating on a non-contact principle and reconfigurable during operation. A method is proposed for calculating the geometric parameters of the device using analytical expressions and considering technological norms necessary to achieve specified characteristics: the natural frequency of resonator oscillations (100 kHz) and the quality factor (20) at atmospheric pressure. Optimal geometric parameters of the device, characteristics of capacitive cells affecting the sensitivity of the device and the quality factor, taking into account air damping, are determined. The accuracy of the calculations is sufficient for designing photomasks without using specialized software. A compact model of a logic microelectromechanical element has been developed, allowing for system analysis of dynamic characteristics and implementation of a functionally complete set of logic operations. The developed design flow can be applied to create logic microelectromechanical elements with the possibility of reprogramming during operation and further cascading of such devices for constructing complex digital circuits. The article is useful for developers of microelectromechanical accelerometers and gyroscopes and proposes an alternative approach to creating three-dimensional models based on a library of parametric components and generating compact models for system analysis.

Keywords

MEMS, micro-electro-mechanical systems resonators, logic gates, system-level modeling, NOR gate, XOR gate

Acknowledgements

This work was supported by the Ministry of Science and Higher Education of the Russian Federation (State Assignment for Universities No. ФГФ3-2023-0005) and carried out using equipment of the Shared Equipment Center of RTU MIREA (Agreement No. 075-15-2021-689 dated 01.09.2021, unique identification number 2296.61321X0010).

For citation: Solov'ev A.A., Pevtsov E.F., Kolchuzhin V.A. Design of the microelectromechanical logic element based on a comb-drive resonator. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2025, vol. 25, no. 3, pp. 508–519 (in Russian). doi: 10.17586/2226-1494-2025-25-3-508-519

Введение

Элементная база современных вычислительных устройств основана на КМОП-технологии, которая практически достигла физического предела уменьшения размеров транзисторов и обладает существенными эксплуатационными ограничениями при экстремальных температурах и ионизирующем излучении. В связи с этим возрастает интерес к альтернативным подходам, среди которых перспективным направлением являются вычисления на основе микроэлектромеханических систем (МЭМС).

Особое внимание представляют МЭМС-резонаторы, используемые в датчиках, приводах, переключателях, фильтрах и логических элементах [1]. Резонансные структуры применяются в механических логических вычислениях из-за устойчивости к вибрациям, возможности перепрограммирования в процессе эксплуатации [2] и бесконтактному принципу работы, в отличие от гальванических переключателей, подверженных деградации контактных площадок [3].

В работах [4, 5] представлены конструкции логических МЭМС-элементов на основе резонаторов, представляющие балку, закрепленную с двух сторон, два и более боковых неподвижных электрода, которые являются управляющими электродами для приведения балки в действие, цифровыми входами и выходами логического вентиля. Однако недостатками данных

устройств являются малые управляющие емкости резонатора, влияющие на чувствительность устройства, а добротность равная 100 при давлении в 800 Па делает невозможным функционирование при атмосферном давлении без дорогостоящего герметичного корпуса.

В настоящей работе рассмотрен процесс проектирования логического МЭМС-элемента на основе гребенчатого микромеханического резонатора с поперечным перемещением электродов с повышенной чувствительностью, работой при атмосферном давлении с заданным демпфированием. В качестве исходных параметров системы использованы собственная частота колебаний резонатора 100 кГц и коэффициент добротности 20 при атмосферном давлении, определенный в ходе исследования [6], обеспечивающий необходимое разделение пиков амплитудно-частотной характеристики (АЧХ) для выполнения логических операций «ИЛИ-НЕ», «Исключающее ИЛИ», «И» и минимальное время перехода между логическими состояниями.

На этапе предварительного проектирования рассчитываются аналитические выражения для определения геометрических параметров устройства с учетом технологических норм, необходимых для достижения заданных характеристик. Аналитические выражения позволяют определить характеристики емкостных ячеек, влияющие на чувствительность устройства, а также добротность с учетом воздушного демпфирования с точностью, достаточной для проектирования фотоша-

блонов, без необходимости использования специализированного программного обеспечения.

Для создания трехмерной модели применен не традиционный метод конечных элементов (МКЭ), обеспечивающий высокую точность, но требующий значительных вычислительных ресурсов, а подход на основе библиотеки параметрических трехмерных компонентов, что сокращает время моделирования при сохранении заданной точности [7].

Результатом процесса проектирования является компактная модель логического МЭМС-элемента для системного анализа динамических характеристик и реализации функционально полного набора логических операций, созданная из трехмерной параметрической модели.

Расчет геометрических размеров МЭМС-резонатора

В работе рассматривается гребенчатый резонатор с поперечным перемещением электродов. В отличие от МЭМС-резонаторов с продольным перемещением электродов, где инерционная масса с электродами смещается вдоль оси гребенчатых структур, в данном устройстве используется поперечное перемещение относительно неподвижных электродов, что обеспечивает высокую чувствительность, поскольку даже при малых перемещениях значительно изменяется емкость.

Разработка конструкции логического МЭМС-резонатора выполнена с учетом технологической реализуемости. Ограничения процесса глубокого ионного травления определяют минимальные размеры элементов, что необходимо учитывать при проектировании. Толщина структуры составляет 50 мкм, минимальная ширина элементов — 3 мкм, минимальный зазор между электродами — 2 мкм.

Определение геометрических параметров инерционной массы является важным этапом проектирования, поскольку ее площадь и форма оказывают прямое влияние на динамические характеристики МЭМС-устройства, включая резонансную частоту и добротность. Размеры МЭМС-устройства ограничены доступной площадью на пластине, например, в случае использования мультипроектной технологии Multi-Project Wafer (MPW). Необходимо учитывать не только размер инерционной массы и гребенчатых электродов, но и подводящих электродов, контактных площадок для соединения с корпусом. При начальном проектировании выбрана прямоугольная форма инерционной массы.

Гребенчатые электроды, размещенные на подвижной инерционной массе и неподвижных боковых электродах, образуют емкостные ячейки. Оптимальное соотношение воздушных зазоров между электродами емкостной ячейки составляет 2,5 при ширине электрода 3 мкм и минимальном зазоре между электродами — 2 мкм [8]. Масса одного электрода гребенчатой структуры, размещенного на инерционной массе, определяется выражением:

$$m_{\text{электрода}} = l_{\text{электрода}} w_{\text{электрода}} \rho h,$$

где $l_{\text{электрода}}$ — длина электрода; $w_{\text{электрода}}$ — ширина электрода; ρ — плотность материала; h — толщина конструкции, определяемая технологическим процессом.

Ширина одной элементарной емкостной ячейки ($w_{\text{ячейки}}$) (рис. 1) включает ширину подвижного и фиксированного электродов и воздушные зазоры между электродами, т. е.

$$w_{\text{ячейки}} = g_1 + 2w_{\text{электрода}} + g_2,$$

где g_1 — зазор между левым электродом статора и электродом ротора; g_2 — зазор между электродом ротора и правым электродом статора.

Количество электродов ($n_{\text{ячеек}}$), размещенных на инерционной массе, рассчитывается с учетом расстояния между гребенками и области крепления подвесов:

$$n_{\text{ячеек}} = \left(\frac{a - d_{\text{гребенок}}}{2w_{\text{ячейки}}} - \frac{w_{\text{подвеса}} - g_1}{w_{\text{ячейки}}} \right) n_{\text{гребенок}},$$

где a — длина центральной части инерционной массы; $d_{\text{гребенок}}$ — расстояние между боковыми электродами; $w_{\text{подвеса}}$ — ширина подвесов; $n_{\text{гребенок}}$ — количество боковых электродов. Масса всех подвижных электродов определяется суммированием масс отдельных электродов $m_{\text{гребенок}} = m_{\text{электрода}} n_{\text{ячеек}}$. Общая масса подвижного электрода определяется как сумма масс центральной части ($m_{\text{инерц}}$) и электродов гребенчатой структуры, размещенных на инерционной массе $m = m_{\text{инерц}} + m_{\text{гребенок}}$.

Для определения зависимости электрической емкости от величины межэлектродного зазора применяется выражение для плоского конденсатора [9]:

$$C_i = \frac{\epsilon_0 \epsilon l_{\text{перекр}} h}{g_i}, \quad (1)$$

где ϵ_0 — электрическая постоянная; ϵ — диэлектрическая проницаемость среды между электродами (в настоящей работе принимаемая за 1); $l_{\text{перекр}}$ — длина перекрытия электродов емкостной ячейки; g_i — межэлектродный зазор, где $i = 1$ для левой, $i = 2$ для правой части емкостной ячейки.

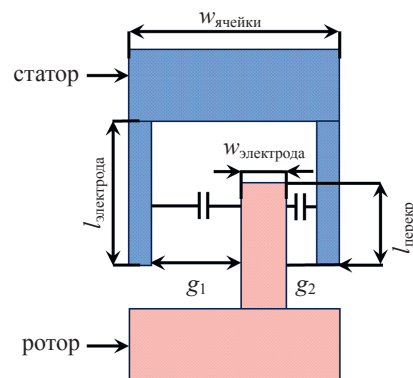


Рис. 1. Схема емкостной ячейки.

$l_{\text{перекр}}$ — длина перекрытия электродов

Fig. 1. Capacitor cell schematic.

$l_{\text{перекр}}$ is the length of the electrode overlap

Формула (1) учитывает только однородное электрическое поле между электродами емкостной ячейки. Однако в реальных конструкциях влияние краевых эффектов может увеличивать емкость до 45 % по сравнению с расчетными значениями [10]. В [11] предложена зависимость емкости левой и правой частей ячейки от краевых эффектов, предполагающая линейное распределение электрического поля в неперекрывающейся области.

$$C_{fi} = \frac{\varepsilon_0 f h}{g_{срi}}, g_{срi} = 0,92 \frac{g_i + \sqrt{g_i^2 + f^2}}{2},$$

где $g_{срi}$ — среднее расстояние между электродами; $i = 1$ для g_1 , $i = 2$ для g_2 ; $f = l_{\text{электрода}} - l_{\text{перекр}}$ — длина неперекрывающейся части электрода; 0,92 — коэффициент, установленный экспериментально, применимый при $f/g \leq 10$.

Емкость фронтальной части ($C_{\text{фронт}}$) ячейки определяется аналогично боковой, но без учета смещения [11]:

$$C_{\text{фронт}} = \varepsilon_0 h \frac{w_{\text{электрода}}}{l_{\text{электрода}} - l_{\text{перекр}}}.$$

Суммарная емкость устройства определяется суммой емкостей правой и левой частей ячейки, краевого поля и фронтальной части, скорректированной на количество ячеек [11]:

$$C = (C_1 + C_2 + C_{f1} + C_{f2} + C_{\text{фронт}}) n_{\text{ячеек}}.$$

Чувствительность емкостной ячейки относительно смещения x определяется как производная емкости по перемещению [12]:

$$\frac{\partial C_i}{\partial x} = \pm \varepsilon_0 l_{\text{перекр}} h \frac{1}{(g_i \mp x)^2}.$$

Общая чувствительность устройства формируется как сумма производных емкостей левой и правой части, скорректированное на число ячеек: $S = \frac{\partial C}{\partial x} = \left(\frac{\partial C_1}{\partial x} + \frac{\partial C_2}{\partial x} \right) n_{\text{ячеек}}$. Конструкция логического МЭМС-элемента на основе резонатора представляет собой инерционную массу, подвешенную на четырех упругих подвесах. Собственная частота колебательной системы f_0 определяется выражением [9]:

$$f_0 = \frac{1}{2\pi} \sqrt{\frac{k_{\text{эфф}}}{m}}, \quad (2)$$

где m — инерционная масса; $k_{\text{эфф}}$ — эффективная жесткость подвесов. Таким образом жесткость системы составляет $k_{\text{эфф}} = 4\pi^2 f_0^2 m$. Для расчета жесткости системы учитываем, что конструкция содержит четыре одинаковых упругих подвеса $k_{\text{эфф}} = 4k_{\text{подвеса}}$.

Жесткость одного упругого подвеса определяется с помощью уравнения прогиба консольной балки под действием силы [9]:

$$k_{\text{подвеса}} = \frac{12E_{\text{Si}}I}{l_{\text{подвеса}}^3},$$

где E_{Si} — модуль упругости кремния; $I = \frac{w_{\text{подвеса}}^3 h}{12}$ — момент инерции для прямоугольного сечения балки, $w_{\text{подвеса}}$ — ширина балки.

Механизм диссипации механической энергии (демпфирование) определяет добротность резонансной конструкции. В настоящей работе рассматривается только воздушное демпфирование, в частности демпфирование сжатой газовой пленки и демпфирование скольжения. Схематическое изображение воздушного демпфирования между электродами гребенки представлено на рис. 2.

Согласно [13], демпфирование, вызванное сжатием воздуха в емкостных ячейках резонатора, зависит от вязкости газа, геометрии канала и количества каналов. Число Кнудсена (Kn) вычисляется по формуле

$$Kn = \frac{\lambda}{g_i},$$

где g_i — воздушные зазоры (g_1 и g_2) между электродами; $\lambda = 66$ нм — средняя длина свободного пробега молекул воздуха при атмосферном давлении.

Эффективная вязкость, учитывающая влияние числа Кнудсена определяется:

$$\eta_{\text{эфф}} = \frac{\eta}{1 + 9,638Kn^{1,159}},$$

где η — вязкость воздуха.

В работе демпфирование сжатия воздуха для случая $l_{\text{перекр}} < h$ для левой и правой сторон ячейки определяется выражением [13]:

$$d_{\text{сжатия}_i} = \left(1 - \frac{0,63l_{\text{перекр}}}{h} + \frac{0,052l_{\text{перекр}}^5}{h^5} \right) \frac{l_{\text{перекр}}^3 h}{g_i^3} \eta_{\text{эфф}} n,$$

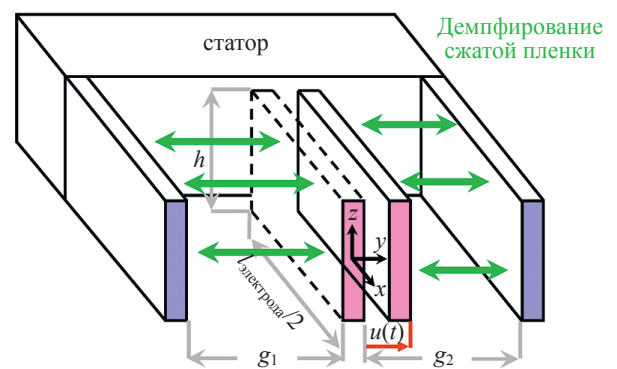


Рис. 2. Схематическое изображение воздушного демпфирования между электродами гребенки. g_1 и g_2 — левый и правый зазоры между электродами статора и ротора; $u(t)$ — перемещение подвижного электрода относительно начального положения в зависимости от времени

Fig. 2. Schematic diagram showing air damping between the comb-drive electrodes.

g_1 and g_2 are the left and right gaps between the stator and rotor electrodes; $u(t)$ — displacement of the movable electrode relative to its initial position as a function of time

где h — толщина структуры; n — количество каналов ($n = 2n_{\text{ячеек}}$).

Общее демпфирование сжатия для всех ячеек резонатора имеет вид

$$D_{\text{сжатия}} = n_{\text{ячеек}} n_{\text{гребенок}} (D_{\text{сжатия1}} + D_{\text{сжатия2}}),$$

где $D_{\text{сжатия1}}$ и $D_{\text{сжатия2}}$ — демпфирование в зазорах g_1 и g_2 емкостной ячейки.

Для вычисления демпфирования от скольжения элементов необходимо учитывать демпфирование от скольжения электродов гребенки сверху и снизу, а затем сложить значения для всех элементов. Длина скользящей части гребенки для расчета демпфирования одинакова сверху и снизу. Для двух гребенок (сверху и снизу) применяется формула:

$$l_{\text{гребенок}} = 2w_{\text{ячейки}} n_{\text{ячеек}}.$$

Если ширина скользящей части не равна 0, то зазор (g) между поверхностями с учетом ширины канала определяется по формуле [13]

$$g = \frac{h}{1 + 8,5 \frac{h}{l_{\text{гребенок}}}},$$

в противном случае $g = h$.

Эффективная вязкость воздуха ($\eta_{\text{эфф}}$) описывается упрощенной формулой [13]:

$$\eta_{\text{эфф}} = \frac{n}{1 + 2Kn}.$$

Число Кнудсена (Kn) определяется как отношение длины свободного пробега молекул газа (λ) к толщине структуры (h): $Kn = \frac{\lambda}{h}$. Демпфирование скольжения зависит от расстояния между поверхностями и определяется формулой Куэтта [13]:

$$D_{\text{скольжения}} = n l_{\text{электрода}} w_{\text{электрода}} \frac{\eta_{\text{эфф}}}{g},$$

где $n = 2n_{\text{гребенок}} n_{\text{ячеек}}$ — количество каналов для электродов гребенки снизу.

Общее демпфирование определяется как сумма демпфирования газовой пленки и демпфирования скольжения $D = D_{\text{сжатия}} + D_{\text{скольжения}}$. Добротность (Q) задается изменением начального перекрытия электродами емкостной ячейки и имеет вид [9]

$$Q = \frac{\sqrt{mk_{\text{эфф}}}}{D}, \quad (3)$$

где D — общее демпфирование.

Параметрические исследования

Разработка МЭМС-устройств требует мультифизического подхода для учета взаимосвязи параметров. После предварительного расчета следует этап параметрического исследования зависимости размеров

элементов конструкции на ключевые характеристики устройства: рабочую частоту, напряжение схлопывания и коэффициент добротности.

Для автоматизации параметрических исследований разработана программа для персонального компьютера на языке MATLAB [14]. Программа может быть интегрирована с другим специализированным программным обеспечением для уточнения механического поведения устройства после первичного анализа.

Высокая резонансная частота необходима для повышения быстродействия логического элемента. Уменьшение массы подвижного электрода m увеличивает рабочую частоту f_0 , при этом важно сохранить достаточную длину инерционной массы для размещения максимального количества емкостных ячеек. Для уменьшения массы также применяется перфорация — создание сквозных отверстий в инерционной массе.

Оптимизация добротности является ключевым этапом проектирования логического МЭМС-элемента, поскольку эти параметры влияют на скорость переключения, энергопотребление и возможность реализации логических операций при атмосферном давлении.

Высокая добротность минимизирует потери энергии и увеличивает амплитуду колебаний, но также увеличивает время переключения из «1» в «0». Снижение добротности уменьшает время переключения, но требует увеличения напряжения для достаточного разделения пиков АЧХ.

Для гребенчатого резонатора с поперечным перемещением добротность регулируется изменением перекрытия электродов емкостной ячейки. Увеличение перекрытия повышает газовое демпфирование, что снижает добротность и позволяет достичь целевого значения 20, установленного в ходе анализа [6].

Для расчетов в качестве подвесов использовались консольные балки, закрепленные одним концом на углах инерционной массы, а другим — на анкерах. В МЭМС, таких как акселерометры и гироскопы, часто применяются П-образные подвесы, состоящие из двух балок и соединяющей перемычки в верхней части (рис. 3). Жесткость П-образного подвеса определяется жесткостью балок и рассчитывается как последовательное соединение этих элементов $\frac{1}{k_{\text{общ}}} = \frac{1}{k_{\text{вертик1}}} + \frac{1}{k_{\text{вертик2}}}$. П-образные подвесы компактнее других конструкций и обеспечивают движение вдоль выбранной оси, так как жесткость в других направлениях значительно выше, что предотвращает нежелательные колебания.

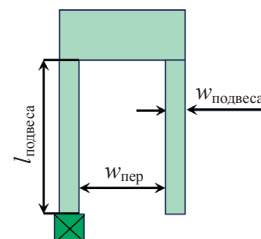


Рис. 3. Конструкция П-образного подвеса. $w_{\text{пер}}$ — ширина перемычки

Fig. 3. Design of the n-shaped suspension. $w_{\text{пер}}$ — width of the bridge

Напряжение схлопывания — напряжение, при котором подвижный электрод притягивается к неподвижному под действием электростатической силы, что приводит к схлопыванию электродов. Для его расчета между парой электродов используется выражение [9]:

$$V_{\text{схл}} = \sqrt{\frac{8k_{\text{эфф}}g^3}{27\varepsilon A}}, \quad (4)$$

где g — начальный зазор между электродами; $A = l_{\text{перекр}}h$ — площадь перекрытия электродов.

Для снижения управляющего напряжения логического элемента необходимо уменьшать напряжение схлопывания. Для этого можно уменьшить жесткость подвесов ($k_{\text{эфф}}$) увеличить площадь перекрытия электродов (A) или уменьшить начальный зазор (g). Поскольку зазор фиксирован на предыдущем шаге, основной метод уменьшать напряжение схлопывания — снижение жесткости подвесов за счет их удлинения. На рис. 4 показана зависимость резонансной частоты (f_0), рассчитываемая с помощью выражения (2), напряжения схлопывания ($V_{\text{схл}}$) (4), и добротности (Q), рассчитан-

ной с помощью выражения (3) от длины подвеса и перекрытия электродов емкостной ячейки.

Таким образом, для достижения ключевых характеристик устройства (собственной частоты колебаний — 100 кГц и коэффициента добротности — 20) рассчитаны длина подвеса ($l_{\text{подвеса}} = 165$ мкм) и перекрытие электродов емкостной ячейки ($l_{\text{перекр}} = 22$ мкм).

В таблице представлены геометрические параметры и характеристики логического МЭМС-элемента на основе гребенчатого резонатора.

Трехмерная модель логического МЭМС-элемента

Анализ МЭМС-устройства требует учета взаимодействия электростатики и механики, а также механизмов диссипации механической энергии. Традиционные сеточные методы, такие как МКЭ, обеспечивают высокую точность, но требуют значительных вычислительных ресурсов. В данной работе использован подход на основе библиотеки параметрических трехмерных компонентов, что сокращает время моделирования при сохранении заданной точности [7].

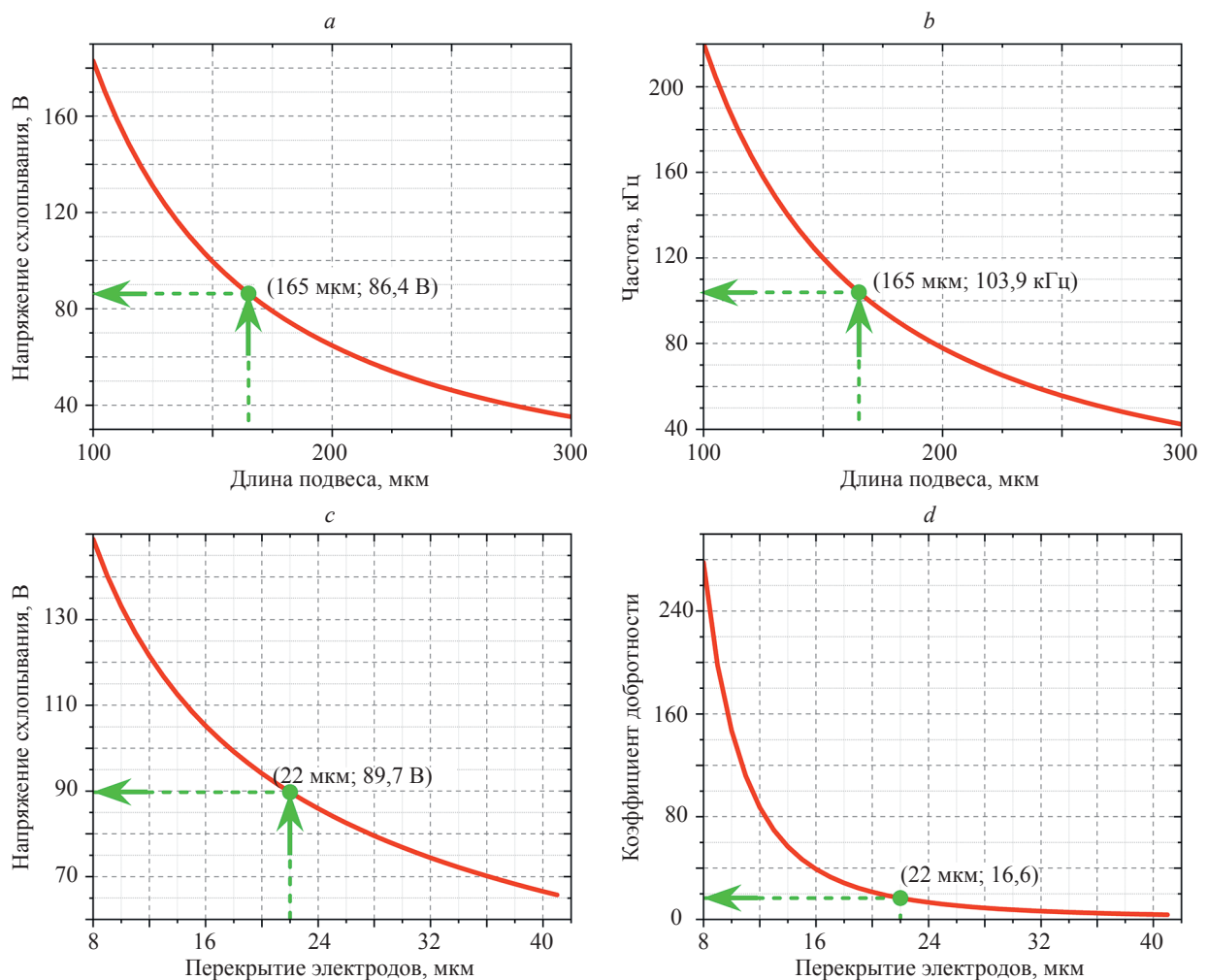


Рис. 4. Параметрический анализ. Зависимости напряжения схлопывания (a) и частоты (b) от длины подвеса; коэффициента напряжения схлопывания (c) и добротности (d) от перекрытия электродов емкостной ячейки

Fig. 4. Parametric analysis. Dependence of pull-in voltage (a) and frequency (b) vs. suspension length; Dependence of pull-in voltage coefficient (c) and quality factor (d) vs. the overlap of capacitive cell electrodes

Таблица. Параметры оптимизированной конструкции
Table. Optimized Design Parameters

Свойство	Значение	Свойство	Значение
Характеристики		Емкостная ячейка	
Частота, кГц	103	Длина электрода, мкм	50
Добротность	17	Ширина электрода, мкм	3
Инерционная масса		Перекрытие электродов, мкм	22
Длина, мкм	360	Зазор 1, мкм	2
Ширина, мкм	60	Зазор 2, мкм	5
Масса, мкг	1,12	Количество ячеек	13×4
Упругие подвесы		Емкость и чувствительность	
Длина, мкм	165	Емкость устройства, пФ	0,56
Ширина, мкм	5	Чувствительность, мкФ/м	0,147

Библиотека содержит базовые элементы МЭМС: балки, пластины, гребенчатые электроды, моделируемые аналитическими уравнениями как многополосники. Поведенческие модели элементов учитывают нелинейную физику [7], технологические нормы (углы боковых стенок, остаточные напряжения, перфорации) и физические эффекты (демпфирование, электростатика, механика) [15]. В отличие от МКЭ, количество степеней свободы снижено за счет решения дифференциальных уравнений только в узловых точках [16]. Параметризация позволяет изменять геометрические размеры (длину подвесов, площадь инерционной массы, перекрытие электродов), оптимизируя добротность и частотные характеристики устройства. Компоненты соединяются в узловых точках, формируя сложные конструкции без построения сетки [17]. Динамика системы описывается дифференциальными уравнениями с матрицами жесткости, массы и демпфирования [16, 18].

Разработанный логический МЭМС-элемент на основе гребенчатого резонатора (рис. 5) включает инерционную массу 5 — жесткой пластины с перфорацией, закрепленной с помощью четырех П-образных упругих подвесов 6, состоящих из 12 балочных элементов [19]

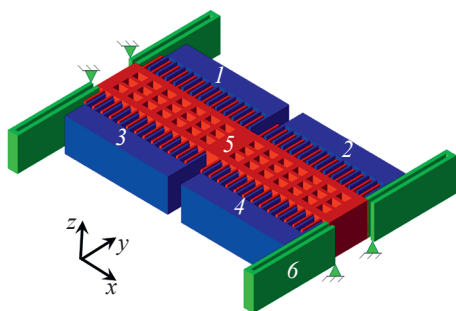


Рис. 5. Трехмерная модель микроэлектромеханического логического элемента.

1–4 — боковые гребенчатые электроды-статоры; 5 — несущая инерционная масса с гребенчатыми электродами-роторами; 6 — П-образные упругие подвесы

Fig. 5. 3D model of the microelectromechanical logic element.
1–4: side comb-drive stator electrodes, 5: supporting inertial mass with comb-drive rotor electrodes, 6: n-shaped spring suspensions

(по три балки на каждый подвес), моделируемых по теории Тимошенко с учетом нелинейных деформаций и обеспечивающих 6 степеней свободы, четырех емкостных гребенчатых структур 1–4, выполняющих функции электростатических приводов, рассчитываемые методом Шварца–Кристоффеля [19].

Модель обладает 6 механическими степенями свободы, электростатические силы рассчитываются в 76 точках интегрирования в гребенках и системой демпфирования, которая учитывает эффект сжимаемой газовой пленки, моделируемый на основе уравнений Рейнольдса с учетом периодичности гребенчатых структур. Данный метод обладает ограничением: из-за библиотечного представления геометрии невозможно добавить нестандартные элементы, например ограничители перемещения несущей массы.

Системная компактная модель

Проектирование МЭМС требует совместного моделирования в составе управляющей схемы для анализа переходных процессов [20]. Для этого используется структурно-поведенческая модель на языках описания аппаратуры Verilog-A, VHDL-AMS.

МКЭ создает модели с тысячами степеней свободы (Degree of Freedom, DoF), что значительно усложняет их применение для системного анализа из-за высокой вычислительной сложности. Модели сокращенного порядка позволяют описать динамическое поведение устройства, сохраняя при этом точность моделирования. Одним из методов редукции является метод модальной суперпозиции, сокращающий размерность модели при сохранении нелинейного электростатического взаимодействия [21].

В настоящей работе генерируется компактная модель из трехмерной модели на основе библиотеки параметрических трехмерных компонентов. Компактная модель содержит матрицы, интерполяционные параметры (например, интерполятор Лагранжа), аппроксимацию нелинейностей и информацию о выводах. Моделирование инициализируется в DC (Direct Current)- или AC (Alternating Current)-состоянии, а на каждом временном шаге восстанавливаются нелиней-

ные электростатические члены высокого порядка, что позволяет учитывать электростатическое смягчение без значительного усложнения вычислений. Число DoF сокращается с нескольких сотен до нескольких десятков параметров. Компактные модели точны только для заданной конфигурации параметров, что требует баланса между скоростью и точностью, учитывая ограниченное количество выбранных нелинейных эффектов. Созданная системная модель [22] экспортируется в среду проектирования интегральных схем как список соединений и схемный символ для совместного моделирования с МЭМС.

Тестирование модели

Целью тестирования является проверка корректности функционирования модели с помощью анализа статического, частотного и динамического откликов структуры, определения величины постоянного смещения инерционной массы и управляющего напряжения, при которых разница в резонансных частотах (Δf) при различных комбинациях на боковых электродах не менее 6 кГц при добротности (Q) около 20 и реализация на модели логических операций «ИЛИ-НЕ», «Исключающее ИЛИ», «И».

Эффект схлопывания возникает при превышении критического напряжения, при котором подвижный электрод притягивается к неподвижному под действием электростатической силы. На рис. 6, а показан анализ схлопывания для двух схем подключения электродов пятиконтактного логического МЭМС-элемента:

- электроды 1, 2 и 3 плавающие, электрод 4 имеет нулевой потенциал, на электрод 5 подается линейно изменяемое напряжение (рис. 5). Напряжение схлопывания составляет 84,5 В, что немного ниже расчетного значения 87,6 В, полученного по формуле (3);
- электроды 1 и 2 плавающие, электроды 3 и 4 имеют нулевой потенциал, на подвижный электрод 5 пода-

ется линейно изменяемое напряжение. Напряжение схлопывания снижается до 59,8 В из-за увеличения площади электростатического взаимодействия между подвижным и боковыми электродами.

- Напряжение логической «1» принято равным 60 В, а «0» — 0 В. Анализ смещения подвижного электрода в зависимости от напряжений на боковых электродах 2 и 3, показал следующие закономерности (рис. 6, б):
- при разности потенциалов «0–1» или «1–0» подвижный электрод притягивается к соответствующему боковому электроду под действием электростатической силы;
- в случае комбинаций «0–0» или «1–1» разность потенциалов между противоположными боковыми электродами компенсируется, что удерживает подвижный электрод в равновесии.

Гармонический анализ используется для исследования реакции устройства на внешнее гармоническое возбуждение. В основе работы логического МЭМС-элемента лежит изменение резонансной частоты подвижного электрода за счет эффекта электростатического смягчения жесткости системы. Изменение входных комбинаций напряжений на боковых электродах (цифровые входы 2 и 3) и на подвижном электроде 5 позволяет динамически управлять резонансной частотой МЭМС-элемента, что обеспечивает выполнение различных логических операций на одном устройстве.

На пятиконтактном логическом МЭМС-элементе три различных резонансных частоты реализуются при следующих комбинациях на боковых электродах 2 и 3 (рис. 7):

- комбинация «0–0»: резонансная частота $f_1 = 75,5$ кГц (операция «ИЛИ-НЕ»);
- комбинация «0–1» или «1–0»: частота $f_2 = 86,6$ кГц (операция «Исключающее ИЛИ»);
- комбинация «1–1»: резонансная частота $f_3 = 93,5$ кГц (операция «И»).

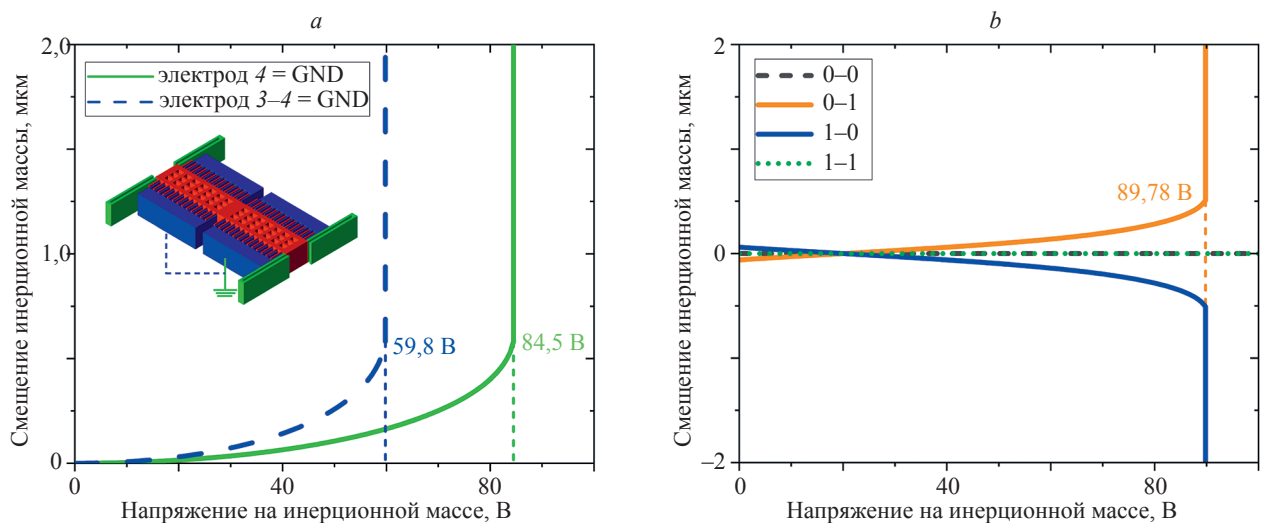


Рис. 6. Анализ эффекта схлопывания при смещении двух (а) и трех (б) электродов.

GND — земля (Ground)

Fig. 6. Pull-in effect analysis for the displacement of (a) two and (b) three electrodes.

GND — Ground

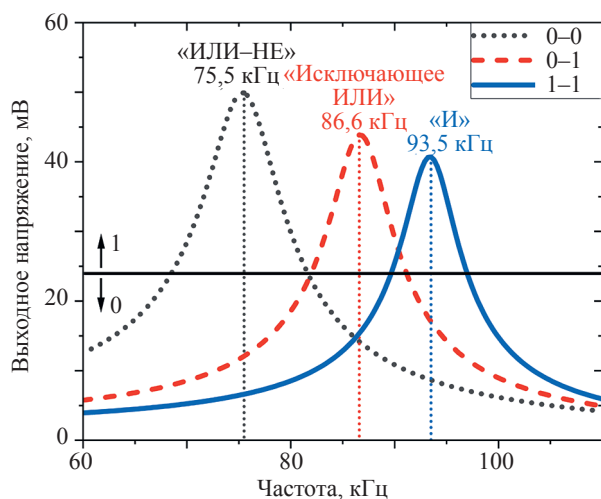


Рис. 7. Амплитудно-частотные характеристики резонатора при различных входных комбинациях на цифровых входах боковых электродов 2 и 3 при $Q = 20$ для напряжений электродов 4 и 5: $V_{\text{схл}4} = 1 \text{ В (AC)}$, $V_{\text{схл}5} = 60 \text{ В (DC)}$

Fig. 7. Frequency response of the resonator for various input combinations at digital inputs 2 and 3, $Q = 20$, $V_4 = 1 \text{ В (AC)}$, $V_5 = 60 \text{ В (DC)}$

На подвижный электрод 5 подается постоянное напряжение 60 В, так как в данном случае разница в частоте Δf между «0–0» и «0–1» максимальная и составляет не менее 6 кГц при $Q = 20$.

Реализация логических операций

Разработанный МЭМС логический элемент способен выполнять базовые логические операции такие, как «И», «ИЛИ-НЕ» и «Исключающее ИЛИ». Схема подключения МЭМС-элемента к внешним источникам представлена на рис. 8.

Для реализации логической операции «И» на электрод 4 подается синусоидальный сигнал с частотой f_3 . Резонанс возникает только при комбинации «1–1», когда частота сигнала возбуждения совпадает с резонансной частотой подвижного электрода. Это приводит к детектированию «высокого» выходного сигнала на чувствительном электроде 1. В остальных случаях выходной сигнал имеет низкую амплитуду. На рис. 9, с показана осциллограмма, подтверждающая реализацию логической операции «И». Аналогично, для реализации логической операции «ИЛИ-НЕ» на электрод 4 подается синусоидальный сигнал с частотой f_1 (рис. 9, а), а при возбуждении с частотой f_2 выполняется логическая операция «Исключающее ИЛИ» (рис. 9, b).

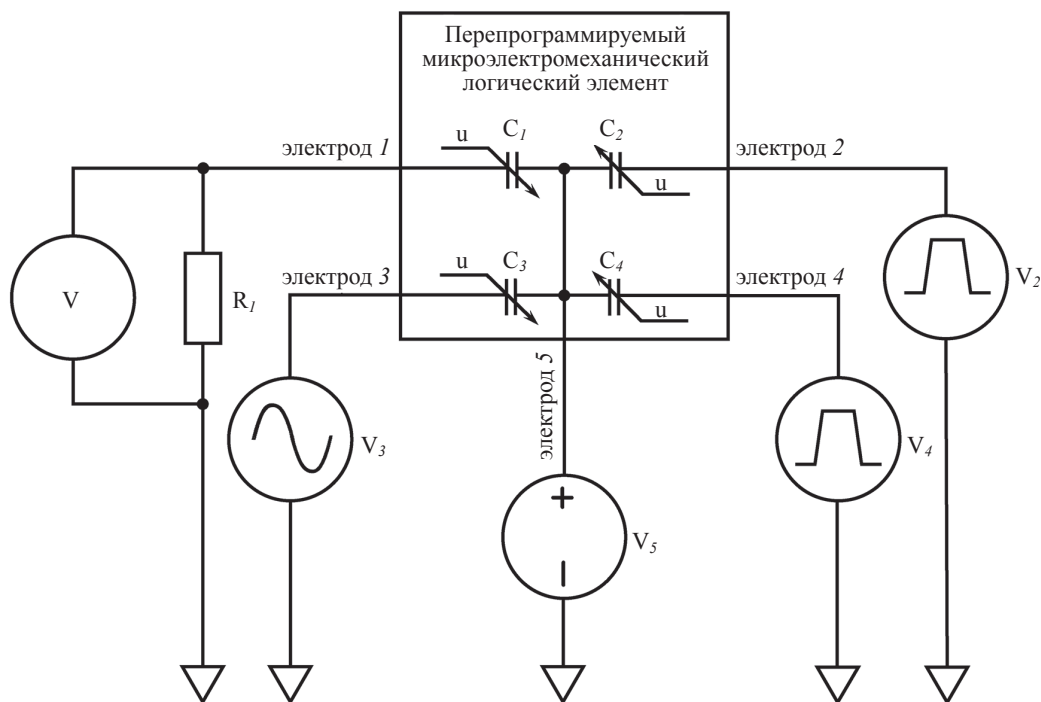


Рис. 8. Подключение микрорезонатора и внешних источников для выполнения логических операций

Fig. 8. Connection of the microresonator and external sources for implementing logic operations

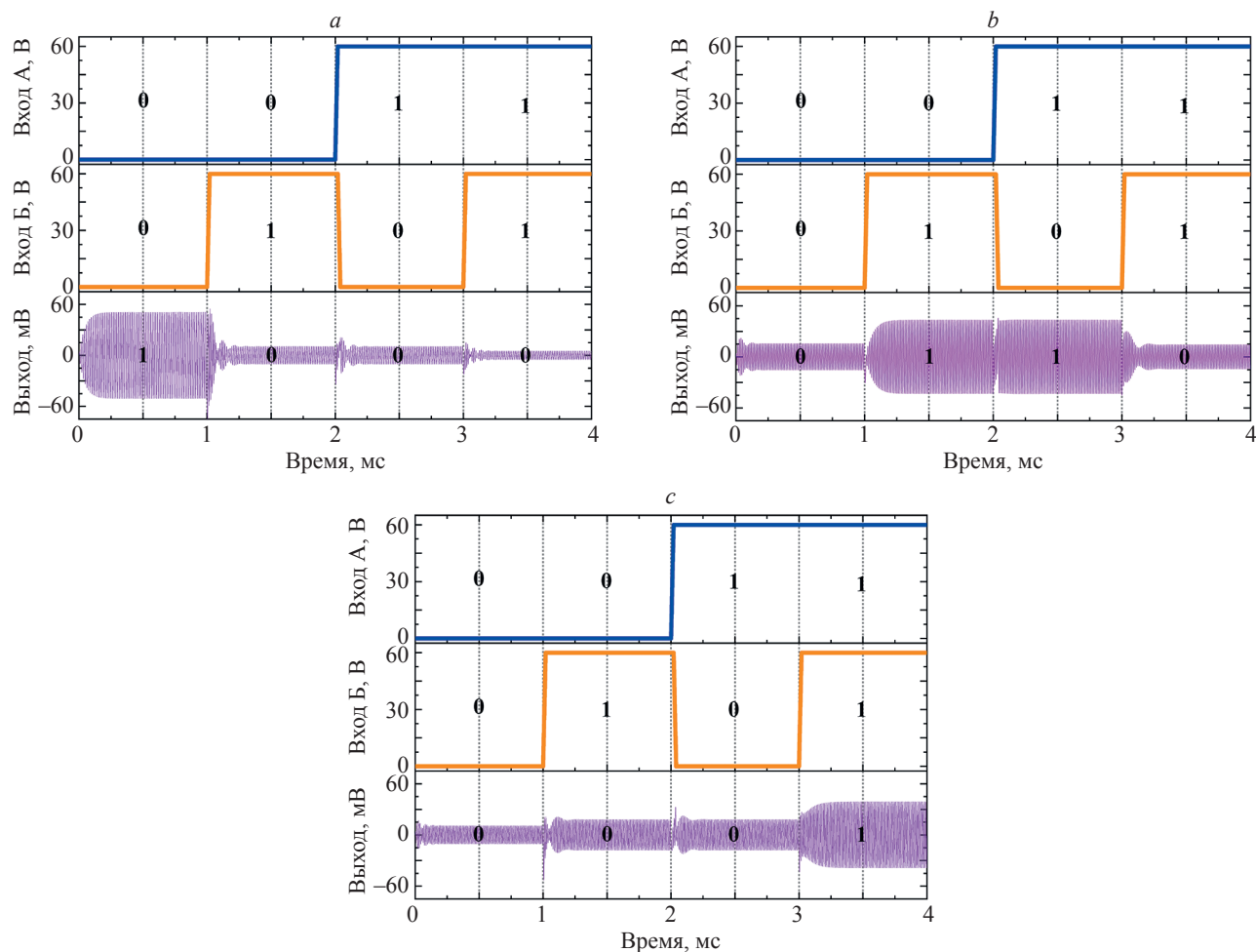


Рис. 9. Осциллограммы, демонстрирующие реализацию логических операций: «ИЛИ-НЕ» (а); «Исключающее ИЛИ» (b); «И» (с)

Fig. 9. Oscilloscope traces demonstrating the implementation of logic operations: NOR (a); XOR (b); AND (c)

Закключение

В работе представлен процесс проектирования логического микроэлектромеханического элемента на основе гребенчатого резонатора. Оптимизированы геометрические параметры емкостных ячеек и подвесов с учетом технологических норм, которые обеспечили собственную частоту колебаний 100 кГц, напряжение схлопывания 84,5 В и добротность 20 при атмосферном давлении достаточные для разделения пиков амплитудно-частотной характеристики и необходимой скорости переключения логических состояний. Чувствительность конструкции составляет 147 нФ/м за счет увеличения емкости до 0,56 пФ.

Анализ динамических характеристик подтвердил возможность выполнения логических операций путем изменения резонансной частоты подвижного электрода

за счет электростатического уменьшения жесткости резонатора.

Разработанный маршрут проектирования может применяться для создания логических МЭМС-элементов с возможностью перепрограммирования в процессе работы и дальнейшего каскадирования таких устройств для построения сложных цифровых схем. Альтернативный подход к созданию трехмерных моделей на основе библиотеки параметрических компонентов и генерации компактных моделей для системного анализа может быть использован для разработки МЭМС-акселерометров и гироскопов. Дальнейшая оптимизация конструкции будет направлена на поиск других форм инерционной массы для размещения на ней большего количества емкостных ячеек и увеличения чувствительности устройства.

Литература

References

1. Nanaiah K.C., Younis M.I. Dual electro-mechanical oscillator for dynamically reprogrammable logic gate. Patent US11031937B2, 2021.
2. Hafiz M.A.A., Kosuru L., Younis M.I. Microelectromechanical reprogrammable logic device // *Nature Communications*. 2016. V. 7. P. 11137. <https://doi.org/10.1038/ncomms11137>
3. Fariborzi H., Chen F., Nathanael R., Chen I-Ru, Hutin L., Lee R., Liu T.-J.K., Stojanovic V. Relays do not leak – CMOS does // *Proc. of the 50th ACM/EDAC/IEEE Design Automation Conference (DAC)*. 2013. P. 1–4.
4. Ilyas S., Ahmed S., Hafiz A.A., Fariborzi H., Younis M. Cascadable microelectromechanical resonator logic gate // *Journal of Micromechanics and Microengineering*. 2019. V. 29. N 1. P. 015007. <https://doi.org/10.1088/1361-6439/aaf0e6>
5. Ilyas S., Hafiz M.A.A., Fariborzi H., Younis M.I. Mechanical resonator based cascadable logic device. Patent US20190341920A1, 2021.
6. Соловьев А.А., Певцов Е.Ф., Колчужин В.А. Системное моделирование мультиточечного микроэлектромеханического логического элемента // *Нано- и микросистемная техника*. 2024. Т. 26. № 6. С. 260–267.
7. Asselot J., Krust A., Parent A., Welham C. High order MEMS models for system design // *Proc. of the IEEE International Symposium on Circuits and Systems (ISCAS)*. 2018. P. 1–5. <https://doi.org/10.1109/iscas.2018.8351644>
8. Gridchin A.V., Kolchuzhin V.A., Gridchin V.A. An optimization of initial gap in electrostatic comb drive // *Proc. on the 13th International Scientific-Technical Conference on Actual Problems of Electronics Instrument Engineering (APEIE)*. 2016. P. 13–15. <https://doi.org/10.1109/APEIE.2016.7802182>
9. Senturia S.D. *Microsystem Design*. Springer, 2005. 689 p.
10. Parker G.W. What is the capacitance of parallel plates? // *Computers in Physics*. 1991. V. 5. P. 534–540. <https://doi.org/10.1063/1.4823017>
11. Maj C., Nazdrowicz J., Stawiński A. Fringing field modelling in MEMS capacitive comb-drive accelerometers // *Methods and tools in CAD — selected issues*. 2021. P. 15–27. https://doi.org/10.24427/978-83-66391-87-1_02
12. Yazdi N., Najafi K., Sallian A. A high-sensitivity silicon accelerometer with a folded-electrode structure // *Journal of Microelectromechanical Systems*. 2003. V. 12. N 4. P. 479–486. <https://doi.org/10.1109/JMEMS.2003.815837>
13. Veijola T., Turowski M. Compact damping models for laterally moving microstructures with gas-rarefaction effects // *Journal of Microelectromechanical Systems*. 2001. V. 10. N 2. P. 263–273. <https://doi.org/10.1109/84.925777>
14. Свидетельство о государственной регистрации программы для ЭВМ № 2024682380 РФ. Программа аналитического расчета геометрических размеров логических МЭМС вентилей на основе гребенчатых резонаторов: № 2024680765. заявлено 10.09.2024. опубликовано 23.09.2024. Бюл. № 10 / Колчужин В.А., Соловьев А.А., Певцов Е.Ф. правообладатель ФГБОУ ВО «МИРЭА – Российский технологический университет».
15. Schröpfer G., Lorenz G., Krust A., Vernay B., Breit S., Mehdaoui A., Sanginario A. MEMS system-level modeling and simulation in smart systems // *Smart Systems Integration and Simulation*. 2016. P. 145–168. https://doi.org/10.1007/978-3-319-27392-1_6
16. Lorenz G., Morris A., Lakkis I. A top-down design flow for MOEMS // *Proceedings of SPIE*. 2001. V. 4408. P. 126–137. <https://doi.org/10.1117/12.425351>
17. Zhang Z., Kamon M., Daniel L. Continuation-based pull-in and lift-off simulation algorithms for microelectromechanical devices // *Journal of Microelectromechanical Systems*. 2014. V. 23. N 5. P. 1084–1093. <https://doi.org/10.1109/jmems.2014.2304967>
18. Lorenz G., Schröpfer G. 3D parametric-library-based MEMS/IC design // *System-Level Modeling of MEMS*. 2013. P. 407–424. <https://doi.org/10.1002/9783527647132.ch17>
19. Haase J., Reitz S., Wünsche S., Schwarz P., Becker U., Lorenz G., Neul R. Netzwerk- und Verhaltensmodellierung eines mikromechanischen Beschleunigungssensors // *Proc. of the Workshop Methoden und Werkzeuge zum Entwurf von Mikrosystemen im Rahmen des 2. Statusseminars zum BMBF-Verbundprojekt Modellbildung für die Mikrosystemtechnik MIMOSYS*. 1997. P. 23–30.
20. Neul R., Becker U., Lorenz G., Schwarz P., Haase J., Wünsche S. A modeling approach to include mechanical microsystem components
1. Nanaiah K.C., Younis M.I. Dual electro-mechanical oscillator for dynamically reprogrammable logic gate. Patent US11031937B2, 2021.
2. Hafiz M.A.A., Kosuru L., Younis M.I. Microelectromechanical reprogrammable logic device. *Nature Communications*, 2016, vol. 7, pp. 11137. <https://doi.org/10.1038/ncomms11137>
3. Fariborzi H., Chen F., Nathanael R., Chen I-Ru, Hutin L., Lee R., Liu T.-J.K., Stojanovic V. Relays do not leak — CMOS does. *Proc. of the 50th ACM/EDAC/IEEE Design Automation Conference (DAC)*, 2013, pp. 1–4.
4. Ilyas S., Ahmed S., Hafiz A.A., Fariborzi H., Younis M. Cascadable microelectromechanical resonator logic gate. *Journal of Micromechanics and Microengineering*, 2019, vol. 29, no. 1, pp. 015007. <https://doi.org/10.1088/1361-6439/aaf0e6>
5. Ilyas S., Hafiz M.A.A., Fariborzi H., Younis M.I. Mechanical resonator based cascadable logic device. Patent US20190341920A1, 2021.
6. Solovlev A.A., Pevtsov E.F., Kolchuzhin V.A. System modeling of a multi-contact microelectromechanical logic element. *Nano- and Microsystems Technology*, 2024, vol. 26, no. 6, pp. 260–267. (in Russian)
7. Asselot J., Krust A., Parent A., Welham C. High order MEMS models for system design. *Proc. of the IEEE International Symposium on Circuits and Systems (ISCAS)*, 2018, pp. 1–5. <https://doi.org/10.1109/iscas.2018.8351644>
8. Gridchin A.V., Kolchuzhin V.A., Gridchin V.A. An optimization of initial gap in electrostatic comb drive. *Proc. on the 13th International Scientific-Technical Conference on Actual Problems of Electronics Instrument Engineering (APEIE)*, 2016, pp. 13–15. <https://doi.org/10.1109/APEIE.2016.7802182>
9. Senturia S.D. *Microsystem Design*. Springer, 2005. 689 p.
10. Parker G.W. What is the capacitance of parallel plates? *Computers in Physics*, 1991, vol. 5, pp. 534–540. <https://doi.org/10.1063/1.4823017>
11. Maj C., Nazdrowicz J., Stawiński A. Fringing field modelling in MEMS capacitive comb-drive accelerometers. *Methods and tools in CAD — selected issues*, 2021, pp. 15–27. https://doi.org/10.24427/978-83-66391-87-1_02
12. Yazdi N., Najafi K., Sallian A. A high-sensitivity silicon accelerometer with a folded-electrode structure. *Journal of Microelectromechanical Systems*, 2003, vol. 12, no. 4, pp. 479–486. <https://doi.org/10.1109/JMEMS.2003.815837>
13. Veijola T., Turowski M. Compact damping models for laterally moving microstructures with gas-rarefaction effects. *Journal of Microelectromechanical Systems*, 2001, vol. 10, no. 2, pp. 263–273. <https://doi.org/10.1109/84.925777>
14. Kolchuzhin V.A., Solovlev A.A., Pevtsov E.F. Program for analytical calculation of geometric dimensions of logical MEMS gates based on comb resonators. Certificate of the State computer program registration RU 2024682380. 2024. (in Russian)
15. Schröpfer G., Lorenz G., Krust A., Vernay B., Breit S., Mehdaoui A., Sanginario A. MEMS system-level modeling and simulation in smart systems. *Smart Systems Integration and Simulation*, 2016, pp. 145–168. https://doi.org/10.1007/978-3-319-27392-1_6
16. Lorenz G., Morris A., Lakkis I. A top-down design flow for MOEMS. *Proceedings of SPIE*, 2001, vol. 4408, pp. 126–137. <https://doi.org/10.1117/12.425351>
17. Zhang Z., Kamon M., Daniel L. Continuation-based pull-in and lift-off simulation algorithms for microelectromechanical devices. *Journal of Microelectromechanical Systems*, 2014, vol. 23, no. 5, pp. 1084–1093. <https://doi.org/10.1109/jmems.2014.2304967>
18. Lorenz G., Schröpfer G. 3D parametric-library-based MEMS/IC design. *System-Level Modeling of MEMS*, 2013, pp. 407–424. <https://doi.org/10.1002/9783527647132.ch17>
19. Haase J., Reitz S., Wünsche S., Schwarz P., Becker U., Lorenz G., Neul R. Netzwerk- und Verhaltensmodellierung eines mikromechanischen Beschleunigungssensors. *Proc. of the Workshop Methoden und Werkzeuge zum Entwurf von Mikrosystemen im Rahmen des 2. Statusseminars zum BMBF-Verbundprojekt Modellbildung für die Mikrosystemtechnik MIMOSYS*, 1997, pp. 23–30. (in German)
20. Neul R., Becker U., Lorenz G., Schwarz P., Haase J., Wünsche S. A modeling approach to include mechanical microsystem components into the system simulation. *Proc. of the Proceedings Design, Automation and Test in Europe*, 1998, pp. 510–517. <https://doi.org/10.1109/DATE.1998.655906>

- into the system simulation // Proc. of the Proceedings Design, Automation and Test in Europe. 1998. P. 510–517. <https://doi.org/10.1109/DATE.1998.655906>
21. Kolchuzhin V., Dotzel W., Mehner J. Challenges in MEMS parametric macro-modeling based on mode superposition technique // Proc. of the 10th International Conference on Thermal, Mechanical and Multi-Physics Simulation and Experiments in Microelectronics and Microsystems, EuroSimE, 2009. P. 1–4. <https://doi.org/10.1109/ESIME.2009.4938481>
22. Kolchuzhin V. Verilog-A_components. [Электронный ресурс]. URL: https://github.com/Kolchuzhin/Verilog-A_components/tree/main/MIREA (дата обращения: 06.02.2025).
21. Kolchuzhin V., Dotzel W., Mehner J. Challenges in MEMS parametric macro-modeling based on mode superposition technique. *Proc. of the 10th International Conference on Thermal, Mechanical and Multi-Physics Simulation and Experiments in Microelectronics and Microsystems, EuroSimE*, 2009, pp. 1–4. <https://doi.org/10.1109/ESIME.2009.4938481>
22. Kolchuzhin V. *Verilog-A_components*. Available at: https://github.com/Kolchuzhin/Verilog-A_components/tree/main/MIREA (accessed: 06.02.2025).

Авторы

Соловьев Александр Анатольевич — старший преподаватель, МИРЭА — Российский технологический университет, Москва, 119454, Российская Федерация, <https://orcid.org/0009-0008-3578-0963>, solovev_aa@mirea.ru

Певцов Евгений Филиппович — кандидат технических наук, доцент, директор дизайн центра, МИРЭА — Российский технологический университет, Москва, 119454, Российская Федерация, <https://orcid.org/0000-0001-6264-1231>, pevtsov@mirea.ru

Колчужин Владимир Анатольевич — кандидат наук, ведущий инженер-конструктор, Qorvo Munich GmbH, Мюнхен, 81829, Германия, <https://orcid.org/0000-0002-6009-1499>, vladimir.kolchuzhin@ieee.org

Authors

Alexander A. Solovlev — Senior Lecturer, Russian Technology University MIREA, Moscow, 119454, Russian Federation, <https://orcid.org/0009-0008-3578-0963>, solovev_aa@mirea.ru

Evgeny F. Pevtsov — PhD, Associate Professor, Director of the Design Center, Russian Technology University MIREA, Moscow, 119454, Russian Federation, <https://orcid.org/0000-0001-6264-1231>, pevtsov@mirea.ru

Vladimir A. Kolchuzhin — PhD, Senior Design Engineer, Qorvo Munich GmbH, Munich, 81829, Germany, <https://orcid.org/0000-0002-6009-1499>, vladimir.kolchuzhin@ieee.org

Статья поступила в редакцию 10.02.2025
Одобрена после рецензирования 02.05.2025
Принята к печати 26.05.2025

Received 10.02.2025
Approved after reviewing 02.05.2025
Accepted 26.05.2025



Работа доступна по лицензии
Creative Commons
«Attribution-NonCommercial»

doi: 10.17586/2226-1494-2025-25-3-520-526

УДК 519.63

Критические нагрузки антисимметричных и смешанных форм защемленной нанопластины при двухосном сжатии

Михаил Васильевич Сухотерин¹✉, Ирина Викторовна Войтко²,
Анна Анатольевна Сосновская³

^{1,2,3} Государственный университет морского и речного флота имени адмирала С.О. Макарова, Санкт-Петербург, 198035, Российская Федерация

¹ sukhoterinmv@gumrf.ru✉, <https://orcid.org/0000-0002-8295-7089>

² voytkoiv@gumrf.ru, <https://orcid.org/0000-0003-2342-2209>

³ sosnovskayaaa@gumrf.ru, <https://orcid.org/0009-0000-6476-6783>

Аннотация

Введение. Изучен процесс вычисления спектра критических нагрузок антисимметричных и смешанных форм равновесия после потери устойчивости защемленной по контуру высокоупругой прямоугольной нанопластины (CCCC-пластина) (C — clamped edge, защемленный край) при двухосном сжатии и различных значениях нелокального параметра Эрингена. **Метод.** Искомые формы закритического равновесия представлены двумя гипербола-тригонометрическими рядами с неопределенными коэффициентами при соответствующих комбинациях нечетных и четных функций. Каждый из рядов подчинялся основному дифференциальному уравнению физического состояния Эрингена, а затем их сумма подчинялась всем граничным условиям задачи. В результате получена бесконечная однородная система линейных алгебраических уравнений относительно одной последовательности неизвестных коэффициентов рядов, содержащая в качестве основного параметра величину сжимающей нагрузки. Для поиска собственных чисел (критических нагрузок) использован впервые предложенный итерационный процесс отыскания нетривиальных решений в сочетании с методом «стрельбы».

Основные результаты. Для ряда значений нелокального параметра e_0A из рабочего диапазона $[0–2]$ теории Эрингена (0 — классическая теория) с шагом 0,25 впервые получен спектр из 10 относительных критических нагрузок. Установлено, что с ростом нелокального параметра критические нагрузки уменьшались. Краевые эффекты не обнаружены. Анализировалась точность компьютерных вычислений. Изменяемыми параметрами вычислительной программы являются относительная сжимающая нагрузка, отношение сторон пластины, значения нелокального параметра Эрингена, число итераций, количество членов в рядах, количество значащих цифр вычислительного процесса. **Обсуждение.** Предложенная методика и полученные численные результаты могут быть использованы при проектировании чувствительных элементов различных датчиков в smart-конструкциях.

Ключевые слова

прямоугольная CCCC-нанопластина, устойчивость, теория Эрингена, ряды Фурье, спектр критических нагрузок

Ссылка для цитирования: Сухотерин М.В., Войтко И.В., Сосновская А.А. Критические нагрузки антисимметричных и смешанных форм защемленной нанопластины при двухосном сжатии // Научно-технический вестник информационных технологий, механики и оптики. 2025. Т. 25, № 3. 520–526. С. doi: 10.17586/2226-1494-2025-25-3-520-526

Critical loads of antisymmetric and mixed forms of buckling of a CCCC-nanoplate under biaxial compression

Mikhail V. Sukhoterin¹✉, Irina V. Voytko², Anna A. Sosnovskaya³

^{1,2,3} Admiral Makarov State University of Maritime and Inland Shipping, Saint-Petersburg, 198035, Russian Federation

¹ sukhoterinmv@gumrf.ru✉, <https://orcid.org/0000-0002-8295-7089>

² voytkoiv@gumrf.ru, <https://orcid.org/0000-0003-2342-2209>

³ sosnovskayaaa@gumrf.ru, <https://orcid.org/0009-0000-6476-6783>

© Сухотерин М.В., Войтко И.В., Сосновская А.А., 2025

Abstract

The process of calculating the spectrum of critical loads of antisymmetric and mixed equilibrium forms after loss of stability of a contour-pinned highly elastic rectangular nanoplate (CCCC-plate) (C — clamped edge) under biaxial compression and various values of the nonlocal Eringen parameter is studied. The desired forms of supercritical equilibrium are represented by two hyperbolic-trigonometric series with indeterminate coefficients for corresponding combinations of odd and even functions. Each of the series obeyed the basic differential equation of the physical state of Eringen, and then their sum obeyed all the boundary conditions of the problem. As a result, an infinite homogeneous system of linear algebraic equations is obtained with respect to a single sequence of unknown coefficients of the series, containing as the main parameter the value of the compressive load. To find eigenvalues (critical loads), the iterative process of finding non-trivial solutions proposed by the authors in combination with the “shooting” method was used. For a number of values of the nonlocal parameter e_0A [nm] from the operating range [0–2] of the Eringen theory (0 is the classical theory) with a step of 0.25, a spectrum of 10 relative critical loads was obtained for the first time. It was found that with an increase in the nonlocal parameter critical loads decreased. No edge effects were detected. The accuracy of computer calculations was analyzed. The variable parameters of the computational program are the relative compressive load, the ratio of the sides of the plate, the values of the non-local Eringen parameter, the number of iterations, the number of members in the rows, the number of significant digits of the computational process. The proposed technique and the numerical results obtained can be used in the design of sensitive elements of various sensors in smart structures.

Keywords

rectangular CCCC-nanoplate, buckling, Eringen theory, Fourier series, critical load spectrum

For citation: Sukhoterina M.V., Voytko I.V., Sosnovskaya A.A. Critical loads of antisymmetric and mixed forms of buckling of a CCCC-nanoplate under biaxial compression. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2025, vol. 25, no. 3, pp. 520–526 (in Russian). doi: 10.17586/2226-1494-2025-25-3-520-526

Введение

В настоящее время нанопластины широко применяются в различных областях техники, а также в биологии и медицине. Особый интерес вызывает их использование в smart-конструкциях в качестве чувствительных элементов различных датчиков. Smart-конструкции представляют собой различные системы, включающие сенсоры, датчики, актуаторы, устройства обработки данных и выработки управляющих сигналов. Такие конструкции могут изменять свои свойства под действием электрических, магнитных или температурных полей и самостоятельно осуществлять ответные действия. Важными задачами при этом являются исследования колебаний и устойчивости, используемых для этих целей нанопластин. Современные технологии изготовления нанопластин позволяют задавать им нужные свойства и, в частности, высокую упругость. Высокоупругие пластины под действием сжимающих нагрузок в их плоскости могут несколько раз терять устойчивость при возрастании нагрузки и приобретать новые формы закритического равновесия. Изменение формы нанопластины через актуаторы фиксируется следящей smart-системой и инициирует соответствующий управляющий сигнал.

Условия закрепления граней нанопластины, так же как и тонкой пластины Кирхгоффа, обычно представляют собой различные сочетания свободно-опертых (S — supported), защемленных (C — clamped) и свободных (F — free) граней.

При анализе изгиба, колебаний и устойчивости наноразмерных структур широко используется нелокальная теория упругости А.С. Эрингена [1, 2], которая учитывает размерный эффект нанообъекта, а также взаимное влияние соседних точек на напряженное состояние, за счет введения так называемого нелокального параметра Эрингена (e_0A). Физическое обоснование этой теории с атомистической точки зрения динамики кристаллической решетки и с помощью метода модели-

рования молекулярной динамики описано в работе [3]. Данным методом в [4] был найден диапазон изменения нелокального параметра Эрингена e_0A от 0 до 2 нм для нанотрубок и нанопластин.

Вопросам устойчивости прямоугольных нанопластин посвящен большой ряд исследований, из которых отметим [5–14].

В [5] применен метод конечных разностей, найдены первые две критические нагрузки и получены соответствующие 3D-формы потери устойчивости SSSS- и CCCC-нанопластин.

Энергетические методы использовались в работах [6–11]. В [6] исследовалась устойчивость нанопластин при двухосном сжатии с учетом деформации поперечного сдвига. Для решения применялся принцип Гамильтона и модифицированная теория парных напряжений. В [7] рассматривалась нанопластина на упругом основании. Задача устойчивости решена методом Ритца. В [8] задача устойчивости магнито-электроупругой функционально градуированной нанопластины с пьезоэлектрическим актуатором решена с использованием тригонометрических функций. Первые (эйлеровые) критические нагрузки для 17 видов закрепления нанопластины найдены в работе [9] при различных значениях нелокального параметра Эрингена. Последующие критические значения не вычислялись. В [10] применялся метод Рэлея–Ритца и принцип Гамильтона для нанопластин средней толщины с учетом сдвиговой теории. В работе [11] устойчивость и колебания нанопластин Кирхгоффа исследованы с помощью модифицированной теории парных напряжений.

В [12] к задаче устойчивости нанопластины применялся метод симплектической суперпозиции. Получены численные результаты для некоторых видов закрепления нанопластин. Этот же метод реализован в работе [13] для консольной пластины на упругом основании в плоском магнитном поле.

При использовании различных приближенных методов остается открытым вопрос об их сходимости к

точному решению задачи, а также о точности вычислений по этим методам.

Целью настоящей работы является построение вычислительного алгоритма, позволяющего с высокой точностью находить спектр критических нагрузок для различных значений нелокального параметра Эрингена при выполнении всех условий задачи. Рассмотрены антисимметричные (A-A) и смешанные (A-S, S-A) формы равновесия после потери устойчивости защемленной по контуру прямоугольной нанопластины. При этом использовались гипербола-тригонометрические ряды, итерационный метод поиска нетривиальных решений бесконечной системы в сочетании с методом «стрельбы» при переборе значений нагрузки. Симметричные формы (S-S) исследовались в [14]. Подобный прием применялся также и в работе [15].

Постановка задачи

Дифференциальное уравнение изогнутой поверхности защемленной прямоугольной нанопластины (рисунок) после потери устойчивости от действия биаксиальных равномерных сжимающих сил одинаковой интенсивности N имеет вид [1, 2, 9]

$$D\nabla^2\nabla^2w + N\nabla^2w - \varepsilon^*N\nabla^2\nabla^2w = 0, \quad (1)$$

где $D = Eh^3/[12(1 - \nu^2)]$ — цилиндрическая жесткость пластины, E — модуль Юнга, h — толщина пластины, ν — коэффициент Пуассона; ∇^2 — двумерный оператор Лапласа; $w(x, y)$ — прогиб пластины; $\varepsilon^* = (e_0A)^2$ — постоянная наноматериала, e_0A — нелокальный параметр, e_0 — безразмерная нелокальная упругая константа Эрингена, A — внутренний характерный (обобщенный) размер нановещества, например, размер атомной решетки, расстояние между гранулами или атомами дискретной структуры (в зависимости от внутреннего строения вещества).

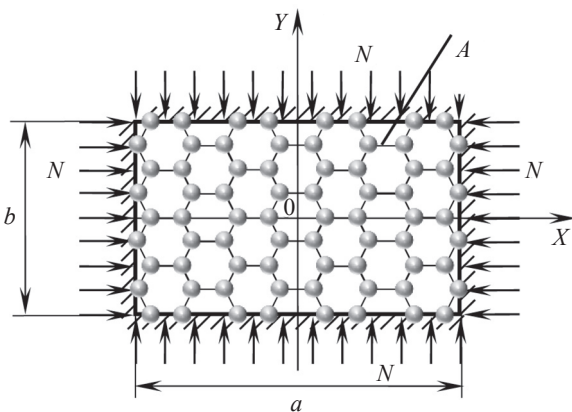


Рисунок. Расчетная схема CCCC-нанопластины.

a и b — размеры пластины в плане; A — внутренний характерный размер нановещества; N — интенсивность сжимающих усилий

Figure. Calculation scheme of CCCC-nanoplates.

Here, a and b are the dimensions of the plate in plan, A is the internal characteristic size of the nanomaterial, and N is the intensity of compressive forces

Отнесем координаты пластины к ее размеру b : $x = X/b$, $y = Y/b$, тогда размеры пластины будут такими: $-\gamma/2 \leq x \leq \gamma/2$, $-1/2 \leq y \leq 1/2$ ($\gamma = a/b$), а уравнение равновесия (1) примет вид

$$(1 - \varepsilon P)\nabla^2\nabla^2w + P\nabla^2w = 0, \quad (2)$$

где $P = Nb^2/D$ — относительная сжимающая нагрузка; $\varepsilon = \varepsilon^*/b^2$ — безразмерная константа наноматериала.

Граничные условия выражают требования отсутствия прогибов и углов поворота защемленных граней:

$$w = 0, \quad \partial w / \partial x \quad \text{при} \quad x = \pm \gamma/2, \quad (3)$$

$$w = 0, \quad \partial w / \partial y \quad \text{при} \quad y = \pm 1/2. \quad (4)$$

Сформулируем задачу: найти некоторый спектр критических нагрузок P_{icr} и получить соответствующие формы (A-A, A-S, S-A) изогнутой поверхности $w_i(x, y)$ нанопластины, удовлетворяющие уравнениям (2)–(4). Данная задача является целью поиска собственных значений и должна иметь бесчисленное множество нетривиальных решений, из которых наиболее важны несколько первых значений.

Построение решения для антисимметричных форм равновесия (A-A-формы)

Искомую функцию прогибов выберем в виде суммы двух гипербола-тригонометрических рядов по двум координатам, содержащих только нечетные функции:

$$w_1(x, y) = \sum_{k=1}^{\infty} (-1)^k [A_k \text{sh}(\alpha_{k1}x) + B_k \text{sh}(\alpha_{k2}x)] \sin(\lambda_k y), \quad (5)$$

$$w_2(x, y) = \sum_{s=1}^{\infty} (-1)^s [C_s \text{sh}(\xi_{s1}y) + D_s \text{sh}(\xi_{s2}y)] \sin(\mu_s x), \quad (6)$$

где $A_k, \alpha_{k1}, B_k, \alpha_{k2}, C_s, \xi_{s1}, D_s, \xi_{s2}$ — неопределенные коэффициенты; $\lambda_k = 2\pi k/\gamma$; $\mu_s = 2\pi s/\gamma$.

Заметим, что функция (5) удовлетворяет условию отсутствия прогибов на гранях $y = \pm 1/2$, а функция (6) — условию отсутствия прогибов на гранях $x = \pm \gamma/2$.

Коэффициенты α_{ki}, ξ_{si} ($i = 1, 2$) найдем из основного уравнения (2). Подставляя в него поочередно ряды (5) и (6) и избавившись от знака суммирования, получим два биквадратных уравнения для определения указанных коэффициентов (как и в [14]):

$$(1 - \varepsilon P)(\alpha_k^4 + \lambda_k^4) + [P - 2(1 - \varepsilon P)\lambda_k^2]\alpha_k^2 - P\lambda_k^2 = 0,$$

$$(1 - \varepsilon P)(\xi_s^4 + \mu_s^4) + [P - 2(1 - \varepsilon P)\mu_s^2]\xi_s^2 - P\mu_s^2 = 0,$$

корни которых

$$\begin{aligned} \alpha_{k1} &= \lambda_k, \quad \alpha_{k2} = \sqrt{\lambda_k^2 - P/(1 - \varepsilon P)}, \\ \xi_{s1} &= \mu_s, \quad \xi_{s2} = \sqrt{\mu_s^2 - P/(1 - \varepsilon P)}. \end{aligned} \quad (7)$$

Отрицательные корни не будем учитывать, так как ряды (5), (6) изменят лишь знак, а формы изменят направление на противоположное.

Потребуем теперь, чтобы ряды (5), (6) удовлетворяли и условиям отсутствия прогибов на гранях $x = \pm \gamma/2$

и $y = \pm 1/2$ соответственно (выражения в квадратных скобках должны при этом обратиться в нуль). Это дает

$$B_k = -A_k \operatorname{sh} \alpha_{k1}^* / \operatorname{sh} \alpha_{k2}^*, D_s = -\operatorname{sh} \xi_{s1}^* / \operatorname{sh} \xi_{s2}^*,$$

где $\alpha_{k1}^* = \alpha_{k1} \gamma / 2$; $\alpha_{k2}^* = \alpha_{k2} \gamma / 2$; $\xi_{s1}^* = \xi_{s1} / 2$; $\xi_{s2}^* = \xi_{s2} / 2$, и функции прогибов примут вид:

$$w_1(x, y) = \sum_{k=1}^{\infty} (-1)^k A_k^* \left[\frac{\operatorname{sh}(\alpha_{k1} x)}{\operatorname{sh} \alpha_{k1}^*} - \frac{\operatorname{sh}(\alpha_{k2} x)}{\operatorname{sh} \alpha_{k2}^*} \right] \sin(\lambda_k y), \quad (8)$$

$$w_2(x, y) = \sum_{s=1}^{\infty} (-1)^s C_s^* \left[\frac{\operatorname{sh}(\xi_{s1} y)}{\operatorname{sh} \xi_{s1}^*} - \frac{\operatorname{sh}(\xi_{s2} y)}{\operatorname{sh} \xi_{s2}^*} \right] \sin(\mu_s x), \quad (9)$$

где $A_k^* = A_k \operatorname{sh} \alpha_{k1}^*$; $C_s^* = C_s \operatorname{sh} \xi_{s1}^*$.

Подчиним функции (8) и (9) вторым условиям (3), (4), т. е. условиям отсутствия поворотов защемленных граней, что дает систему двух уравнений для определения коэффициентов A_k^* и C_s^* :

$$\begin{cases} \sum_{k=1}^{\infty} (-1)^k A_k^* (\alpha_{k1} \operatorname{cth} \alpha_{k1}^* - \alpha_{k2} \operatorname{cth} \alpha_{k2}^*) \sin(\lambda_k y) + \\ + \sum_{s=1}^{\infty} C_s^* \mu_s \left(\frac{\operatorname{sh}(\xi_{s1} y)}{\operatorname{sh} \xi_{s1}^*} - \frac{\operatorname{sh}(\xi_{s2} y)}{\operatorname{sh} \xi_{s2}^*} \right) = 0, \\ \sum_{k=1}^{\infty} A_k^* \lambda_k \left(\frac{\operatorname{sh}(\alpha_{k1} x)}{\operatorname{sh} \alpha_{k1}^*} - \frac{\operatorname{sh}(\alpha_{k2} x)}{\operatorname{sh} \alpha_{k2}^*} \right) + \\ + \sum_{s=1}^{\infty} (-1)^s C_s^* (\xi_{s1} \operatorname{cth} \xi_{s1}^* - \xi_{s2} \operatorname{cth} \xi_{s2}^*) \sin(\mu_s x) = 0. \end{cases} \quad (10)$$

Чтобы упростить систему (10), надо второй ряд в первом уравнении представить рядом Фурье по $\sin(\lambda_k y)$, а первый ряд во втором уравнении — рядом Фурье по $\sin(\mu_s x)$. Затем можно избавиться от знаков суммирования.

Разложим гиперболические функции в соответствующие ряды Фурье

$$\operatorname{sh}(\alpha_k x) = \frac{4 \operatorname{sh} \alpha_k^*}{\gamma} \sum_{s=1}^{\infty} \frac{(-1)^s \mu_s \sin(\mu_s x)}{\alpha_k^2 + \mu_s^2};$$

$$\operatorname{sh}(\xi_s y) = 4 \operatorname{sh} \xi_s^* \sum_{k=1}^{\infty} \frac{(-1)^k \lambda_k \sin(\lambda_k y)}{\xi_s^2 + \lambda_k^2}$$

и подставим их в систему (10). С учетом формулы (7) после перестановки знаков суммирования и освобождения от знака внешней суммы получим систему уравнений

$$\begin{cases} A_k^* (\alpha_{k1} \operatorname{cth} \alpha_{k1}^* - \alpha_{k2} \operatorname{cth} \alpha_{k2}^*) + \\ + 4 \lambda_k P_\varepsilon \sum_{s=1}^{\infty} \frac{C_s^* \mu_s}{(\xi_{s1}^2 + \lambda_k^2)(\xi_{s2}^2 + \lambda_k^2)} = 0, \\ \frac{4}{\gamma} \mu_s P_\varepsilon \sum_{k=1}^{\infty} \frac{A_k^* \lambda_k}{(\alpha_{k1}^2 + \mu_s^2)(\alpha_{k2}^2 + \mu_s^2)} + \\ + C_s^* (\xi_{s1} \operatorname{cth} \xi_{s1}^* - \xi_{s2} \operatorname{cth} \xi_{s2}^*) = 0, \end{cases} \quad (11)$$

где $P_\varepsilon = P/(1 - \varepsilon P)$.

Из второго уравнения системы (11) выразим C_s^* через A_k^* :

$$C_s^* = \frac{-4 P_\varepsilon \mu_s}{\gamma \theta_s} \sum_{k=1}^{\infty} \frac{A_k^* \lambda_k}{\sigma_{ks} (\sigma_{ks} - P_\varepsilon)}, \quad (12)$$

где $\theta_s = \xi_{s1} \operatorname{cth} \xi_{s1}^* - \xi_{s2} \operatorname{cth} \xi_{s2}^*$; $\sigma_{ks} = \lambda_k^2 + \mu_s^2$.

Подставим (12) в первое уравнение системы (11) и получим разрешающее однородное линейное алгебраическое уравнение:

$$A_k^* = \frac{16 P_\varepsilon^2 \lambda_k}{\gamma \varphi_k} \sum_{s=1}^{\infty} \frac{\mu_s^2}{\sigma_{ks} (\sigma_{ks} - P_\varepsilon)} \sum_{k=1}^{\infty} \frac{\lambda_k A_k^*}{\sigma_{ks} (\sigma_{ks} - P_\varepsilon)}, \quad (13)$$

где $\varphi_k = \alpha_{k1} \operatorname{cth} \alpha_{k1}^* - \alpha_{k2} \operatorname{cth} \alpha_{k2}^*$.

Уравнение (13) фактически представляет собой бесконечную однородную систему линейных алгебраических уравнений относительно одной последовательности неизвестных коэффициентов A_k^* . В качестве параметра система (12) содержит величину сжимающей нагрузки P . Искомые критические значения (собственные числа задачи) обеспечат нетривиальные решения этой системы и соответствующие формы закритического равновесия.

Система (13) записана в стандартной форме, позволяющей считать коэффициенты A_k^* под знаком суммы в правой части равенства предыдущей итерацией и организовать процесс последовательных приближений для отыскания коэффициентов последующей итерации при заданном значении нагрузки. Если нагрузка подобрана такой, что является критической, то, начиная с некоторой итерации, все соответствующие коэффициенты будут совпадать. Этого можно достичь простым перебором значений P , используя «метод стрельбы». Данный итерационный процесс более эффективен, чем традиционный метод поиска собственных чисел из условия равенства нулю определителя бесконечной однородной системы линейных алгебраических уравнений.

Сходимость числовых рядов разрешающего уравнения

Оценим порядок и быстроту сходимости рядов, входящих в выражение (13).

Если все коэффициенты A_{k0}^* в правой части для начального приближения положить равными $1/\lambda_k$, то внутренний ряд мажорируется рядом, точная сумма которого известна [16, 17]:

$$\sum_{k=1}^{\infty} \frac{1}{(\lambda_k^2 + \mu_s^2)^2} = -\frac{1}{2 \mu_s^4} + \frac{1}{8 \mu_s^3} \left(\operatorname{cth} \mu_s^* + \frac{\mu_s^*}{\operatorname{sh}^2 \mu_s^*} \right), \quad (14)$$

где $\mu_s^* = \mu_s / 2$. Следовательно оценка суммы ряда (14) по индексу s :

$$\sum_{k=1}^{\infty} \frac{1}{(\lambda_k^2 + \mu_s^2)^2} = O\left(\frac{1}{\mu_s^3}\right). \quad (15)$$

С учетом (15) внешняя сумма в формуле (13) будет мажорироваться рядом

$$\sum_{s=1}^{\infty} \frac{1}{\mu_s(\lambda_k^2 + \mu_s^2)\theta_s}. \quad (16)$$

Оценим теперь θ_s :

$$\theta_s = \xi_{s1} \operatorname{cth} \xi_{s1}^* - \xi_{s2} \operatorname{cth} \xi_{s2}^* \approx \xi_{s1} - \xi_{s2} = \frac{\xi_{s1}^2 - \xi_{s2}^2}{\xi_{s1} + \xi_{s2}} = \frac{P_\varepsilon}{\mu_s + \sqrt{\mu_s^2 - P_\varepsilon}} = \frac{P_\varepsilon}{\mu_s(1 + \sqrt{1 - P_\varepsilon/\mu_s^2})},$$

поэтому $\theta_s = 0(1/\mu_s)$. Тогда мажорантный ряд (16) преобразуется к виду, аналогичному (14) с соответствующей оценкой его суммы:

$$\sum_{s=1}^{\infty} \frac{1}{(\lambda_k^2 + \mu_s^2)^2} = 0\left(\frac{1}{\lambda_k^3}\right).$$

Оценка $\phi_k = 0(1/\lambda_k)$ аналогична θ_s . Тогда порядок коэффициентов в левой части формулы (13) будет $A_k^* = 0(1/\lambda_k)$, т. е. такой же, как и для начальных значений A_{k0}^* .

Это означает, что ряды в разрешающей системе (13) быстро сходятся: не хуже ряда с общим членом $1/k^3$ или $1/s^3$.

Решение для антисимметрично-симметричных форм (А-S-формы)

Функции прогибов будут содержать в случае А-S-формы нечетные функции по переменной x и четные — по переменной y :

$$w_1(x, y) = \sum_{k=1,3,\dots}^{\infty} (-1)^{\tilde{k}} [A_k \operatorname{sh}(\alpha_{k1}x) + B_k \operatorname{sh}(\alpha_{k2}x)] \cos(\lambda_k y),$$

$$w_2(x, y) = \sum_{s=1}^{\infty} (-1)^{\tilde{s}} [C_s \operatorname{ch}(\xi_{s1}y) + D_s \operatorname{ch}(\xi_{s2}y)] \sin(\mu_s x),$$

где $\lambda_k = \pi k$ ($k = 1, 3, \dots$); $\mu_s = 2\pi s/\gamma$ ($s = 1, 2, \dots$); $\tilde{k} = (k+1)/2$; α_k, ξ_s — коэффициенты, которые определяются формулами (7). Отрицательные значения коэффициентов не рассматривались, так как формы равновесия отличаются только противоположным направлением.

По аналогии со случаем А-А-формы после аналогичных действий была получена разрешающая система

$$A_k^* = \frac{16P_\varepsilon^2 \lambda_k}{\gamma \phi_k} \sum_{s=1}^{\infty} \frac{\mu_s^2}{\sigma_{ks}(\sigma_{ks} - P_\varepsilon) \theta_s} \sum_{k=1,3,\dots}^{\infty} \frac{\lambda_k A_k^*}{\sigma_{ks}(\sigma_{ks} - P_\varepsilon)}, \quad (17)$$

где $A_k^* = A_k \operatorname{sh} \alpha_{k1}^*$; $\sigma_{ks} = \lambda_k^2 + \mu_s^2$; $\phi_k = \alpha_{k1} \operatorname{cth} \alpha_{k1}^* - \alpha_{k2} \operatorname{cth} \alpha_{k2}^*$; $\theta_s = \xi_{s1} \operatorname{th} \xi_{s1}^* - \xi_{s2} \operatorname{th} \xi_{s2}^*$.

Система (17) имеет такое же выражение, как и (13), которое отличается тем, что имеет $k = 1, 3, \dots$, и в выражении для θ_s функции cth заменены на th .

Решение для симметрично-антисимметричных форм (S-A-формы)

Функции прогибов выбираются такими

$$w_1(x, y) = \sum_{k=1}^{\infty} (-1)^k [A_k \operatorname{ch}(\alpha_{k1}x) + B_k \operatorname{ch}(\alpha_{k2}x)] \sin(\lambda_k y),$$

$$w_2(x, y) = \sum_{s=1,3,\dots}^{\infty} (-1)^{\tilde{s}} [C_s \operatorname{sh}(\xi_{s1}y) + D_s \operatorname{sh}(\xi_{s2}y)] \cos(\mu_s x),$$

где $\lambda_k = 2\pi k$ ($k = 1, 2, \dots$); $\mu_s = \pi s/\gamma$ ($s = 1, 3, \dots$); $\tilde{s} = (s+1)/2$.

В случае S-A-формы по сравнению с А-А-формой все нечетные и четные функции поменялись местами.

Разрешающая система будет аналогична выражению (17):

$$A_k^* = \frac{16P_\varepsilon^2 \lambda_k}{\gamma \phi_k} \sum_{s=1,3,\dots}^{\infty} \frac{\mu_s^2}{\sigma_{ks}(\sigma_{ks} - P_\varepsilon) \theta_s} \sum_{k=1}^{\infty} \frac{\lambda_k A_k^*}{\sigma_{ks}(\sigma_{ks} - P_\varepsilon)},$$

где $A_k^* = A_k \operatorname{ch} \alpha_{k1}^*$; $\phi_k = \alpha_{k1} \operatorname{th} \alpha_{k1}^* - \alpha_{k2} \operatorname{th} \alpha_{k2}^*$; $\theta_s = \xi_{s1} \operatorname{cth} \xi_{s1}^* - \xi_{s2} \operatorname{cth} \xi_{s2}^*$.

Численные результаты. Обсуждение

Для получения спектра критических нагрузок и соответствующих форм равновесия были разработаны вычислительные программы в среде Maple, в которых был предусмотрен вывод на печать коэффициентов A_k^* и C_s^* на каждой итерации при заданной нагрузке. Программы позволяли менять отношение сторон нанопластины, число итераций и количество членов в рядах. Основным изменяемым параметром являлась сжимающая нагрузка, перебором которой методом «стрельбы» выявлялись критические ее значения. Для каждого значения нелокального параметра $e_0 A \in [0-2]$ нм с шагом 0,25 получены первые 10 относительных критических нагрузок и соответствующие формы (здесь не приводятся) закритического равновесия для квадратной нанопластины со стороной 5 нм (включая результаты для симметричных форм, полученные в работе [14]). Значение $e_0 A = 0$ соответствует классической пластине Кирхгофа.

Заметим, что для квадратной пластины в силу симметрии S-A- и А-S-форм и критические нагрузки совпадают.

В таблице приведены значения первых 10 относительных критических нагрузок для различных значений нелокального параметра $e_0 A$, а также для сравнения значения первых критических нагрузок, полученных в работе [9].

При реализации вычислительного процесса число итераций, количество членов в рядах и количество значащих цифр увеличивалось для оценки точности вычислений критических нагрузок и коэффициентов гипербола-тригонометрических рядов. Окончательно приняты значения этих параметров соответственно 20, 99 и 30. Это обеспечивало совпадение до 4–5 значащих цифр соответствующих коэффициентов последних соседних итераций при «попадании» на критическую нагрузку, которая при этом содержала четыре знака после запятой. Дальнейшее увеличение этих параметров не оказывало существенного влияния на точность вычислений.

Данные, приведенные в таблице, показывают, что с увеличением нелокального параметра критические нагрузки убывают. При анализе полученных 3D-форм равновесия каких-либо краевых эффектов не обнаруже-

Таблица. Критические безразмерные сжимающие нагрузки $P = Nb^2/D$ квадратной нанопластины для S-S-, A-A- и S-A-форм равновесия при различных значениях нелокального параметра

Table. Critical dimensionless compressive loads $P = Nb^2/D$ of a square CCCC-nanoplate for S-S, A-A и S-A equilibrium forms at different values of the nonlocal parameter

Форма	e_0A , нм								
	0	0,25	0,5	0,75	1,0	1,25	1,5	1,75	2,0
S-S	52,3450	46,2875	34,3594	24,0361	16,9193	12,2543	9,1656	7,0619	5,5833
S-A	92,1244	74,8790	47,9504	29,9806	19,6638	13,6324	9,9152	7,4988	5,8529
A-A	128,2100	97,0900	56,1805	33,0036	20,9206	14,2248	10,2250	7,6746	5,9595
S-S	167,0292	117,8276	62,5509	35,1038	21,7453	14,6013	10,4180	7,7828	6,0246
S-A	189,5720	128,6170	65,4665	36,0035	22,0872	14,7547	10,4959	7,8263	6,0505
S-S	246,3270	152,4470	71,1255	37,6511	22,6965	15,0241	10,6316	7,9014	6,0953
A-A	279,0810	164,3875	73,6207	38,3389	22,9447	15,1324	10,6857	7,9313	6,1131
S-A	326,6430	179,8095	76,5612	39,1214	23,2226	15,2529	10,7456	7,9642	6,1327
S-S	434,7870	208,3342	81,3010	40,3226	23,6407	15,4321	10,8342	8,0128	6,1614
A-A	630,4320	244,7240	86,3092	41,5175	24,3512	15,6040	10,9187	8,0589	6,1886
P_1 [9]	52,4550	46,3739	—	24,0594	16,9308	—	—	—	5,5846

Примечание. Прочерки в таблице означают, что эти результаты в работе [9] отсутствуют.

но. Возможно, они будут проявляться, если уравнение состояния Эрингена (1) брать в более точной интегральной форме [14].

В работе [9] получены только эйлеровые значения. Последующие критические нагрузки в этой работе не вычислялись. Сравнение показывает очень хорошее совпадение, хотя результаты [9] несколько выше соответствующих значений настоящей работы. Заметим, что энергетические методы дают, как правило, завышенные результаты при вычислениях.

Закключение

Полученные результаты могут быть использованы в проектировании smart-конструкций с нанопластинами в качестве чувствительных элементов различных датчиков, реагирующих на биаксиальное сжатие пластины в ее срединной плоскости.

Предполагается в дальнейшем исследовать устойчивость нанопластин с использованием разрешающего уравнения состояния Эрингена в интегральной форме для выявления возможных краевых эффектов.

Литература

1. Eringen A.C. Linear theory of nonlocal elasticity and dispersion of plane waves // *International Journal of Engineering Science*. 1972. V. 10. N 5. P. 425–435. [https://doi.org/10.1016/0020-7225\(72\)90050-X](https://doi.org/10.1016/0020-7225(72)90050-X)
2. Eringen A.C. On differential equations of nonlocal elasticity and solutions of screw dislocation and surface waves // *Journal of Applied Physics*. 1983. V. 54. N 9. P. 4703–4710. <https://doi.org/10.1063/1.332803>
3. Chen Y., Lee J.D., Eskandarian A. Atomistic viewpoint of the applicability of microcontinuum theories // *International Journal of Solids and Structures*. 2004. V. 41. N 8. P. 2085–2097. <https://doi.org/10.1016/j.ijsolstr.2003.11.030>
4. Duan W.H., Wang C.M., Zhang Y.Y. Calibration of nonlocal scaling effect parameter for free vibration of carbon nanotubes by molecular dynamics // *Journal of Applied Physics*. 2007. V. 101. N 2. P. 024305. <https://doi.org/10.1063/1.2423140>
5. Ravari M.R.K., Talebi S., Shahidi A.R. Analysis of the buckling of rectangular nanoplates by use of finite-difference method // *Meccanica*. 2014. V. 49. N 6. P. 1443–1455. <https://doi.org/10.1007/s11012-014-9917-x>
6. Malikan M. Analytical predictions for the buckling of a nanoplate subjected to non-uniform compression based on the four-variable plate theory // *Journal of Applied and Computational Mechanics*. 2017. V. 3. N 3. P. 218–228. <https://doi.org/10.22055/jacm.2017.21757.1115>
7. Bastami M., Behjat B. Ritz solution of buckling and vibration problem of nanoplates embedded in an elastic medium // *Sigma Journal of Engineering and Natural Sciences*. 2017. V. 35. N 2. P. 285–302.
8. Ebrahimi F., Barati M.R. Buckling analysis of piezoelectrically actuated smart nanoscale plates subjected to magnetic field // *Journal of Intelligent Material Systems and Structures*. 2017. V. 28. N 11. P. 1472–1490. <https://doi.org/10.1177/1045389x16672569>
9. Wang Z., Xing Y., Sun Q., Yang Y. Highly accurate closed-form solutions for free vibration and eigenbuckling of rectangular

References

1. Eringen A.C. Linear theory of nonlocal elasticity and dispersion of plane waves. *International Journal of Engineering Science*, 1972, vol. 10, no. 5, pp. 425–435. [https://doi.org/10.1016/0020-7225\(72\)90050-X](https://doi.org/10.1016/0020-7225(72)90050-X)
2. Eringen A.C. On differential equations of nonlocal elasticity and solutions of screw dislocation and surface waves. *Journal of Applied Physics*, 1983, vol. 54, no. 9, pp. 4703–4710. <https://doi.org/10.1063/1.332803>
3. Chen Y., Lee J.D., Eskandarian A. Atomistic viewpoint of the applicability of microcontinuum theories. *International Journal of Solids and Structures*, 2004, vol. 41, no. 8, pp. 2085–2097. <https://doi.org/10.1016/j.ijsolstr.2003.11.030>
4. Duan W.H., Wang C.M., Zhang Y.Y. Calibration of nonlocal scaling effect parameter for free vibration of carbon nanotubes by molecular dynamics. *Journal of Applied Physics*, 2007, vol. 101, no. 2, pp. 024305. <https://doi.org/10.1063/1.2423140>
5. Ravari M.R.K., Talebi S., Shahidi A.R. Analysis of the buckling of rectangular nanoplates by use of finite-difference method. *Meccanica*, 2014, vol. 49, no. 6, pp. 1443–1455. <https://doi.org/10.1007/s11012-014-9917-x>
6. Malikan M. Analytical predictions for the buckling of a nanoplate subjected to non-uniform compression based on the four-variable plate theory. *Journal of Applied and Computational Mechanics*, 2017, vol. 3, no. 3, pp. 218–228. <https://doi.org/10.22055/jacm.2017.21757.1115>
7. Bastami M., Behjat B. Ritz solution of buckling and vibration problem of nanoplates embedded in an elastic medium. *Sigma Journal of Engineering and Natural Sciences*, 2017, vol. 35, no. 2, pp. 285–302.
8. Ebrahimi F., Barati M.R. Buckling analysis of piezoelectrically actuated smart nanoscale plates subjected to magnetic field. *Journal of Intelligent Material Systems and Structures*, 2017, vol. 28, no. 11, pp. 1472–1490. <https://doi.org/10.1177/1045389x16672569>

- nanoplates // *Composite Structures*. 2019. V. 210. P. 822–830. <https://doi.org/10.1016/j.compstruct.2018.11.094>
10. Chwal M., Muc A. Buckling and free vibrations of nanoplates—comparison of nonlocal strain and stress approaches // *Applied Sciences*. 2019. V. 9. N 7. P. 1409. <https://doi.org/10.3390/app9071409>
 11. Shahraki M.E., Jam J.E. Investigating the buckling and vibration of a Kirchhoff rectangular nanoplate using modified couple stress theory // *University Proceedings. Volga Region. Physical and Mathematical Sciences*. 2023. N 4 (68). P. 75–89. <https://doi.org/10.21685/2072-3040-2023-4-7>
 12. Zheng X., Huang M., An D., Zhou C., Li R. New analytic bending, buckling, and free vibration solutions of rectangular nanoplates by the symplectic superposition method // *Scientific Reports*. 2021. V. 11. N 1. P. 2939. <https://doi.org/10.1038/s41598-021-82326-w>
 13. Wang W., Rong D., Xu C., Zhang J., Xu X., Zhou Z. Accurate buckling analysis of magnetically affected cantilever nanoplates subjected to in-plane magnetic fields // *Journal of Vibration Engineering & Technologies*. 2020. V. 8. N 4. P. 505–515. <https://doi.org/10.1007/s42417-019-00106-3>
 14. Сухотерин М.В., Сосновская А.А. Потеря устойчивости защемленной по контуру прямоугольной нанопластины // *Научно-технический вестник информационных технологий, механики и оптики*. 2024. Т. 24. № 4. С. 629–636. <https://doi.org/10.17586/2226-1494-2024-24-4-629-636>
 15. Сухотерин М.В., Распутин Е.И., Пижурин Н.Ф. Смешанные формы свободных колебаний прямоугольной CFCF-пластины // *Научно-технический вестник информационных технологий, механики и оптики*. 2023. Т. 23. № 2. С. 413–421. <https://doi.org/10.17586/2226-1494-2023-23-2-413-421>
 16. Сухотерин М.В. Изгиб консольных пластин поперечной нагрузкой: автореферат диссертации на соискание ученой степени кандидата физико-математических наук / Ленинградский Политехнический институт им. М.И. Калинина. Л., 1978. 16 с.
 17. Абрамян Б.Л. Об одной осесимметричной задаче для сплошного весомого цилиндра конечной длины // *Механика твердого тела*. 1983. № 1. С. 55–62.
 9. Wang Z., Xing Y., Sun Q., Yang Y. Highly accurate closed-form solutions for free vibration and eigenbuckling of rectangular nanoplates. *Composite Structures*, 2019, vol. 210, pp. 822–830. <https://doi.org/10.1016/j.compstruct.2018.11.094>
 10. Chwal M., Muc A. Buckling and free vibrations of nanoplates—comparison of nonlocal strain and stress approaches. *Applied Sciences*, 2019, vol. 9, no. 7, pp. 1409. <https://doi.org/10.3390/app9071409>
 11. Shahraki M.E., Jam J.E. Investigating the buckling and vibration of a Kirchhoff rectangular nanoplate using modified couple stress theory. *University Proceedings. Volga Region. Physical and Mathematical Sciences*, 2023, no. 4 (68), pp. 75–89. <https://doi.org/10.21685/2072-3040-2023-4-7>
 12. Zheng X., Huang M., An D., Zhou C., Li R. New analytic bending, buckling, and free vibration solutions of rectangular nanoplates by the symplectic superposition method. *Scientific Reports*, 2021, vol. 11, no. 1, pp. 2939. <https://doi.org/10.1038/s41598-021-82326-w>
 13. Wang W., Rong D., Xu C., Zhang J., Xu X., Zhou Z. Accurate buckling analysis of magnetically affected cantilever nanoplates subjected to in-plane magnetic fields. *Journal of Vibration Engineering & Technologies*, 2020, vol. 8, no. 4, pp. 505–515. <https://doi.org/10.1007/s42417-019-00106-3>
 14. Sukhoterina M.V., Sosnovskaya A.A. Instability of a rectangular CCCC-nanoplate. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2024, vol. 24, no. 4, pp. 629–636. (in Russian). <https://doi.org/10.17586/2226-1494-2024-24-4-629-636>
 15. Sukhoterina M.V., Rasputina E.I., Pizhurina N.F. Mixed forms of free oscillations of a rectangular CFCFplate. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2023, vol. 23, no. 2, pp. 413–421. (in Russian). <https://doi.org/10.17586/2226-1494-2023-23-2-413-421>
 16. Sukhoterina M.V. *Bending of cantilever plates under transverse load*. Abstract of a dissertation for the degree of candidate of physical and mathematical sciences. Leningrad, 1978, 16 p. (in Russian)
 17. Abramian B.L. On an axisymmetric problem for a solid weighty cylinder of finite length. *Mechanics of Solids*, 1983, no. 1, pp. 55–62. (in Russian)

Авторы

Сухотерин Михаил Васильевич — доктор технических наук, доцент, заведующий кафедрой, Государственный университет морского и речного флота имени адмирала С.О. Макарова, Санкт-Петербург, 198035, Российская Федерация, <https://orcid.org/0000-0002-8295-7089>, sukhoterimv@gumrf.ru

Войтко Ирина Викторовна — кандидат технических наук, доцент, доцент, Государственный университет морского и речного флота имени адмирала С.О. Макарова, Санкт-Петербург, 198035, Российская Федерация, <https://orcid.org/0000-0003-2342-2209>, voytkoiv@gumrf.ru

Сосновская Анна Анатольевна — старший преподаватель, Государственный университет морского и речного флота имени адмирала С.О. Макарова, Санкт-Петербург, 198035, Российская Федерация, <https://orcid.org/0009-0000-6476-6783>, sosnovskayaaa@gumrf.ru

Authors

Mikhail V. Sukhoterina — D.Sc., Associate Professor, Head of Department, Admiral Makarov State University of Maritime and Inland Shipping, Saint-Petersburg, 198035, Russian Federation, <https://orcid.org/0000-0002-8295-7089>, sukhoterimv@gumrf.ru

Irina V. Voytko — PhD, Associate Professor, Associate Professor, Admiral Makarov State University of Maritime and Inland Shipping, Saint-Petersburg, 198035, Russian Federation, <https://orcid.org/0000-0003-2342-2209>, voytkoiv@gumrf.ru

Anna A. Sosnovskaya — Senior Lecturer, Admiral Makarov State University of Maritime and Inland Shipping, Saint-Petersburg, 198035, Russian Federation, <https://orcid.org/0009-0000-6476-6783>, sosnovskayaaa@gumrf.ru

Статья поступила в редакцию 19.12.2024

Одобрена после рецензирования 17.03.2025

Принята к печати 22.05.2025

Received 19.12.2024

Approved after reviewing 17.03.2025

Accepted 22.05.2025



Работа доступна по лицензии
Creative Commons
«Attribution-NonCommercial»

doi: 10.17586/2226-1494-2025-25-3-527-535

УДК 519.233.22

**О свойствах компромиссных M -оценок,
оптимизирующих весовую L_2 -норму функции влияния**
Даниил Валерьевич Лисицин¹, Константин Викторович Гаврилов²✉

^{1,2} Новосибирский государственный технический университет, Новосибирск, 630073, Российская Федерация

¹ lisitsin@ami.nstu.ru, <https://orcid.org/0009-0007-0723-6285>

² aenigma77@mail.ru ✉, <https://orcid.org/0009-0001-9252-9942>

Аннотация

Введение. В работе развивается теория M -оценок, оптимизирующих весовую L_2 -норму функции влияния. Указанный критерий качества оценивания является достаточно общим и, кроме того, позволяет получать решения, относящиеся к классу сниженных оценок, т. е. обладающие свойством устойчивости к асимметричному засорению. Такие оценки исследовались в рамках локально устойчивого подхода д.т.н. А.М. Шурыгина, основанного на анализе показателя неустойчивости оценки (L_2 -нормы функции влияния), и его подхода, основанного на модели байесовского точечного засорения. В работе исследуется компромиссное семейство оценок, для которого оптимизируемый функционал представляет собой выпуклую линейную комбинацию двух базовых критериев. Компромиссное семейство подобно условно оптимальному семейству оценок, предложенному А.М. Шурыгиным, но используемые критерии могут быть квадратами весовых L_2 -норм функции влияния с произвольными заранее известными весовыми функциями. **Метод.** Рассмотренная предметная область до настоящего времени оставалась малоизученной. При проведении исследований применялась ранее разработанная авторами настоящей работы теория, которая описывает свойства оценок, оптимизирующих весовую L_2 -норму функции влияния. **Основные результаты.** Получен ряд свойств компромиссных оценок, показана единственность элементов семейства. Отдельно рассмотрен член семейства, доставляющий равные значения критериям: показано, что данная оценка соответствует седловой точке оптимизируемого функционала, а также является минимаксным решением относительно базовых критериев на множестве всех регулярных оценочных функций. **Обсуждение.** Построенная теория проиллюстрирована на примере задачи оценивания математического ожидания нормального распределения в условиях целенаправленного вредоносного воздействия на набор данных (аналогичного атаке отравления данных во вредоносном машинном обучении).

Ключевые слова

M -оценки, робастная статистика, функция влияния, устойчивые оценки, сниженные оценки, седловая точка, отравление данных

Ссылка для цитирования: Лисицин Д.В., Гаврилов К.В. О свойствах компромиссных M -оценок, оптимизирующих весовую L_2 -норму функции влияния // Научно-технический вестник информационных технологий, механики и оптики. 2025. Т. 25, № 3. С. 527–535. doi: 10.17586/2226-1494-2025-25-3-527-535

**On the properties of compromise M -estimators optimizing weighted L_2 -norm
of the influence function**

Daniil V. Lisitsin¹, Konstantin V. Gavrilov²✉

^{1,2} Novosibirsk State Technical University, Novosibirsk, 630073, Russian Federation

¹ lisitsin@ami.nstu.ru, <https://orcid.org/0009-0007-0723-6285>

² aenigma77@mail.ru ✉, <https://orcid.org/0009-0001-9252-9942>

Abstract

The paper develops a theory of M -estimators optimizing the weighted L_2 -norm of the influence function. The specified criterion of the estimation quality is quite general and, in addition, allows obtaining solutions related to the class of redescending estimators, i.e., possessing the property of stability to asymmetric contamination. Such estimators, in particular, were studied within the framework of the locally stable approach of A.M. Shurygin, based on the analysis

of the estimator instability functional (L_2 -norm of the influence function), or his approach based on the model of a series of samples with random point contamination (point Bayesian contamination model). In this paper, a compromise family of estimators is studied for which the optimized functional is a convex linear combination of two basic criteria. The compromise family is similar to the conditionally optimal family of estimators proposed by A.M. Shurygin, but the criteria used can be squares of the weighted L_2 -norms of the influence function with arbitrary pre-known weight functions. The considered subject area has remained little-studied to date. In the course of the research, we used a theory we had developed earlier, which describes the properties of estimators that optimize the weighted L_2 -norm of the influence function. As a result of the study, a number of properties of compromise estimators were obtained, and the uniqueness of the family elements was shown. A family member that delivers equal values to the criteria was considered separately: it was shown that this estimator corresponds to the saddle point of the optimized functional, and is also a minimax solution with respect to the basic criteria on the set of all regular score functions. The constructed theory is illustrated using the example of the problem of mathematical expectation estimating of a normal distribution under conditions of targeted malicious influence on a data set (similar to a data poisoning attack in malicious machine learning).

Keywords

M -estimators, robust statistics, influence function, stable estimates, redescending estimators, saddle point, data poisoning

For citation: Lisitsin D.V., Gavrilov K.V. On the properties of compromise M -estimators optimizing weighted L_2 -norm of the influence function. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2025, vol. 25, no. 3, pp. 527–535 (in Russian). doi: 10.17586/2226-1494-2025-25-3-527-535

Введение

Современные подходы к решению задачи оценивания параметров статистических моделей [1] позволяют обеспечить устойчивость получаемых решений к возможным отклонениям реальной ситуации от принятых в модели предположений [2–4]. Среди таких отклонений часто рассматривают *засорение*, под которым будем понимать присутствие в массиве данных посторонних (засоряющих) наблюдений; с теоретической точки зрения в результате получается смесь исходного распределения наблюдений и засоряющего распределения. Одним из наиболее трудных для формализации являлось асимметричное засорение симметрично распределенных данных, поскольку априорная информация о возможном искажении, как правило, отсутствует. Классические робастные подходы в этой ситуации не обладают достаточной устойчивостью. Например, для математического ожидания нормально-го распределения оптимальными оказываются оценка Хьюбера — оценка максимального правдоподобия для плотности, нормальной в центре и экспоненциально убывающей на хвостах, и медиана — оценка максимального правдоподобия для распределения Лапласа, также имеющего экспоненциально убывающую плотность. Однако на практике хвосты могут быть более тяжелыми [3].

Для решения этой проблемы были предложены *сниженные оценки* [3, 4], в которых влияние периферийных наблюдений значительно снижено по сравнению с классическими робастными решениями или равно нулю. Однако все они получены в большей или меньшей степени эвристически.

Между тем наш соотечественник д.т.н. Александр Михайлович Шурыгин предложил довольно удачную *модель байесовского точечного засорения* (БТЗ) [2], которая позволяет описывать окрестность модельного распределения, возникающую в результате асимметричного засорения, и, в частности, рассматривать ситуации целенаправленного вредоносного воздействия (атаки) на набор данных [5].

Модель БТЗ предполагает наличие серии выборок, каждая из которых содержит засорение, сосредоточен-

ное в одной точке. Эта точка является фиксированной в пределах одной выборки, но имеет некоторое распределение по серии выборок. Каждому распределению засоряющей точки соответствует наилучшая оценка, полученная в результате минимизации ее асимптотического квадратичного отклонения [2]. Получающиеся на практике оценки часто оказываются сниженными. В общем случае свойства оценок зависят от соотношения между модельной плотностью и плотностью распределения засоряющей точки (плотностью БТЗ). При неудачном выборе плотности БТЗ оценки могут не быть сниженными и даже вообще устойчивыми — например, когда оба распределения нормальные, но второе имеет меньшую дисперсию [6].

В рамках модели БТЗ являются оптимальными многие известные сниженные оценки параметра сдвига, среди них: бивес-оценка Тьюки, оценки Эндрюса, Бернуллы, Смита, Мешалкина (Уэлша), Фэйра, Рамсея, оценка Хьюбера типа урезанного среднего, обобщенные оценки Шарбонье [2, 5, 7].

Подход на основе БТЗ может быть сведен к анализу весовой L_2 -нормы функции влияния [4], где плотность БТЗ выступает в роли весовой функции. Поскольку последняя не обязательно должна быть нормированной, такая интерпретация удобна тем, что распространяет понятие наилучшей оценочной функции на весовые функции, не являющиеся плотностью, в том числе неинтегрируемые [5].

Таким образом, указанный критерий качества оценивания может служить теоретической основой [8] для конструирования широкого спектра оптимальных устойчивых оценок, имеющих важное практическое значение, в том числе робастных решений, которые характеризуются заданием непараметрической («полноразмерной») окрестности модельного распределения [3, 4]. При этом наиболее продуктивный путь построения оптимальных оценок заключается в разумном выборе весовой функции и последующем нахождении для нее оценочной функции, минимизирующей весовую L_2 -норму функции влияния. В частности, весовая функция может определяться как результат оптимизации самостоятельного критерия, например, некоторой меры неопределенности [5, 7].

Обзор основных результатов

В настоящей работе исследуются свойства *компромиссных оценок*, оптимальных с точки зрения компромиссного критерия в виде выпуклой линейной комбинации двух *базовых критериев* [9]. И компромиссный, и оба базовых критерия представляют собой квадраты весовых L_2 -норм функции влияния [8]. Для базовых критериев весовые функции заранее известны, а их выпуклая линейная комбинация выступает в роли весовой функции компромиссного критерия. Такая формулировка — способ задания семейства, подобного *условно оптимальным оценкам* [2, 10, 11], поскольку оптимизируемый функционал для них тоже может быть записан в виде выпуклой линейной комбинации нормированных функционалов *асимптотической дисперсии и неустойчивости* (квадрата L_2 -нормы функции влияния) [12].

Априорный выбор параметра линейной комбинации на практике может оказаться затруднительным. Важным результатом работы является возможность нахождения оценки, которая бы разрешала такую неопределенность. Указанная оценка соответствует седловой точке оптимизируемого функционала и названа *равнооптимальной* (ОРО). Она также удовлетворяет решению минимаксной задачи относительно базовых критериев.

С помощью компромиссного семейства обычно удается исправить рассмотренную ситуацию, когда используемая плотность БТЗ приводит к неустойчивой оценке. В результате можно перейти к компромиссу между исходным критерием качества оценки и критерием с константной весовой функции. Похожий способ обеспечения устойчивости рассмотрен в работе [13].

Критерий, аналогичный рассматриваемому компромиссному критерию, возникает в случае применения асимптотики сужающихся окрестностей засорения [3, 4, 14] к модели БТЗ. Полученные результаты позволяют решить задачу нахождения минимаксной оценки для такой модели при неизвестном размере окрестности, выражающемся через параметр линейной комбинации (подобно задаче поиска минимаксного размера окрестности в [14]).

Данная работа является продолжением исследований, опубликованных в работе [8]. Помимо изложения теории, в работе приведен пример построения устойчивых оценок параметра сдвига нормального распределения в условиях отвлекающего воздействия на набор данных.

Элементы теории устойчивого оценивания

Пусть x_1, \dots, x_m — наблюдения случайной величины ξ , которая моделирует измеряемую числовую характеристику в некотором эксперименте. Для ее описания используется модельное распределение с плотностью $f(x, \theta)$, заданной в прямоугольнике $X \times \Theta$, где $X \subseteq R$ — множество значений переменной x (это возможные значения измеряемой характеристики), $\Theta \subseteq R$ — множество значений параметра плотности θ и R — множество действительных чисел. M -оценка неизвестного параметра θ может определяться [1, 3] как решение оптимизационной задачи

$$\hat{\theta} = \arg \min_{\theta} \sum_{i=1}^m \rho(x_i, \theta),$$

где $\rho(x_i, \theta): X \times \Theta \rightarrow R$ — некоторая функция, называемая функцией потерь.

Условие равенства нулю производной по θ оптимизируемой функции служит альтернативной (хотя и не эквивалентной [1]) формулировкой задачи M -оценивания:

$$\sum_{i=1}^m \psi(x_i, \hat{\theta}) = 0,$$

называемой *оценочным уравнением*, где функция

$$\psi(x, \theta) = c(\theta) \partial \rho(x, \theta) / \partial \theta \quad (1)$$

называется *оценочной функцией* для параметра θ ; $c(\theta)$ — произвольная непрерывная функция, не равная нулю для всех $\theta \in \Theta$. Множитель $c(\theta)$ задает семейство *эквивалентных оценочных функций*.

Также оценочная функция (1) при модельном распределении должна удовлетворять условию асимптотической несмещенности, которое принимает вид [1, 15]

$$\mathbf{E} \psi(\xi, \theta) = \int_X \psi(x, \theta) f(x, \theta) dx = 0, \quad (2)$$

где \mathbf{E} — оператор математического ожидания по плотности $f(x, \theta)$.

Дифференцируя (2) по θ и допуская возможность изменения порядка дифференцирования и интегрирования, можно записать следующие равенства [3, 15]:

$$\begin{aligned} N(\theta) &= -\lim_{t \rightarrow \theta} \frac{d}{dt} \mathbf{E} \psi(\xi, t) = -\mathbf{E} \frac{\partial}{\partial \theta} \psi(\xi, \theta) = \\ &= \int_X \psi(x, \theta) \frac{\partial}{\partial \theta} f(x, \theta) dx. \end{aligned} \quad (3)$$

Потребуем, чтобы в окрестности истинного значения параметра θ выполнялось условие асимптотической несмещенности (2), были справедливы равенства (3), функция $N(\theta)$ была непрерывной и не равной нулю [10]. Более полный набор условий регулярности, обеспечивающий состоятельность и асимптотическую нормальность оценок, приведен, например в [5, 16]. Специфика случая конечного X подробно рассмотрена в [17].

Асимптотическая дисперсия [1] M -оценки $\hat{\theta}$ определяется выражением

$$V(\psi) = \frac{1}{N^2(\theta)} \int_X \psi^2(x, \theta) f(x, \theta) dx = \mathbf{E} \text{IF}^2(\xi, \theta),$$

где $\text{IF}(x, \theta) = \psi(x, \theta) / N(\theta)$ — функция влияния Хампеля [4]. Функция влияния является показателем количественной робастности и отражает воздействие на оценку бесконечно малого точечного засорения. Величина $\gamma^* = \sup_{x \in X} |\text{IF}(x, \theta)|$ называется *чувствительностью к большой ошибке*, для качественно робастной оценки $\gamma^* < \infty$ [4].

Неустойчивостью оценки [2, 18, 19] называется величина

$$W(\psi) = \frac{1}{N^2(\theta)} \int_X \psi^2(x, \theta) dx = \int_X IF^2(x, \theta) dx.$$

Показатель $W(\psi)$ представляет собой квадрат L_2 -нормы функции влияния, его минимизация приводит к оценке максимальной устойчивости (ОМУ) [2].

Рассмотрим в качестве показателя качества оценивания квадрат весовой L_2 -нормы функции влияния с весом $s(x, \theta)$

$$U(\psi, s) = \frac{1}{N^2(\theta)} \int_X \psi^2(x, \theta) s(x, \theta) dx = \int_X IF^2(x, \theta) s(x, \theta) dx. \quad (4)$$

Такой функционал возникает, например, как асимптотическое квадратичное отклонение оценки в модели БТЗ [2, 5, 9, 16, 19], где весовая функция $s(x, \theta)$ представляет собой плотность распределения засоряющей точки.

В [2] показано, что функционал (4) достигает минимума по ψ на функции

$$\psi_s(x, \theta) = c(\theta) \left[\frac{\partial}{\partial \theta} \ln f(x, \theta) + \beta(\theta) \right] \frac{f(x, \theta)}{s(x, \theta)}, \quad (5)$$

где $c(\theta)$ — функция, имеющая тот же смысл, что и в (1); $\beta(\theta)$ — функция, определяемая из условия (2) путем подстановки в него оценочной функции (5) и решения получившегося уравнения относительно $\beta(\theta)$. Оценочную функцию (5) назовем оптимальной для заданной функции s .

Другие функции, зависящие от оценочной функции, далее отметим нижним индексом « s », если в них подставлена оптимальная оценочная функция (5). Результат (5) по построению справедлив для функций $s(x, \theta) \geq 0$, и далее будут рассматриваться только такие весовые функции. Для краткости записи выражений будем опускать аргументы функций там, где они понятны. Например, (5) можно записать в виде $\psi_s = c(\partial f / \partial \theta + \beta f) / s$. Перед обсуждением свойств компромиссных оценок приведем важные в настоящей работе результаты, доказанные в [8].

Пусть k — некоторый скалярный параметр, от которого зависит весовая функция s . Предположим, что все приведенные далее функции непрерывно зависят от параметра на своей области определения. Штрихом обозначим производную функций по данному параметру. Также для удобства введем обозначение $U(k) = U(\psi_s, s)$, поскольку это функция параметра.

Теорема 1. Для оценочной функции (5) справедливо

$$U'(k) = \frac{1}{N_s^2} \int_X \psi_s^2 s' dx = \int_X IF_s^2 s' dx.$$

Теорема 2. Для оценочной функции (5) справедливо неравенство

$$U''(k) \leq \int_X IF_s^2 s'' dx.$$

Причем равенство в нем достигается, только если функции s и s' не зависят от x . Это следует из усло-

вия коллинеарности функций ϕ и ϕ' [8, формула (6)]. Иными словами, неравенство в теореме 2 обращается в равенство, только когда весовые функции всех членов семейства не зависят от x (т. е. все члены семейства — ОМУ). Данный случай не представляет интерес, поэтому в качестве дополнительного условия регулярности примем, что неравенство в теореме 2 строгое.

Свойства компромиссного семейства оценок

Определим семейство компромиссных оценок как решение, оптимизирующее компромиссный критерий

$$U(\psi, k) = (1 - k)U_1(\psi) + kU_2(\psi), \quad (6)$$

где $U_1(\psi) = U(\psi, s_1)$ и $U_2(\psi) = U(\psi, s_2)$ — базовые критерии качества оценки, которым соответствуют весовые функции $s_1(x, \theta)$ и $s_2(x, \theta)$ (не зависящие от k); k — параметр, задающий семейство и определяющий компромисс между базовыми критериями.

Компромиссный критерий (6) является частным случаем (4) с весовой функцией

$$s = (1 - k)s_1 + ks_2, \quad (7)$$

и, таким образом, результат оптимизации критерия (6) имеет вид (5).

Определение 1. Компромиссной оценкой называется оценка с оценочной функцией вида (5) и весовой функцией (7), причем функции s , s_1 и s_2 должны быть одновременно неотрицательными.

Исходя из данного определения, для функции (7) выбирается диапазон допустимых значений параметра k . Очевидно, этот диапазон обязательно содержит отрезок $[0; 1]$, который традиционно определяет компромисс [9]. Более того, отрезок $[0; 1]$ соответствует парето-оптимальным решениям задачи двухкритериальной оптимизации [20]. Это непосредственно следует из условия оптимальной формулировки семейства, которая сводится к оптимизационной задаче $U_1(\psi) + \lambda U_2(\psi) \rightarrow \min_{\psi}$, где $\lambda \geq 0$ — параметр, задающий семейство. В текущей работе рассматривается расширенное множество значений параметра k , которое может содержать члены $k < 0$ и $k > 1$.

Наряду с $U(k)$, введем обозначения

$$U_1(k) = U_1(\psi_s) = U(\psi_s, s_1), \quad U_2(k) = U_2(\psi_s) = U(\psi_s, s_2).$$

Свойство 1. Применим теорему 1 к семейству оценок с весовой функцией (7):

$$U'(k) = \int_X IF_s^2 s' dx = \int_X IF_s^2 s'(s_2 - s_1) dx = U_2(k) - U_1(k).$$

Свойство 2. Продифференцируем равенство (6) по k и приравняем выражению, определяемому свойством 1. После приведения подобных получим

$$(1 - k)U_1'(k) + kU_2'(k) = 0.$$

Свойство 3. Применим теорему 2 к компромиссным оценкам с весовой функцией (7), учитывая, что $s'' = 0$. Получим $U''(k) < 0$.

Свойство 4. Согласно свойству 3 величина $U'(k)$ является убывающей функцией параметра k .

Свойство 5. Из свойств 1 и 3 следует, что при любом k имеет место $U'_2(k) < U'_1(k)$.

Свойство 6. Из свойств 2 и 5 следует, что $U_1(k)$ и $U_2(k)$ являются монотонными функциями параметра k на каждом из следующих промежутков: $k \leq 0$, $0 \leq k \leq 1$, $k \geq 1$. Рассмотрим эти промежутки подробнее.

Для $k \leq 0$ и $k \geq 1$, согласно свойству 2 функции $U'_1(k)$ и $U'_2(k)$ имеют совпадающие промежутки знакопостоянства и одинаковые знаки на них. Кроме того, данные функции в силу свойства 5 не имеют общих точек, а следовательно, одновременно являются положительными либо отрицательными. Таким образом, функции $U_1(k)$ и $U_2(k)$ возрастают или убывают одновременно. Причем, поскольку функционал (6) минимизируется, функции $U_1(k)$ и $U_2(k)$ являются убывающими при $k \leq 0$ и возрастающими при $k \geq 1$.

Для $0 \leq k \leq 1$, согласно свойству 2, функции $U'_1(k)$ и $U'_2(k)$ имеют совпадающие промежутки знакопостоянства и разные знаки на них. Кроме того, данные функции в силу свойства 5 не имеют общих точек, а следовательно, одна из них положительная, другая отрицательная. Заметим, что $U_1(k)$ достигает минимума, совпадающего с минимумом функционала (6), при $k = 0$, а $U_2(k)$ — аналогично при $k = 1$. Следовательно, функция $U_1(k)$ возрастает, а $U_2(k)$ — убывает.

Следующий результат устанавливает единственность членов компромиссного семейства оценок.

Теорема 3. Решение задачи оптимизации функционала (6) имеет взаимно однозначное соответствие с семейством компромиссных оценочных функций.

Доказательство. При каждом фиксированном k функционалу (6) соответствует выражение (4) с весовой функцией (7), а потому, согласно [2], задача его оптимизации имеет единственное решение вида (5). Нужно показать, что каждому такому оптимальному решению отвечает единственное значение k , определяющее компромисс.

Допустим, двум различным значениям k в выражении (6) отвечает одна оптимальная функция влияния IF_0 . Согласно свойству 1 для данных значений k

$$U'(k) = \int_X IF_0^2 (s_2 - s_1) dx,$$

где правая часть равенства одна и та же, однако левая часть согласно свойству 4 является монотонной функцией от k , следовательно принимает разные значения для разных k . Пришли к противоречию. Теорема доказана.

Равнооптимальная оценка

Ранее для условно оптимального семейства [2, 11] было введено понятие ОРО [10] как члена семейства, который доставляет равные значения относительно характеристикам устойчивости $stb \psi = W_{ОМУ}/W(\psi)$ и эффективности $eff \psi = V_{ОМП}/V(\psi)$, где $W_{ОМУ}$ и $V_{ОМП}$ — неустойчивость ОМУ и асимптотическая дисперсия ОМП. Распространим это понятие на компромиссные оценки.

Определение 2. ОРО называется оценка из компромиссного семейства, удовлетворяющая условию $U_1(k) = U_2(k)$.

Значение параметра k , при котором достигается это равенство, обозначим $k_{ОРО}$. Обратим внимание, что существование ОРО в семействе не гарантируется.

Свойство 7. Если ОРО существует, она единственна и соответствует решению максиминной задачи относительно аргументов функционала (6):

$$k_{ОРО} = \arg \max_k \min_{\psi} U(\psi, k).$$

Соответствующее ОРО значение критерия (6) обозначим $U_{ОРО} = \max_k U(k)$.

Действительно, согласно свойству 1 условие $U_1(k) = U_2(k)$ эквивалентно условию $U'(k) = 0$, которое вместе со свойством 3 образует достаточное условие максимума функции $U(k)$. С учетом свойства 4 условие $U'(k) = 0$ выполняется не более чем для единственного значения k , и оно является допустимым, поскольку функции $U(k)$, $U_1(k)$ и $U_2(k)$ неотрицательные (они равны в этой точке).

Свойство 8. Если ОРО существует, то ее оценочная функция $\psi_{ОРО}$ является решением минимаксной задачи относительно аргументов функционала (6):

$$\psi_{ОРО} = \arg \max_k \min_{\psi} U(\psi, k),$$

причем в точке решения функционал (6) принимает одно и то же значение, равное $U(\psi_{ОРО}, k)$ независимо от значения k .

Действительно, используя (6), можно записать

$$U(\psi, k) = U_1(\psi) + k[U_2(\psi) - U_1(\psi)] \leq U_1(\psi) + |k| |U_2(\psi) - U_1(\psi)|. \quad (8)$$

Второе слагаемое мажорирующего функционала в (8) достигает минимума, равного 0, при условии

$$U_1(\psi) = U_2(\psi) \quad (9)$$

для любого k (т. е. значение k является неопределенным). Полученному условию, в частности, отвечает ОРО, однако, кроме ОРО, это условие могло бы выполняться и для оценочных функций, не принадлежащих компромиссному семейству. Исходя из этого, нужно убедиться, что решение принадлежит данному семейству.

Когда условие (9) выполнено, неравенство в (8) обращается в равенство, в правой части которого остается только первое слагаемое, поэтому задача принимает вид

$$U_1(\psi) = (1 - k)U_1(\psi) + kU_2(\psi) = U(\psi, k) \rightarrow \min_{\psi},$$

ее решение для каждого k по определению принадлежит компромиссному семейству.

Теорема 4. ОРО соответствует седловой точке функционала (6) относительно аргументов ψ и k , при этом имеет место

$$U(\psi_{ОРО}, k) = U(\psi_{ОРО}, k_{ОРО}) \leq U(\psi, k_{ОРО}).$$

Утверждение теоремы 4 следует из свойств 7 и 8.

Свойство 9. Если $0 \leq k_{\text{ОРО}} \leq 1$, то

$$U_{\text{ОРО}} = \min_k \max \{U_1(k), U_2(k)\} = \max_k \min \{U_1(k), U_2(k)\}.$$

Справедливость свойства 9 следует из свойства 6 для отрезка $0 \leq k \leq 1$.

Теорема 5. Если $0 \leq k_{\text{ОРО}} \leq 1$, то

$$\Psi_{\text{ОРО}} = \arg \min_{\Psi} \max \{U_1(\Psi), U_2(\Psi)\}.$$

Доказательство. По условию теоремы 5 значение k , соответствующее решению оптимизационной задачи в свойстве 8, принадлежит отрезку $[0; 1]$, поэтому решение этой задачи не изменится, если множество значений k , на котором производится оптимизация, ограничить указанным отрезком, т. е.

$$\Psi_{\text{ОРО}} = \arg \min_{\Psi} \max_{k \in [0; 1]} U(\Psi, k).$$

С учетом представления (6) очевидно, что минимизируемые функционалы в последнем выражении и в правой части утверждения теоремы 5 совпадают для каждой Ψ :

$$\max \{U_1(\Psi), U_2(\Psi)\} = \max_{k \in [0; 1]} U(\Psi, k).$$

Теорема 5 доказана.

Свойство 10. Если $\min_{\Psi} U_1(\Psi) = \min_{\Psi} U_2(\Psi)$, то $0 < k_{\text{ОРО}} < 1$.

Действительно, по условию имеем $U_1(0) = U_2(1)$. Рассмотрим разность $u(k) = U_2(k) - U_1(k)$. В силу свойства 6 на промежутке $0 \leq k \leq 1$ функция $U_1(k)$ возрастает, а $U_2(k)$ убывает, поэтому

$$u(0) = U_2(0) - U_1(0) = U_2(0) - U_2(1) > 0$$

и

$$u(1) = U_2(1) - U_1(1) = U_1(0) - U_1(1) < 0,$$

причем функция $u(k)$ убывает. Следовательно, значение $k_{\text{ОРО}}$, являющееся решением уравнения $u(k) = 0$, принадлежит интервалу $0 < k < 1$.

Данное свойство, в частности, делает удобным применение базовых критериев вида $\text{stb}_s^{-1}\Psi = U(\Psi, s)/U(\Psi_s, s)$ при некотором U , определяемом (4) [9]. Такие базовые критерии имеют минимальное значение, равное 1.

Обратим внимание, что для практического применения ОРО базовые критерии предварительно следует привести к единой шкале путем нормирования — в данном случае, умножением (делением) на положительные константы. С физической точки зрения оба слагаемых в (6) должны быть безразмерными или иметь одинаковую размерность, что обычно тоже обеспечивается нормированием [21]. Нормирование базовых критериев производится исследователем на основе имеющейся априорной информации и является неформальной процедурой. При недостатке априорной информации исследователь на свое усмотрение может воспользоваться некоторыми эвристиками: например, критерии могут нормироваться к своим минимальным значениям (это

будут функционалы вида $\text{stb}_s^{-1}\Psi$), или могут использоваться функционалы, которым соответствуют весовые функции s_1 и s_2 , являющиеся плотностями.

ОРО представляет собой объективный выбор одного члена компромиссного семейства, и, таким образом, разрешается неопределенность в выборе k . Получившийся компромисс удовлетворяет ряду свойств оптимальности. Так, теорема 4 гарантирует, что значение критерия (6) будет наилучшим при неблагоприятном значении $k = k_{\text{ОРО}}$, а любое другое значение k не ухудшит (правда, и не улучшит) значение этого критерия, т. е. ОРО — так называемое гарантирующее решение. Теорема 5 показывает, что базовые критерии одновременно не могут быть улучшены. Важность теоремы 5 состоит еще и в том, что она позволяет сформулировать ОРО с использованием только базовых критериев, без привлечения компромиссного.

Пример

Рассмотрим задачу оценивания математического ожидания нормального распределения в условиях БТЗ [2], плотность которого нормальная. Пусть модельная плотность имеет математическое ожидание θ и дисперсию σ_0^2 :

$$f(x, \theta) = \frac{1}{\sigma_0 \sqrt{2\pi}} \exp \left[-\frac{(x - \theta)^2}{2\sigma_0^2} \right], \quad (10)$$

а плотность БТЗ — математическое ожидание $\theta + \mu$ и дисперсию σ^2 :

$$s(x, \theta) = \frac{1}{\sigma \sqrt{2\pi}} \exp \left[-\frac{(x - \theta - \mu)^2}{2\sigma^2} \right]. \quad (11)$$

Выбор нормальной плотности БТЗ можно обосновать непараметрическим способом (подобно тому, как это делается в теории робастности) — путем оптимизации некоторых информационных мер при дополнительных ограничениях по аналогии с изложенным в [5]. При этом главным отличием от [5] является необходимость ограничения на величину μ . Например, в соответствии с принципом максимальной энтропии можно максимизировать дифференциальную энтропию плотности $s(x, \theta)$ в классе распределений с фиксированным μ и дисперсией, ограниченной сверху величиной σ^2 .

Случай несовпадающих центров плотностей (10) и (11) интересен тем, что позволяет моделировать целенаправленное вредоносное воздействие (атаку) на набор данных. Обработка данных в подобных условиях все чаще упоминается в работах по робастности, по-видимому, из-за интенсивного развития такого направления, как вредоносное машинное обучение, в котором использование робастных методов считается перспективным [5, 22]. Рассматриваемый в настоящей работе способ воздействия в рамках этого направления называется *атакой отравления данных*. В связи с этим отметим, что некоторые специалисты [22] связывают трудности исследований в данном направлении со сложностью систем глубокого обучения: по их мнению, вначале целесообразно обратиться к изучению более простых моделей, и данный пример можно рассматри-

вать как шаг к реализации этой идеи на основе подхода А.М. Шурыгина.

В данном случае цель злоумышленника — вызвать смещение оценок математического ожидания, а воздействие моделируется точечным засорением как простейшим непараметрическим способом достигнуть этой цели (вообще, точечное засорение оказывается в некотором смысле наихудшим [15]). Статистик, полагая неизвестную ему засоряющую точку случайной, выбирает на основе некоторой априорной информации значения μ и σ^2 , определяющие ее распределение (подобный подход использован в [23]). Так, могут быть выбраны $\mu > 0$ или $\mu < 0$ в зависимости от того, что, по мнению статистика, выгодно злоумышленнику. Но если информации о цели злоумышленника нет, то статистик вынужден положить $\mu = 0$, получая известные оценки Мешалкина [2, 24]. В последнем случае предполагать $\sigma^2 \leq \sigma_0^2$ вряд ли имеет смысл, поскольку тогда в среднем искажение не приведет к ухудшению оценок.

Изучим робастность оптимальной оценки (5) с точки зрения значения γ^* . При $\sigma^2/\sigma_0^2 \leq 2/3$ модель не регулярна, при $2/3 < \sigma^2/\sigma_0^2 \leq 1$ оценкам соответствует неограниченное значение γ^* , но даже при $\sigma^2 > \sigma_0^2$ они могут иметь слишком большое, хотя и конечное значение γ^* .

В соответствии со сделанным анализом, как минимум в случае $\mu \neq 0$, защищенность оценок от воздействия злоумышленника не гарантирует их робастности. Однако робастность полезна в присутствии, наряду со злонамеренным, отличного от него «природного» искажения (а также как страховка от ошибок спецификации плотности БТЗ). Будем считать, что природа безразлична к задаче статистика, однако стремится к увеличению неопределенности, поэтому проявляет себя в равномерном распределении, максимизирующем дифференциальную энтропию [5]. Как отражающий качество оценивания при несобственной равномерной (равной 1) плотности БТЗ может рассматриваться функционал $W(\psi)$ (подробнее в [5]).

Итак, для достижения баланса между двумя аспектами устойчивости — защищенностью и робастностью — воспользуемся компромиссным функциона-

лом (6), в котором базовыми функционалами выберем критерий защищенности $U_1(\psi) = \text{stb}_s^{-1}\psi$, где s — плотность (11), и критерий робастности $U_2(\psi) = \text{stb}^{-1}\psi$. Функционалы $\text{stb}_s\psi$ [9] и $\text{stb}\psi$ [2] принимают значения в диапазоне $[0; 1]$, подобно $\text{eff}\psi$, и удобны в качестве относительных характеристик защищенности и робастности.

Предположим, что степени воздействия злоумышленника и природы статистику неизвестны, и он выбирает ОРО как гарантирующее решение.

На рисунке для $\theta = 0$ и $\sigma_0^2 = 1$ сравниваются функции влияния трех оценок: ОРО, оптимальной оценки для плотности БТЗ (11) и ОМУ (сплошные линии 1, 2 и 3). Также представлены графики плотностей (10) и (11), умноженных на 4 — пунктирные линии 4 и 5. Рисунок *a* соответствует случаю $\mu = 0,82$, $\sigma^2 = 0,6724$ с $k_{\text{ОРО}} \approx 0,249$, $\text{stb}_s\psi_{\text{ОРО}} = \text{stb}\psi_{\text{ОРО}} \approx 0,155$; рисунок *b* — случаю $\mu = 5$, $\sigma^2 = 4$ с $k_{\text{ОРО}} \approx 0,454$, $\text{stb}_s\psi_{\text{ОРО}} = \text{stb}\psi_{\text{ОРО}} \approx 0,665$.

В случае, соответствующем рисунку *a*, оптимальная оценка для плотности (11) неробастна ($\gamma^* = \infty$), поэтому необходимо использовать, например, компромиссную оценку. Однако ОРО имеет хотя и ограниченное, но достаточно большое значение γ^* , модуль функции влияния уже не мал в той области на левой части графика, где еще мала плотность $f(x, \theta)$. Возможно, в этом случае выгодно было бы использовать $k > k_{\text{ОРО}}$. Отметим характерное поведение функций влияния для этих двух оценок — прогиб под центральной областью плотности $s(x, \theta)$ для уменьшения ее влияния. Напротив, функция влияния ОМУ достигает максимального значения недалеко от центра $s(x, \theta)$, поэтому с большой вероятностью может оказаться слишком чувствительной к воздействию злоумышленника.

В случае, соответствующем рисунку *b*, оптимальная оценка для плотности (11) имеет, хотя и ограниченное, но достаточно большое значение γ^* , модуль функции влияния уже не мал в той области на левой части графика, где еще мала плотность $f(x, \theta)$. ОРО там выглядит более предпочтительной — значение γ^* не слишком большое и модуль функции влияния ненамного больше модуля функции влияния ОМУ. Однако ОМУ явно

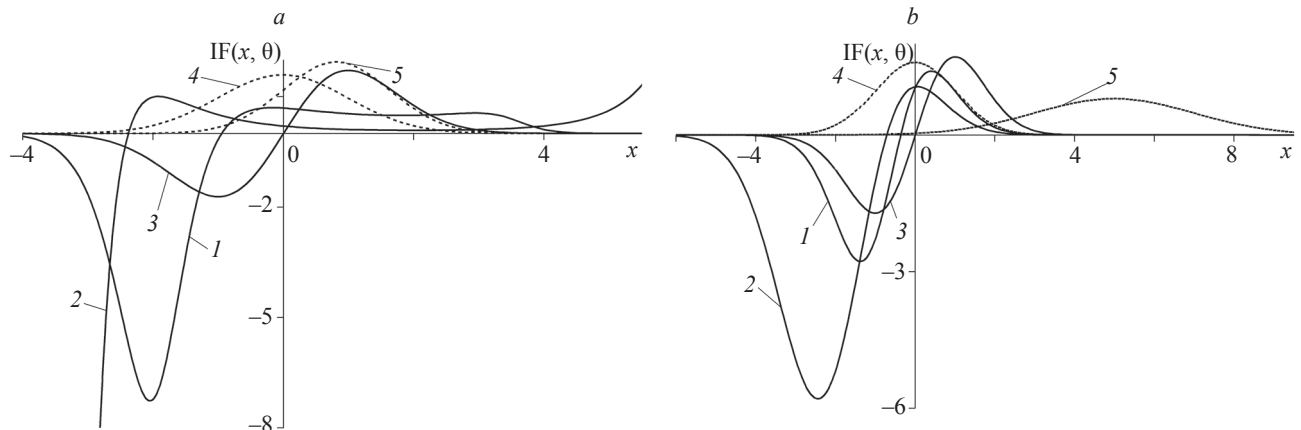


Рисунок. Графики некоторых функций влияния
Figure. Graphs of some influence functions

меньше защищена от искажения, сосредоточенного в правой части графика.

Малое значение функционалов $\text{stb}_s \psi_{\text{ОРО}}$ и $\text{stb} \psi_{\text{ОРО}}$ в случае, соответствующем рисунку, a , говорит о слишком большой потере качества с точки зрения критериев защищенности и робастности. В случае, соответствующем рисунку, b , потеря качества значительна, однако намного меньше. В то же время при $\sigma^2 \rightarrow \infty$ оценочные функции ψ_s и $\psi_{\text{ОРО}}$ стремятся к оценочной функции ОМУ, поэтому значение функционалов $\text{stb}_s \psi_{\text{ОРО}}$ и $\text{stb} \psi_{\text{ОРО}}$ будет стремиться к 1.

Наконец, отметим, что данные построения можно распространить на модель линейной регрессии и нейросетевые модели, включая смеси экспертов [25].

Заключение

В работе проведено теоретическое исследование компромиссного семейства оценок, которое обладает большей общностью, чем предложенное А.М. Шурыгиным семейство условно оптимальных оценок, и отличается способом параметризации. Получены следующие результаты:

- выведен ряд свойств семейства компромиссных оценок;
- показана единственность членов семейства компромиссных оценок;

Литература

1. Боровков А.А. Математическая статистика. Новосибирск: Наука. Сиб. предприятие РАН: Изд-во Ин-та математики, 1997. 771 с.
2. Шурыгин А.М. Прикладная стохастика: робастность, оценивание, прогноз. М.: Финансы и статистика, 2000. 223 с.
3. Huber P., Ronchetti E. *Robust Statistics*. John Wiley & Sons, 2009. 354 p. <https://doi.org/10.1002/9780470434697>
4. Hampel F., Ronchetti E., Rousseeuw P., Stahel W. *Robust Statistics: The Approach Based on Influence Functions*. John Wiley & Sons, 2005. 536 p. <https://doi.org/10.1002/9781118186435>
5. Лисицин Д.В., Гаврилов К.В. Использование принципа максимальной энтропии для конструирования робастных оценок при байесовском точечном засорении. Часть I // Прикладная математика и вопросы управления. 2024. № 1. С. 55–72. <https://doi.org/10.15593/2499-9873/2024.1.04>
6. Лисицин Д.В. Устойчивые методы оценивания параметров статистических моделей. Новосибирск: Изд-во НГТУ, 2013. 76 с.
7. Лисицин Д.В., Гаврилов К.В. Использование принципа максимальной энтропии для конструирования робастных оценок при байесовском точечном засорении. Часть II // Прикладная математика и вопросы управления. 2024. № 2. С. 18–33. <https://doi.org/10.15593/2499-9873/2024.2.02>
8. Лисицин Д.В., Гаврилов К.В. О свойствах M -оценок, оптимизирующих весовую L_2 -норму функции влияния // Научно-технический вестник информационных технологий, механики и оптики. 2024. Т. 24. № 2. С. 267–275. <https://doi.org/10.17586/2226-1494-2024-24-2-267-275>
9. Лисицин Д.В. Устойчивое оценивание параметров модели по многомерным неоднородным неполным данным // Научный вестник НГТУ. 2013. Т. 50. № 1. С. 17–30.
10. Лисицин Д.В., Гаврилов К.В. Об устойчивом оценивании параметров модели при асимметричном засорении данных // Научный вестник НГТУ. 2008. Т. 30. № 1. С. 33–41.
11. Shevlyakov G.L., Oja H. *Robust Correlation: Theory and Applications*. John Wiley & Sons, 2016. 352 p. <https://doi.org/10.1002/9781119264507>
12. Гаврилов К.В., Веретельникова Е.Л. Об одном способе выбора компромисса в семействе условно оптимальных оценок //

- введено понятие равнооптимальной оценки, показано, что ей соответствует седловая точка оптимизируемого функционала;
- показано, что равнооптимальная оценка является минимаксным решением относительно базовых критериев на множестве всех регулярных оценочных функций;
- построенная теория применена к задаче оценивания математического ожидания нормального распределения в условиях целенаправленного вредоносного воздействия на набор данных.

Среди членов компромиссного семейства особую роль играет равнооптимальная оценка. В силу своих свойств оптимальности она может представлять интерес в качестве решения, когда необходим некоторый априорный выбор одной оценки из семейства.

В ближайшие планы входит применение полученных результатов к задаче о минимаксной оценке при неизвестном размере окрестности байесовского точечного засорения в асимптотике сужающихся окрестностей, а также распространение построений из примера на задачу оценивания параметров линейной регрессии. В дальнейшем планируется распространение этих построений и на нейросетевые модели, включая смеси экспертов.

References

1. Borovkov A.A. *Mathematical Statistics*. Amsterdam: Gordon and Breach, 1998, 570 p.
2. Shurygin A. M. *Applied Stochastics: Robustness, Estimation, Prediction*. Moscow, Finansy i statistika Publ., 2000, 223 p. (in Russian)
3. Huber P., Ronchetti E. *Robust Statistics*. John Wiley & Sons, 2009, 354 p. <https://doi.org/10.1002/9780470434697>
4. Hampel F., Ronchetti E., Rousseeuw P., Stahel W. *Robust Statistics: The Approach Based on Influence Functions*. John Wiley & Sons, 2005, 536 p. <https://doi.org/10.1002/9781118186435>
5. Lisitsin D.V., Gavrilov K.V. The use of maximum entropy principle to construct robust estimators under point Bayesian contamination. Part I. *Applied Mathematics and Control Sciences*, 2024, no. 1, pp. 55–72. (in Russian). <https://doi.org/10.15593/2499-9873/2024.1.04>
6. Lisitsin D.V. *Robust Methods for Parameters Estimating of Statistical Models*. Novosibirsk, NSTU Publ., 2013, 76 p. (in Russian)
7. Lisitsin D.V., Gavrilov K.V. The use of maximum entropy principle to construct robust estimators under point Bayesian contamination. Part II. *Applied Mathematics and Control Sciences*, 2024, no. 2, pp. 18–33. (in Russian). <https://doi.org/10.15593/2499-9873/2024.2.02>
8. Lisitsin D.V., Gavrilov K.V. On the properties of M -estimators optimizing weighted L_2 -norm of the influence function. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2024, vol. 24, no. 2, pp. 267–275. (in Russian). <https://doi.org/10.17586/2226-1494-2024-24-2-267-275>
9. Lisitsin D.V. Robust estimation of model parameters in presence of multivariate nonhomogeneous incomplete data. *Science Bulletin of the Novosibirsk State Technical University*, 2013, vol. 50, no. 1, pp. 17–30. (in Russian)
10. Lisitsin D.V., Gavrilov K.V. On stable estimation of models parameters in presence of asymmetric data contamination. *Science Bulletin of the Novosibirsk State Technical University*, 2008, vol. 30, no. 1, pp. 33–41. (in Russian)
11. Shevlyakov G.L., Oja H. *Robust Correlation: Theory and Applications*. John Wiley & Sons, 2016, 352 p. <https://doi.org/10.1002/9781119264507>
12. Gavrilov K.V., Veretelnikova E.L. On one way of choice of compromise in the family of conditionally optimal estimates //

- Вестник Томского государственного университета. Управление, вычислительная техника и информатика. 2024. № 67. С. 60–68. <https://doi.org/10.17223/19988605/67/7>
13. Coretto P., Hennig C. Robust improper maximum likelihood: tuning, computation, and a comparison with other methods for robust Gaussian clustering // *Journal of the American Statistical Association*. 2016. V. 111. N 516. P. 1648–1659. <https://doi.org/10.1080/01621459.2015.1100996>
 14. Rieder H., Kohl M., Ruckdeschel P. The cost of not knowing the radius // *Statistical Methods and Applications*. 2008. V. 17. N 1. P. 13–40. <https://doi.org/10.1007/s10260-007-0047-7>
 15. Смоляк С.А., Титаренко Б.П. Устойчивые методы оценивания: статистическая обработка неоднородных совокупностей. М.: Статистика, 1980. 210 с.
 16. Лисицин Д.В., Гаврилов К.В. Максиминная задача оценивания параметров в условиях байесовского точечного засорения // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. 2023. № 62. С. 56–64. <https://doi.org/10.17223/19988605/62/6>
 17. Лисицин Д.В., Гаврилов К.В. Оценивание параметров распределения ограниченной случайной величины, робастное к нарушению границ // Научный вестник НГТУ. 2016. Т. 63. № 2. С. 70–89. <https://doi.org/10.17212/1814-1196-2016-2-70-89>
 18. Shevlyakov G., Morgenthaler S., Shurygin A. Redescending *M*-estimators // *Journal of Statistical Planning and Inference*. 2008. V. 138. N 10. P. 2906–2917. <https://doi.org/10.1016/j.jspi.2007.11.008>
 19. Shurygin A.M. New approach to optimization of stable estimation // *Proc. of the First US/Japan Conference on the Frontiers of Statistical Modeling: An Informational Approach*. 1994. V. 3. P. 315–340. https://doi.org/10.1007/978-94-011-0854-6_15
 20. Ногин В.Д. Принятие решений в многокритериальной среде: количественный подход. М.: Физматлит, 2005. 176 с.
 21. Подиновский В.В. Идеи и методы теории важности критериев в многокритериальных задачах принятия решений. М.: Наука, 2019. 103 с.
 22. Li F., Lai L., Cui Sh. *Machine Learning Algorithms: Adversarial Robustness in Signal Processing*. Springer, 2022. 104 p. <https://doi.org/10.1007/978-3-031-16375-3>
 23. Großhans M., Sawade C., Brückner M., Scheffer T. Bayesian games for adversarial regression problems // *Proc. of the 30th International Conference on Machine Learning. International Conference on Machine Learning*. 2013, V. 28. P. 55–63. <https://doi.org/10.5555/3042817.3042943>
 24. Айвазян С.А., Енюков И.С., Мешалкин Л.Д. Прикладная статистика: Основы моделирования и первичная обработка данных. М.: Финансы и статистика, 1983. 471 с.
 25. Хайкин С. Нейронные сети: полный курс. М.: Вильямс, 2020. 1104 с.
 12. Gavrilov K.V., Veretel'nikova E.L. On one way to choose a compromise in a family of conditionally optimal estimators. *Tomsk State University Journal of Control and Computer Science*, 2024, no. 67, pp. 60–68. (in Russian). <https://doi.org/10.17223/19988605/67/7>
 13. Coretto P., Hennig C. Robust improper maximum likelihood: tuning, computation, and a comparison with other methods for robust Gaussian clustering. *Journal of the American Statistical Association*, 2016, vol. 111, no. 516, pp. 1648–1659. <https://doi.org/10.1080/01621459.2015.1100996>
 14. Rieder H., Kohl M., Ruckdeschel P. The cost of not knowing the radius. *Statistical Methods and Applications*, 2008, vol. 17, no. 1, pp. 13–40. <https://doi.org/10.1007/s10260-007-0047-7>
 15. Smolyak S.A., Titarenko B.P. *Stable Estimation Methods: Statistical Processing of Heterogeneous Aggregates*. Moscow, Statistika Publ., 1980, 210 p. (in Russian)
 16. Lisitsin D.V., Gavrilov K.V. Maximin problem of parameter estimation in conditions of point Bayesian contamination. *Tomsk State University Journal of Control and Computer Science*, 2023, no. 62, pp. 56–64. (in Russian). <https://doi.org/10.17223/19988605/62/6>
 17. Lisitsin D.V., Gavrilov K.V. Estimation of distribution parameters of a bounded random variable robust to bound disturbance. *Science Bulletin of the Novosibirsk State Technical University*, 2016, vol. 63, no. 2, pp. 70–89. (in Russian). <https://doi.org/10.17212/1814-1196-2016-2-70-89>
 18. Shevlyakov G., Morgenthaler S., Shurygin A. Redescending *M*-estimators. *Journal of Statistical Planning and Inference*, 2008, vol. 138, no. 10, pp. 2906–2917. <https://doi.org/10.1016/j.jspi.2007.11.008>
 19. Shurygin A.M. New approach to optimization of stable estimation. *Proc. of the First US/Japan Conference on the Frontiers of Statistical Modeling: An Informational Approach*, 1994, vol. 3, pp. 315–340. https://doi.org/10.1007/978-94-011-0854-6_15
 20. Nogin V.D. *Decision Making in Multicriteria Environment: a Quantitative Approach*. Moscow, Fizmatlit Publ., 2004, 176 p. (in Russian)
 21. Podinovskiy V.V. *Ideas and Methods of Criteria Importance Theory in Multicriteria Decision-Making Problems*. Moscow, Nauka Publ., 2019, 103 p. (in Russian)
 22. Li F., Lai L., Cui Sh. *Machine Learning Algorithms: Adversarial Robustness in Signal Processing*. Springer, 2022, 104 p. <https://doi.org/10.1007/978-3-031-16375-3>
 23. Großhans M., Sawade C., Brückner M., Scheffer T. Bayesian games for adversarial regression problems. *Proc. of the 30th International Conference on Machine Learning. International Conference on Machine Learning*, 2013, vol. 28, pp. 55–63. <https://doi.org/10.5555/3042817.3042943>
 24. Aivazyazyan S.A., Yenyukov I.S., Meshalkin, L.D. *Applied Statistics: Bases of Modelling and Initial Data Processing*. Moscow, Finansy i statistika, Publ., 1983, 471 p. (in Russian)
 25. Haykin S. *Neural Networks: A Comprehensive Foundation*. Prentice Hall, 1998, 842 p.

Авторы

Лисицин Даниил Валерьевич — доктор технических наук, профессор, Новосибирский государственный технический университет, Новосибирск, 630073, Российская Федерация, [sc 57199594779](https://orcid.org/0009-0007-0723-6285), <https://orcid.org/0009-0007-0723-6285>, lisitsin@ami.nstu.ru

Гаврилов Константин Викторович — кандидат технических наук, доцент, Новосибирский государственный технический университет, Новосибирск, 630073, Российская Федерация, [sc 58966384700](https://orcid.org/0009-0001-9252-9942), <https://orcid.org/0009-0001-9252-9942>, aenigma77@mail.ru

Authors

Daniil V. Lisitsin — D.Sc., Full Professor, Novosibirsk State Technical University, Novosibirsk, 630073, Russian Federation, [sc 57199594779](https://orcid.org/0009-0007-0723-6285), <https://orcid.org/0009-0007-0723-6285>, lisitsin@ami.nstu.ru

Konstantin V. Gavrilov — PhD, Associate Professor, Novosibirsk State Technical University, Novosibirsk, 630073, Russian Federation, [sc 58966384700](https://orcid.org/0009-0001-9252-9942), <https://orcid.org/0009-0001-9252-9942>, aenigma77@mail.ru

Статья поступила в редакцию 25.02.2025
Одобрена после рецензирования 08.04.2025
Принята к печати 25.05.2025

Received 25.02.2025
Approved after reviewing 08.04.2025
Accepted 25.05.2025



Работа доступна по лицензии
Creative Commons
«Attribution-NonCommercial»

doi: 10.17586/2226-1494-2025-25-3-536-544

УДК 004.942

Комбинированный подход к детектированию неисправностей в сложных технических системах на основе модели бонд-графа

Валентин Алексеевич Дмитриев¹, Мария Яковлевна Марусина²✉

¹ Ассоциация «Цифровые технологии в промышленности», Санкт-Петербург, 191181, Российская Федерация

^{1,2} Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация

¹ valentinswsu@mail.ru, <https://orcid.org/0009-0007-5568-498X>

² myamarusina@itmo.ru✉, <https://orcid.org/0000-0002-5611-8812>

Аннотация

Введение. Разработан и исследован новый подход к детектированию неисправностей для сложных технических систем, позволяющий определить и классифицировать одиночные и множественные одновременные неисправности. Решена задача надежной и своевременной идентификации как одиночных, так и множественных одновременных неисправностей в условиях ограниченного доступа к размеченным данным. Угроза безопасной работе автономной техники является типичной проблемой для полевых условий эксплуатации, в которых традиционные подходы, основанные на моделях или данных, применяемых по отдельности, оказываются неэффективными. **Метод.** В работе представлен комбинированный подход к детектированию неисправностей. Предложенное решение сочетает в себе аналитическую модель бонд-графа и сверточную нейронную сеть. Бонд-граф формирует остатки — разницу между значениями, рассчитанными по физическим законам системы, и измерениями с датчиков. Остатки анализируются сверточной нейронной сетью, которая обучается выявлять и классифицировать неисправности по их характерным особенностям. Для учета неопределенности параметров (например, сопротивления или емкости) использовано линейное дробное преобразование. Такой подход позволяет объединить априорные знания о физике системы с возможностями глубокого обучения. **Основные результаты.** Эффективность подхода оценена на симуляторе гидравлической системы управления поворотным механизмом автономной техники. В симуляцию добавлен гауссовский шум для имитации реальных условий. Эксперименты включали начальные, ступенчатые, одиночные и множественные неисправности. Тесты с разными объемами обучающих данных при выборке менее 128 образцов показали более высокую эффективность предложенного комбинированного подхода по сравнению с классическими методами машинного обучения, такими как Random Forest или K-Nearest Neighbors). Предложено решение для детектирования неисправностей в гидравлических системах управления автономной техникой. **Обсуждение.** Разработанный подход особенно эффективен при ограниченных данных, что делает его пригодным для полевых условий. Он позволяет своевременно выявлять и классифицировать неисправности (например, утечки в клапанах или сбой электромагнитных задвижек), что снижает риск отказов и обеспечивает безопасность автономной техники. Результаты могут быть адаптированы и внедрены для электрических, механических и прочих сложных технических систем.

Ключевые слова

детектирование неисправностей, бонд-граф, сверточная нейронная сеть, линейное дробное преобразование, технические системы

Ссылка для цитирования: Дмитриев В.А., Марусина М.Я. Комбинированный подход к детектированию неисправностей в сложных технических системах на основе модели бонд-графа // Научно-технический вестник информационных технологий, механики и оптики. 2025. Т. 25, № 3. С. 536–544. doi: 10.17586/2226-1494-2025-25-3-536-544

Combined approach to fault detection in complex technical systems based on bond-graph model

Valentin A. Dmitriev¹, Maria Ya. Marusina²✉

¹ Association of Digital Technologies in Industry, Saint Petersburg, 191181, Russian Federation

^{1,2} ITMO University, Saint Petersburg, 197101, Russian Federation

¹ valentinswsu@mail.ru, <https://orcid.org/0009-0007-5568-498X>

² myamarusina@itmo.ru✉, <https://orcid.org/0000-0002-5611-8812>

Abstract

A new fault detection approach for complex technical systems has been developed and investigated, enabling the identification and classification of single and multiple simultaneous faults. The challenge of reliable and timely identification of both single and multiple simultaneous faults under conditions of limited access to labeled data has been addressed. A threat to the safe operation of autonomous equipment is a common challenge in the field operating conditions where traditional model-based or data-driven approaches, used individually, prove to be ineffective. This work presents a hybrid approach to fault detection. The proposed solution combines an analytical bond-graph model and a Convolutional Neural Network (CNN). The bond-graph generates residuals — the difference between values calculated based on the system physical laws and sensor measurements. The residuals are then analyzed by the CNN which is trained to detect and classify faults based on their characteristic features. Linear Fractional Transformation is employed to account for parameter uncertainties (e.g., resistance or capacitance). This approach allows combining a priori knowledge of the system physics with the capabilities of deep learning. The effectiveness of the approach was evaluated on a simulator of a hydraulic steering control system for autonomous equipment. Gaussian noise was added to the simulation to simulate real-world conditions. The experiments included incipient, abrupt, single, and multiple faults. Tests with varying amounts of training data, using sample sizes less than 128, demonstrated the higher effectiveness of the proposed hybrid approach compared to classical machine learning methods (such as Random Forest or K-Nearest Neighbors). A solution is proposed for fault detection in hydraulic control systems of autonomous equipment. The developed approach is particularly effective with limited data, making it suitable for field conditions. It allows for timely detection and classification of faults (e.g., valve leaks or solenoid valve failures), which reduces the risk of failures and ensures the safety of autonomous equipment. The results can be adapted and implemented for electrical, mechanical, and other complex technical systems.

Keywords

fault detection, bond-graph, convolutional neural network, linear fractional transformation, technical systems

For citation: Dmitriev V.A., Marusina M.Ya. Combined approach to fault detection in complex technical systems based on bond-graph model. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2025, vol. 25, no. 3, pp. 536–544 (in Russian). doi: 10.17586/2226-1494-2025-25-3-536-544

Введение

Целью данной работы является создание надежного и эффективного подхода детектирования неисправностей (Fault Detection and Isolation, FDI), способного преодолеть ограничения существующих подходов для сложных технических систем управления, применяемых в автономной технике. При этом неисправностью принято считать состояние объекта, в котором хотя бы один параметр объекта не соответствует хотя бы одному из требований, установленных в документации на этот объект¹. Существующие методы детектирования, основанные на физическом моделировании [1–5], требуют точных аналитических моделей, построение которых затруднительно. Более того, они часто неспособны адекватно учитывать неопределенности параметров и внешние возмущения, приводящие к ложным срабатываниям. В свою очередь, методы, основанные на данных [6], такие как нейронные сети, требуют значительных объемов размеченных данных, что не всегда выполнимо в полевых условиях. В контексте автономной техники [7, 8], где надежность является критическим фактором, применение как физических моделей,

так и моделей, основанных на данных, в чистом виде часто оказывается недостаточным для решения задачи раннего обнаружения и классификации неисправностей [9], особенно в случаях одновременного возникновения нескольких неисправностей, или при развитии неисправности в начальной стадии. Для решения этой задачи предлагается новый комбинированный подход, который сочетает преимущества как физических методов, так и методов, основанных на данных [10–12]. Разработанный подход включает построение бонд-графа (Bond-Graph, BG) [13] системы для генерации остатков разницы ожидаемых и фактических значений в системе (далее — остатков), а также модели сверточных нейронных сетей (Convolutional Neural Network, CNN) для детектирования и классификации неисправностей. В работе дано общее описание задачи, представлена схема типового элемента гидравлической системы управления, основы построения модели BG и полная модель BG гидравлической системы управления. Показана применимость методов, основанных на модели BG (через построение таблицы аналитических избыточных отношений (Analytical Redundancy Relations, ARR), а также применимость методов, основанных на данных. Представлена схема работы сформированного комбинированного подхода (BG-CNN) к детектированию неисправностей и пример его работы. Далее проведено сравнение эффективности предложен-

¹ ГОСТ Р 27.102-2021 Надежность в технике. Надежность объекта. Термины и определения. Применяется с 01.01.2022 взамен ГОСТ 27.002-2015.

ного подхода с методом без предварительного выделения остаточных данных, а также продемонстрирована способность разработанного подхода работать как с одиночными, так и с множественными одновременными неисправностями.

Постановка задачи

Основная цель алгоритмов FDI состоит в том, чтобы обнаруживать, отслеживать и анализировать неисправности, возникающие в системе. Эти алгоритмы не только определяют наличие неисправности, но и оценивают степень ухудшения производительности системы и выявляют коренные причины возникших неисправностей. Достигается это посредством непрерывного мониторинга изменений в физических свойствах системы, которые проявляются в виде наблюдаемых явлений [14]. FDI традиционно делятся на три категории: методы, основанные на физических моделях системы; методы, использующие данные, собранные с системы; а также комбинированные и гибридные методы, сочетающие оба подхода. Далее будет дано обоснование выбора метода линейного дробного преобразования (Linear Fractional Transformation, LFT-BG) среди методов, базирующихся на физических моделях, и подхода глубокого обучения среди методов, основанных на данных. В качестве примера применения рассмотрена упрощенная схема гидравлической системы управления (рис. 1).

При моделировании BG система рассматривается как единый блок (рис. 1, *b*), состоящий из инерциального элемента (I), диссипативного элемента (R) и емкостного элемента (C). Входные данные внешних для системы источников могут выражаться как элементы источника усилия (S_e), а в текущем случае — источника потока (S_f). Для многопортовых элементов применяются блоки трансформатора (TF), связывающего величины одного и того же типа с входом и выходом, а именно преобразует поток в поток и усилие в усилие, а в приведенном примере использовался гиратор в модульном исполнении (MGY), связывающего поток с одной стороны с усилием с другой [14]. Модульность гиратора указывает на изменяемое значение его коэффициента. Система ограничений и связей представляется в виде 0 и 1, где узел типа 0 показывает общее усилие и, как следствие, только одну входящую причинную связь, а узел типа 1 — напротив, общий поток

и одну выходящую причинную связь. При этом усилия на каждой связи, действующей на 0-переход, равны, а потоки на данных связях в сумме равны 0. Связи на 1-переходе имеют противоположные взаимосвязи потоков и усилий, усилия в этом случае в сумме равны 0, а потоки, равны между собой. Также стоит отметить, что символы 0 и 1 выбраны как нейтральные по отношению к физической области модели. Также связи дополнительно пронумерованы от «2» до «9», начиная с источника потока, заканчивая связью «9» на диссипативном элементе рассеивания энергии в накопителе (R_u). Зная данные равенства потоков и усилий, возможно сформировать модель BG (рис. 1, *b*) гидравлической системы, представленной на рис. 1, *a*.

На рис. 1, *a* переменные представлены следующими обозначениями: p_A — исходное давления на начальном участке гидравлической системы (условной точке «А»); T — крутящий момент двигателя, а ω — угловая скорость; условное обозначение « w » представляет массовый расход компрессора, w_1 — на входе и w_2 — на выходе компрессора соответственно, индексы давления аналогичны — p_1 и p_2 ; p_c — повышающее давление в компрессоре, а p_d — результирующее давление на выходе из компрессора; p_p — давление в накопителе, V_p — объем накопителя; p_B — выходное давление элемента гидравлической системы управления.

В подходах, основанных на физике, существует различие между количественными и качественными методами, основанное на объеме априорных знаний, доступных о системе. Качественные методы используют внутреннюю структуру модели системы, причинно-следственные связи и отношения, основанные на правилах, для формулирования диагностических выводов. Эти выводы затем используются для передачи информации, связанной с неисправностями, потенциальным кандидатам на диагностику. Качественные методы анализа систем часто основываются на графических представлениях, таких как диаграфы (с ребрами, представляющими причинные связи), двудольные графы и графы связей [2]. Структура этих графовых моделей является гибкой и допускает описание различных типов взаимосвязей. Свойства графа, в частности связность, могут быть использованы для определения отслеживаемости, то есть возможности мониторинга отдельных частей системы. Также формальные определения структурной наблюдаемости и управляемости системы могут

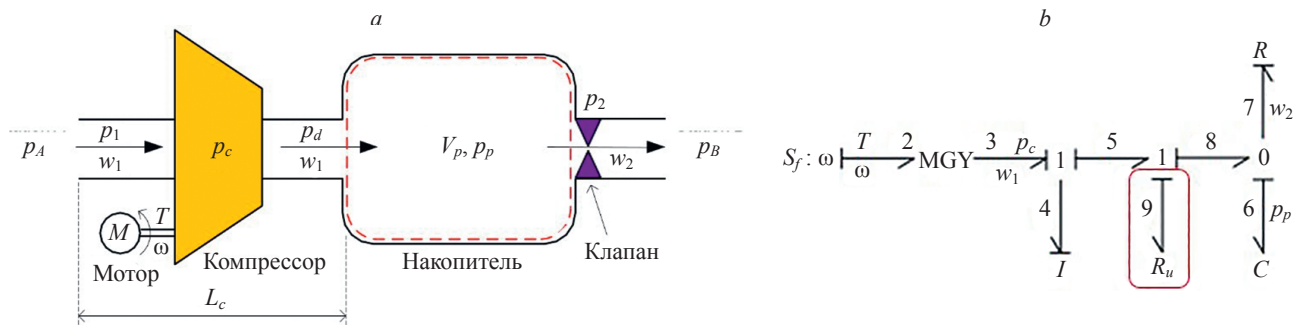


Рис. 1. Схема элемента гидравлической системы управления: функциональная схема (*a*); соответствующая модель бонд-графа (*b*)

Fig. 1. Schematic diagram of the hydraulic control system element: functional diagram (*a*); corresponding bond-graph model (*b*)

быть сформулированы на основе анализа графа. Тем не менее, ключевым ограничением этих методов является их качественный характер, что затрудняет их применение для идентификации и классификации разнообразных неисправностей. В отличие от качественных методов, количественные методы предлагают более точные способы обнаружения неисправностей. Эти методы используют математические модели, описывающие систему через отношения между входами и выходами, и включают в себя такие методы, как наблюдатели, анализ пространства четности и оценка параметров. Рассмотрим каждый из этих методов.

Метод наблюдателей предусматривает сравнение результатов измерений процесса с их оценками, полученными на основе модели системы, для формирования остатка [3]. С целью повышения робастности к шуму и неопределенности используются множественные наблюдатели, чувствительные к различным типам неисправностей [4]. При этом методы, основанные на наблюдателях, сталкиваются со сложностями в определении местоположения неисправности в рамках модели, а также не всегда обеспечивают сходимость для нелинейных процессов. Метод оценки параметров основывается на использовании методов системной идентификации для вычисления параметров модели на основе измеренных входных и выходных данных. Далее полученные параметры сравниваются с эталонными значениями, что позволяет обнаруживать неисправности. Недостатком подхода является зависимость от постоянного изменения входных данных системы, что может создавать проблемы при анализе стационарных процессов.

Проанализировав существующие методы FDI, основанные на физике, можно сделать вывод о том, что BG является заслуживающим внимания вариантом из-за присущих ему преимуществ, вытекающих как из качественных, так и из количественных аспектов. Взаимодействие между узлами в BG описывается с использованием двух обобщенных переменных: усилие — $e(t)$ (в физических понятиях это может быть — сила, напряжение, давление) и поток — $f(t)$ (скорость, ток, расход) [15]. На рис. 2 показано определение направления мощности для элемента A, посылающего энергию на B, и связанного с этим усилия, показанного над полустрелкой, а также потока — по определению, под полустрелкой. Такое обозначение принято называть «связь». При этом причинно-следственную связь усилия и потока графически обозначают вертикальной чертой. Если черта находится перед полустрелкой (связи 2, 5, 6, 8 и 9, рис. 1, b), то мощность направлена от потока к усилию, а в случае расположения после полустрелки — от усилия к потоку.

Применение методов, основанных на физической модели системы

Модель BG формируется с помощью физических законов, описывающих динамику гидравлической системы управления, а также на основе понимания взаимосвязей между компонентами системы (рис. 1, b). Добавляя к модели элементы неисправностей, как

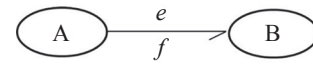


Рис. 2. Направление мощности бонд-графа и связанные с ним определения усилия и потока: элемент «B» получает энергию от «A»

Fig. 2. Bond-graph power direction and associated definitions of force and flow: element “B” receives power from “A”

например элемент Ru, отражающий сопротивление дросселя (связь 9, рис. 1, b), можно создать диагностическую модель бонд-графа (Diagnostics Bond-Graph, DBG). DBG позволяет использовать количественные FDI путем создания аналитических избыточных отношений (ARR) за счет использования подхода, основанного на анализе причинно-следственного пути [2]. ARR — математические выражения, которые связывают измеряемые значения с ожидаемыми, рассчитанными на основе модели.

ARR могут быть записаны в терминах только известных для системы значений

$$ARR = \{U, S, \vartheta\},$$

где U — набор входных данных для системы (известный); S — измерения датчика (известные); ϑ — набор значений параметров (известный). Для реализации ARR были выполнены следующие этапы: создана модель BG системы; преобразована DBG путем дуализации датчиков (определения местоположения датчика); выведены ARR, исключая неизвестные переменные. ARR-модель выдает числовую оценку, которая называется остаточным сигналом (r_i). Набор этих сигналов ($r = \{r_1, r_2, \dots, r_q\}$) показывает, как работает система. В идеале эти остатки близки к 0, когда система функционирует нормально, и заметно отличаются от 0 при возникновении неисправностей. Для оценки применимости подхода на основе модели BG, строится матрица сигнатур неисправностей (Fault Signature Matrices, FSM), представленная в виде таблицы. Строки FSM соответствуют компонентам системы (например, насос, клапан, датчик), а столбцы соответствуют ARR. Элементы FSM (γ_{ij}) являются булевыми переменными (0 или 1), указывающими на чувствительность соответствующего остатка ARR_j к компоненту системы E_m ($\gamma = 1$ — если ARR_j чувствителен к компоненту E_m ; $\gamma = 0$ — если не чувствителен). Сигнатура неисправности компонента E_i хранится в строке FSM. Если вектор остатков совпадает с сигнатурой компонента, то можно локализовать неисправность в этом компоненте.

Для схемы на рис. 1 исследование проводилось относительно 8-ми возможных неисправностей: давление штока; силы, действующие на поршень; сопротивление поршня; гидравлическое сопротивление накопителя; гидравлическое сопротивление клапана; сопротивление утечки и внешние силы. Результаты пространства остатков и датчиков, описанных неисправностей, представлены в разделе «Комбинированный подход».

Ключевое преимущество BG заключается в его применимости к нелинейным системам. Однако метод имеет ограничение: предполагается, что параметры

Таблица. Структурный анализ матрицы признаков неисправностей (FSM) в частном и общем видах (сокращенный пример заполнения)

Table. Structural analysis of the fault sign matrix (FSM) in private and general form

Частный вид матрицы			Общий вид матрицы			
Наименования строк	Наименования столбцов		Наименования строк	Наименования столбцов		
	ARR_1	ARR_2		ARR_1	...	ARR_j
p	1	0	E_1	γ_{11}	...	γ_{1j}
R	1	1
F	1	0	E_m	γ_{m1}	...	γ_{mj}

Примечание. p — давление штока; R — гидравлическое сопротивление; F — силы, действующие на поршень.

модели (например, сопротивление клапанов, емкость накопителя, инерция жидкости) известны точно и остаются постоянными. Однако в реальных системах эти параметры подвержены неопределенностям. Для решения этой проблемы используется метод линейного BG с дробным преобразованием (LFT-BG), где параметры моделируются с учетом неопределенностей. LFT-BG расширяет классический BG, моделируя параметры в виде:

$$\theta_i = \theta_{in}(1 + \delta_i), \delta_i \in [-\Delta, \Delta],$$

где Δ — максимальная относительная неопределенность; θ_i — параметры системы; θ_{in} — номинальное значение параметров системы; δ_i — номинальный коэффициент неопределенности системы.

Применение методов, основанных на данных

Для обработки остаточных сигналов и выявления закономерностей неисправностей были использованы CNN [16–18]. Несмотря на то, что остаточные сигналы (r_1, r_2) представляют собой временные ряды, выбор CNN обусловлен рядом факторов. Для выявления пространственных паттернов в данных остатки преобразуются в спектрограммы с использованием кратковременного преобразования Фурье. При этом преимущество CNN заключается в том, что спектрограммы имеют структуру, аналогичную изображениям, где сверточные слои эффективно выделяют локальные особенности (например, резкие изменения амплитуды, характерные для утечек). Кроме этого CNN, как архитектура глубокого обучения, обладает способностью самостоятельно извлекать признаки из данных. Однако, чтобы корректно оценить преимущества CNN перед другими, более простыми методами машинного обучения, сравнение проведено в равных условиях. Это означает, что все сравниваемые модели использовали одни и те же входные данные, а именно — остатки, сгенерированные моделью BG. Использование остатков является важным этапом предобработки, который позволяет учесть физику процесса и повысить качество детектирования. Сравнение CNN с другими моделями, примененными к «сырым» данным, было бы некорректным, поскольку CNN изначально предназначена для работы со структурированными данными, такими как изображения или спектрограммы остатков. Основной принцип работы CNN заключается в применении операций свертки

к входным данным, где каждый слой CNN преобразует вход (X), создавая выходные признаки (Z_k).

$$Z_k = \sigma(W_k * X + B_k),$$

где «*» — оператор 2D-свертки; W_k — веса, связанные со сверточным ядром; B_k — элемент смещения; $\sigma(x)$ — функция активации: $\sigma(x) = \max(0, x)$; k — обозначение число ядер.

Для уменьшения размерности и фильтрации шума используются слои максимального пулинга (maxpooling). Данная операция выделяет наиболее значимые признаки, создавая карту признаков $Z_{k_{pooled}}$ с использованием окна $s \times s$:

$$Z_{k_{pooled}} = \text{maxpool}(Z_k);$$

$$Z_{k_{pooled}}(i, j) = \max(Z_k[i: i + s, j: j + s]).$$

Полученная карта признаков после пулинга преобразуется в одномерный вектор с помощью слоя сглаживания. Это $Z_{Flatten}$, использующийся полносвязной сетью для классификации неисправностей. Если карта признаков $Z_{k_{pooled}}$ имеет H_p столбцов и W_p строк, то вектор $Z_{Flatten}$ формируется следующим образом:

$$Z_{Flatten} = [Z_{k_{pooled}}(1, 1), Z_{k_{pooled}}(1, 2), \dots, Z_{k_{pooled}}(H_p, 1), Z_{k_{pooled}}(H_p, 2), \dots, Z_{k_{pooled}}(H_p, W_p)].$$

Затем $Z_{Flatten}$ подается на вход полносвязной сети, которая имеет M нейронов на выходном слое, где M — количество классов неисправностей. Выход сети определяется с помощью функции softmax:

$$\hat{Y} = \text{softmax}(\Phi Z_{Flatten} + \mathbf{b}),$$

$$\text{softmax}(\mathbf{x}_i) = \frac{e^{x_i}}{\sum_{j=1}^M e^{x_j}},$$

где Φ — веса; \mathbf{b} — смещение полносвязной сети; e^{x_i} — экспоненциальное значение i -го элемента вектора \mathbf{x} , а e^{x_j} — аналогичное значение j -го элемента, используемое для нормализации. Функция softmax преобразует выходы полносвязной сети в нормализованное распределение вероятностей, облегчая интерпретацию результатов, при этом полносвязная сеть используется для классификации признаков, извлеченных сверточными слоями.

Процесс обучения CNN заключается в настройке параметров сети (весов и смещений) путем минимизации функции потерь:

$$Loss = -\frac{1}{n1} \sum_{i=1}^{n1} \sum_{j=1}^M Y_{ij} \log(\hat{Y}_{ij}),$$

где $n1$ — число обучающих выборок; Y_{ij} и \hat{Y}_{ij} — целевые и предсказанные значения для i -ой выборки и j -ого класса. В этом процессе используются обучающие данные X_{train} с соответствующими метками Y_{train} . Для тестирования используются тестовые данные X_{test} . Для задачи мультиклассовой классификации применяется метрика F1-score, как наиболее сбалансированная метрика, учитывающая как точность, так и полноту.

Таким образом, для обучения CNN используется набор остаточных сигналов, которые формируются из разработанной модели BG, проходя через процедуру свертки, пулинга, и сглаживания для создания входных данных для классификатора на основе полносвязной сети. Это позволяет применить возможности глубокого обучения для эффективного обнаружения и классификации неисправностей на основе чувствительных к неисправностям остаточных сигналов.

Комбинированный подход

Как показано на рис. 3, предложенный комбинированный подход BG-CNN состоит из блока генерации остатков BG и блока классификации неисправностей на основе CNN. Чувствительные к неисправностям остатки генерируются из модели BG [19] и используются в качестве входных данных для модели CNN. Этот подход имеет две фазы: во-первых, помеченные сигналы остатков неисправностей используются для автономного обучения CNN. Во второй фазе сигналы остатков в режиме реального времени, генерируемые моделью BG, используются обученной моделью CNN для классификации типов неисправности.

Схема работы предложенного подхода состоит из трех основных этапов.

Этап 1. Строится модель гидравлической системы с использованием BG. В полученной модели учитываются

неопределенности параметров с помощью LFT-BG. Модель позволяет рассчитать *остатки* — разницу между ожидаемыми значениями, рассчитанными моделью, и измеренными значениями, полученными с датчиков. В системе генерируется два остатка: r_1 , связанный с электрической частью, и r_2 — гидравлической частью. Далее остаточные сигналы r_1 и r_2 используются для обучения CNN. Каждый остаток, представленный в виде одномерного временного ряда $X \in R^{(N \times 100 \times 2)}$, с которым потом и работает CNN. Преобразование происходит за счет предобработки данных с целью устранения несоответствия в масштабах сигналов от датчиков и актуаторов, а также минимизации влияния шумов на работу модели.

Этап 2. Проводится предобработка в три шага: нормализация данных, фильтрация шумов, а также сегментация временных рядов. Нормализация данных представлена следующим решением:

$$X_{norm} = \frac{X - \mu}{\sigma},$$

где μ — среднее значение параметра; σ — стандартное отклонение по обучающей выборке. Фильтрация шумов выполняется посредством применения фильтра Калмана, а сегментация временных рядов происходит в два этапа: для начала сигналы разбиваются на окна длиной 100 отсчетов, что соответствует 100 мс при выбранной частоте дискретизации в 1 кГц, далее сформированные окна перекрываются на 50 %, что позволяет увеличить размер обучающей выборки. Длительность измерений и интервал дискретизации выбраны на основе анализа спектральной плотности мощности для сигналов датчиков и остатков. Выбранные значения обеспечивают достаточное разрешение для обнаружения характерных частот, связанных с неисправностями, и при этом не приводят к избыточному объему данных.

В процессе обучения CNN изучает взаимосвязь между этими остаточными сигналами и различными типами неисправностей (включая их одиночные и множественные проявления). В рассмотренном случае количество классов неисправностей для одиночных проявлений равно 6, и для множественных — 8. После

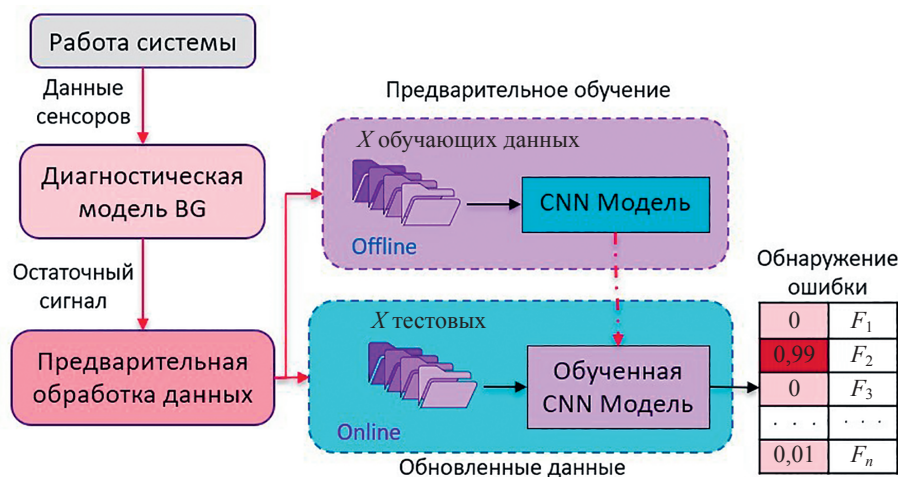


Рис. 3. Структурная схема комбинированного подхода

Fig. 3. Structure diagram of the combined approach

обучения CNN может использоваться в режиме реального времени.

Этап 3. Выполняется обнаружение ошибки. Когда новые данные поступают от датчиков, модель BG генерирует остаточные сигналы, которые передаются обученной CNN, и она определяет, присутствует ли какая-либо неисправность, и если да, то какая.

Таким образом, предложенный подход сочетает в себе преимущества использования физической модели для получения устойчивых к шумам остатков и возможности машинного обучения для извлечения скрытых закономерностей и классификации сложных нелинейностей в данных, что позволяет добиться лучшей точности и надежности детектирования неисправностей в гидравлических системах. Для сравнения на двух диаграммах (рис. 4) представлены результаты подхода FDI на остаточных данных и на данных с датчиков, где наглядно видно, что с применением сформированного подхода детектирование неисправностей существенно упрощается. На рис. 4, *a* ось « ARR_1 » представляет значения параметров в приведенном масштабе, а ось « ARR_2 » — значения пространства остатков, на рис. 4, *b* представлены нормализованные данные пространства датчиков относительно времени. Рассматривался следующий список параметров неисправностей: Healthy — случайный набор проверочных данных, параметры остатков которых приближены к 0; I/W — силы, действующие на поршень; I — внешние силы; K — сопротивление утечки; Re/Rm — результирующее сопротивление; Re — гидравлическое сопротивление накопителя; Rm — гидравлическое сопротивление клапана; W — сопротивление поршня.

Обсуждение

Для оценки эффективности предложенного комбинированного подхода, было проведено его сравнение с методами, основанными только на BG или только на CNN, а также с традиционными методами машинного обучения (ML), такими как Random Forest (RF),

K-Nearest Neighbors (KNN), Artificial Neural Networks (ANN) и метод опорных векторов (Support Vector Machines или просто SVM). Важно отметить, что все сравниваемые методы, включая RF, KNN, ANN и SVM, использовали в качестве входных данных те же остатки, сгенерированные моделью BG, что и предложенный подход (BG-CNN). Это обеспечивает честное и объективное сравнение, поскольку все модели работают с одинаковой предобработанной информацией, учитывая физику системы. Проведена оптимизация гиперпараметров как предложенного подхода, так и методов для сравнения. Использовался метод случайного поиска (Random Search). Оптимизировались следующие гиперпараметры: для CNN — количество сверточных слоев, количество фильтров в каждом слое, размер ядра свертки, функция активации, метод пулинга, learning rate; для RF — количество деревьев, максимальная глубина дерева, количество признаков для разделения; для KNN — количество соседей; для ANN — количество скрытых слоев, количество нейронов в каждом слое, функция активации, learning rate. Результаты сравнительного анализа, полученные на едином наборе данных в 2384 образца, объединяющем все типы неисправностей (начальные, ступенчатые, одиночные и множественные одновременные), показали явное превосходство предложенного подхода над всеми остальными методами, особенно при малом количестве обучающих данных. Результаты представлены на рис. 5.

При увеличении размера обучающей выборки эффективность CNN, RF и KNN приближается к комбинированному подходу, но достигает близких значений лишь при выборке выше 128. Важно отметить, что разработанный подход BG-CNN показал наилучшую способность обнаруживать начальные неисправности на ранних этапах, а также изолировать множественные одновременные неисправности, которые представляют наибольшую сложность для традиционных методов FDI. Была проведена оценка влияния различных параметров модели на производительность, включая размер

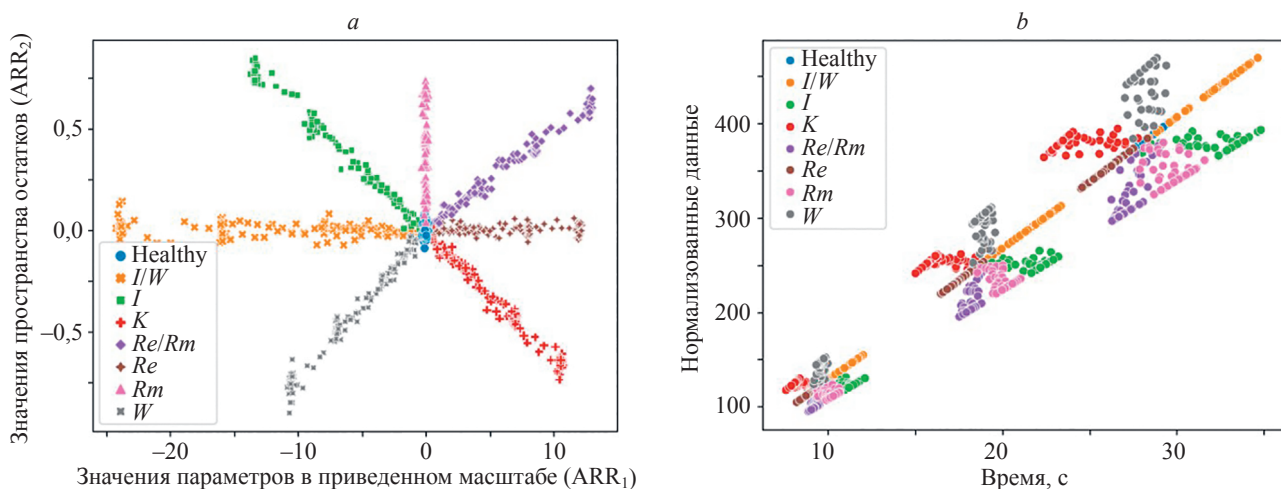


Рис. 4. Результаты детектирования неисправностей гидравлической системы: на основе остаточных данных (*a*); на основе данных датчиков (*b*)

Fig. 4. Results of hydraulic system fault detection: based on residual data (*a*); based on sensor data (*b*)

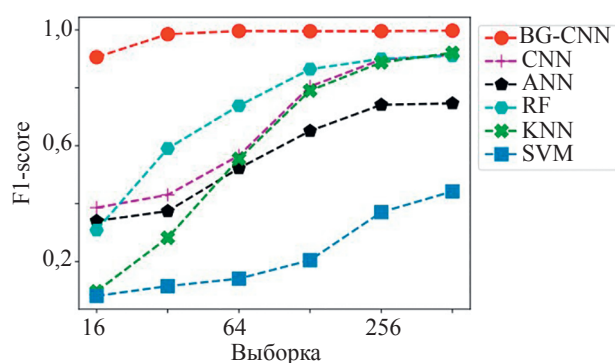


Рис. 5. Сравнение эффективности методов детектирования неисправностей

Fig. 5. Comparison of the fault detection methods effectiveness

обучающей выборки. Результаты показали, что подход BG-CNN сохраняет точность, превышающую точность CNN, RF и KNN, как отдельных методов, не менее, чем на 10 % при различных значениях параметров и количестве измерений, большем, чем 128. Проведено исследование чувствительности модели к параметрическим неопределенностям, что показало следующее:

- модель с LFT-BG демонстрирует большую точность по сравнению с подходом с константными значениями;
- результаты не зависят от размера обучающей выборки, но наибольший прирост заметен при размере выборки 32;
- подход BG-CNN показывает лучшие результаты в обнаружении неисправностей на ранних этапах.

Литература

1. Bouamama B.O., Biswas G., Loureiro R., Merzouki R. Graphical methods for diagnosis of dynamic systems: Review // *Annual reviews in control*. 2014. V. 38. N 2. P. 199–219. <https://doi.org/10.1016/j.arcontrol.2014.09.004>
2. Isermann R. *Fault-Diagnosis Applications: Model-Based Condition Monitoring: Actuators, Drives, Machinery, Plants, Sensors, and Fault-Tolerant Systems*. Springer Science & Business Media, 2011. 354 p.
3. Chen J., Patton R.J. *Robust Model-Based Fault Diagnosis for Dynamic Systems*. Springer Science & Business Media, 2012. V. 3. 356 p.
4. Ding S.X. *Model-Based Fault Diagnosis Techniques: Design Schemes, Algorithms and Tools*. Springer Science & Business Media, 2012. 504 p.
5. Akhenak A., Duviella E., Bako L., Lecoeuche S. Online fault diagnosis using recursive subspace identification: application to a dam-gallery open channel system // *Control Engineering Practice*. 2013. V. 21. N 6. P. 797–806. <https://doi.org/10.1016/j.conengprac.2013.02.013>
6. Lei Y., Yang B., Jiang X., Jia F., Li N., Nandi A.K. Applications of machine learning to machine fault diagnosis: a review and roadmap // *Mechanical Systems and Signal Processing*. 2020. V. 138. P. 106587. <https://doi.org/10.1016/j.ymssp.2019.106587>
7. Xu G., Yu Z., Lu N., Lv G. High-gain observer-based sliding mode control for hydraulic excavators // *Journal of Harbin Engineering University*. 2021. V. 42. N 6. P. 885–892.
8. Shen W., Yuan X., Liu M. Event-triggered control for hydraulic position tracking system with extended state observer // *Journal of Mechanical Engineering*. 2022. V. 58. N 8. P. 274–284. <https://doi.org/10.3901/JME.2022.08.274>
9. Nasiri S., Khosravani M.R., Weinberg K. Fracture mechanics and mechanical fault detection by artificial intelligence methods: a review // *Engineering Failure Analysis*. 2017. V. 81. P. 270–293. <https://doi.org/10.1016/j.engfailanal.2017.07.011>

Заключение

Разработан комбинированный подход детектирования неисправностей (FDI) для гидравлических систем управления, интегрирующий графы связей (BG) с линейным дробным преобразованием (LFT-BG) для генерации устойчивых остатков, и сверточные нейронные сети (CNN) для их классификации. Подход BG с LFT-BG позволил создать аналитическую модель и сгенерировать остатки, чувствительные к неисправностям, а CNN, в свою очередь, эффективно выявила закономерности для классификации разных типов неисправностей, включая их начальные, ступенчатые и множественные проявления. Предложенный комбинированный подход BG-CNN продемонстрировал превосходство над другими распространенными методами машинного обучения, примененными к тем же остаточным данным, особенно в условиях ограниченных обучающих данных, обеспечивая более высокую точность детектирования неисправностей, а также более высокую скорость обучения. Эксперименты показали, что подход LFT-BG с CNN эффективно обрабатывает нелинейности и позволяет создавать устойчивую к шуму систему детектирования. Практическая значимость заключается в повышении надежности автономной техники, где точное детектирование неисправностей гидравлических систем критично. Дальнейшие исследования должны быть направлены на адаптацию подхода к другим системам, оптимизацию архитектуры CNN и обучение на малых выборках, а также на изучение влияния параметров модели и применение подхода к реальным промышленным системам.

References

1. Bouamama B.O., Biswas G., Loureiro R., Merzouki R. Graphical methods for diagnosis of dynamic systems: Review. *Annual reviews in control*, 2014, vol. 38, no. 2, pp. 199–219. <https://doi.org/10.1016/j.arcontrol.2014.09.004>
2. Isermann R. *Fault-Diagnosis Applications: Model-Based Condition Monitoring: Actuators, Drives, Machinery, Plants, Sensors, and Fault-Tolerant Systems*. Springer Science & Business Media, 2011, 354 p.
3. Chen J., Patton R.J. *Robust Model-Based Fault Diagnosis for Dynamic Systems*. Springer Science & Business Media, 2012, vol. 3, 356 p.
4. Ding S.X. *Model-Based Fault Diagnosis Techniques: Design Schemes, Algorithms and Tools*. Springer Science & Business Media, 2012. 504 p.
5. Akhenak A., Duviella E., Bako L., Lecoeuche S. Online fault diagnosis using recursive subspace identification: application to a dam-gallery open channel system. *Control Engineering Practice*, 2013, vol. 21, no. 6, pp. 797–806. <https://doi.org/10.1016/j.conengprac.2013.02.013>
6. Lei Y., Yang B., Jiang X., Jia F., Li N., Nandi A.K. Applications of machine learning to machine fault diagnosis: a review and roadmap. *Mechanical Systems and Signal Processing*, 2020, vol. 138, pp. 106587. <https://doi.org/10.1016/j.ymssp.2019.106587>
7. Xu G., Yu Z., Lu N., Lv G. High-gain observer-based sliding mode control for hydraulic excavators. *Journal of Harbin Engineering University*, 2021, vol. 42, no. 6, pp. 885–892.
8. Shen W., Yuan X., Liu M. Event-triggered control for hydraulic position tracking system with extended state observer. *Journal of Mechanical Engineering*. 2022, vol. 58, no. 8, pp. 274–284. (in Chinese). <https://doi.org/10.3901/JME.2022.08.274>
9. Nasiri S., Khosravani M.R., Weinberg K. Fracture mechanics and mechanical fault detection by artificial intelligence methods: a review. *Engineering Failure Analysis*, 2017, vol. 81, pp. 270–293. <https://doi.org/10.1016/j.engfailanal.2017.07.011>

10. Hu Z., Zhao G., Li F., Zhou D. Fault diagnosis for nonlinear dynamical system based on adaptive unknown input observer // *Control and Decisions*. 2016. V. 31. N 5. P. 901–906.
11. Liu Y., Sun C., Chen Y., Lin G. Design and application on sliding mode controller for roller microporous system based on disturbance observer // *Forging and Stamping Technology*. 2022. V. 47. N 3. P. 116–120.
12. Bohagen B., Gravdahl J.T. Active surge control of compression system using drive torque // *Automatica*. 2008. V. 44. N 4. P. 1135–1140. <https://doi.org/10.1016/j.automatica.2007.11.002>
13. Дмитриев В.А., Марусина М.Я. Особенности построения бонд-графа шагающих роботов // *Известия высших учебных заведений. Приборостроение*. 2024. Т. 67. № 2. С. 195–199. <https://doi.org/10.17586/0021-3454-2024-67-2-195-199>
14. Rahman A., Hasan N., Zaki M. Modelling and validation of electric vehicle drive line architecture using Bond Graph // *Test Engineering and Management*. 2020. V. 82. P. 15154–15167.
15. Gonzalez-Avalos G., Gallegos N.B., Ayala-Jaimes G., Garcia A.P. Modeling and simulation in Multibond Graphs applied to three-phase electrical systems // *Applied Sciences*. 2023. V. 13. N 10. P. 5880. <https://doi.org/10.3390/app13105880>
16. Li F., Wu Z., Li J., Lai Z., Zhao B., Min C. A multi-step CNN-based estimation of aircraft landing gear angles // *Sensors*. 2021. V. 21. N 24. P. 8440. <https://doi.org/10.3390/s21248440>
17. Chen J., Xu Q., Guo Y., Chen R. Aircraft landing gear retraction/extension system fault diagnosis with 1-D dilated convolutional neural network // *Sensors*. 2022. V. 22. N 4. P. 1367. <https://doi.org/10.3390/s22041367>
18. Saucedo-Dorantes J.J., Delgado-Prieto M., Osornio-Rios R.A., Romero-Troncoso R.J. Industrial data-driven monitoring based on incremental learning applied to the detection of novel faults // *IEEE Transactions on Industrial Informatics*. 2020. V. 16. N 9. P. 5985–5995. <https://doi.org/10.1109/TII.2020.2973731>
19. Mo H., Li Y. Fault diagnosis based on interval analytic redundancy relation // *Journal of Nanjing University of Aeronautics and Astronautics*. 2021. V. 53. N 6. P. 972–980.
10. Hu Z., Zhao G., Li F., Zhou D. Fault diagnosis for nonlinear dynamical system based on adaptive unknown input observer. *Control and Decisions*, 2016, vol. 31, no. 5, pp. 901–906.
11. Liu Y., Sun C., Chen Y., Lin G. Design and application on sliding mode controller for roller microporous system based on disturbance observer. *Forging and Stamping Technology*, 2022, vol. 47, no. 3, pp. 116–120. (in Chinese)
12. Bohagen B., Gravdahl J.T. Active surge control of compression system using drive torque. *Automatica*, 2008, vol. 44, no. 4, pp. 1135–1140. <https://doi.org/10.1016/j.automatica.2007.11.002>
13. Dmitriev V.A., Marusina M.Ya. Features of constructing a bond graph of walking robots. *Journal of Instrument Engineering*, 2024, vol. 67, no. 2, pp. 195–199. (in Russian). <https://doi.org/10.17586/0021-3454-2024-67-2-195-199>
14. Rahman A., Hasan N., Zaki M. Modelling and validation of electric vehicle drive line architecture using Bond Graph. *Test Engineering and Management*, 2020, vol. 82, pp. 15154–15167.
15. Gonzalez-Avalos G., Gallegos N.B., Ayala-Jaimes G., Garcia A.P. Modeling and simulation in Multibond Graphs applied to three-phase electrical systems. *Applied Sciences*, 2023, vol. 13, no. 10, pp. 5880. <https://doi.org/10.3390/app13105880>
16. Li F., Wu Z., Li J., Lai Z., Zhao B., Min C. A multi-step CNN-based estimation of aircraft landing gear angles. *Sensors*, 2021, vol. 21, no. 24, pp. 8440. <https://doi.org/10.3390/s21248440>
17. Chen J., Xu Q., Guo Y., Chen R. Aircraft landing gear retraction/extension system fault diagnosis with 1-D dilated convolutional neural network. *Sensors*, 2022, vol. 22, no. 4, pp. 1367. <https://doi.org/10.3390/s22041367>
18. Saucedo-Dorantes J.J., Delgado-Prieto M., Osornio-Rios R.A., Romero-Troncoso R.J. Industrial data-driven monitoring based on incremental learning applied to the detection of novel faults. *IEEE Transactions on Industrial Informatics*, 2020, vol. 16, no. 9, pp. 5985–5995. <https://doi.org/10.1109/TII.2020.2973731>
19. Mo H., Li Y. Fault diagnosis based on interval analytic redundancy relation. *Journal of Nanjing University of Aeronautics and Astronautics*, 2021, vol. 53, no. 6, pp. 972–980.

Авторы

Дмитриев Валентин Алексеевич — руководитель практики, Ассоциация «Цифровые технологии в промышленности», Санкт-Петербург, 191181, Российская Федерация; аспирант, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, <https://orcid.org/0009-0007-5568-498X>, valentinswsu@mail.ru

Марусина Мария Яковлевна — доктор технических наук, профессор, профессор, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, <https://orcid.org/0000-0002-5611-8812>, myamarusina@itmo.ru

Authors

Valentin A. Dmitriev — Head of the Practice, Association of Digital Technologies in Industry, Saint Petersburg, 191181, Russian Federation; PhD Student, ITMO University, Saint Petersburg, 197101, Russian Federation, <https://orcid.org/0009-0007-5568-498X>, valentinswsu@mail.ru

Maria Ya. Marusina — D.Sc., Full Professor, ITMO University, Saint Petersburg, 197101, Russian Federation, <https://orcid.org/0000-0002-5611-8812>, myamarusina@itmo.ru

Статья поступила в редакцию 05.02.2025

Одобрена после рецензирования 02.04.2025

Принята к печати 22.05.2025

Received 05.02.2025

Approved after reviewing 02.04.2025

Accepted 22.05.2025



Работа доступна по лицензии
Creative Commons
«Attribution-NonCommercial»

doi: 10.17586/2226-1494-2025-25-3-545-553

УДК 004.02

Методы извлечения k -меров и признаков из наборов метагеномных графов де Брейна на основе информации о классах образцов

Артём Борисович Иванов¹✉, Анатолий Абрамович Шалыто²,
Владимир Игоревич Ульяновцев³

¹ Федеральный научно-клинический центр физико-химической медицины им. академика Ю. М. Лопухина
Федерального медико-биологического агентства, Москва, 119435, Российская Федерация

^{1,2,3} Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация

¹ abivanov@itmo.ru✉, <https://orcid.org/0000-0002-7997-0637>

² anatoly.shalyto@gmail.com, <https://orcid.org/0000-0002-2723-2077>

³ ulyantsev@itmo.ru, <https://orcid.org/0000-0003-0802-830X>

Аннотация

Введение. Рассмотрена задача сравнительного анализа наборов метагеномных образцов с использованием графов де Брейна. Для повышения точности работы классификационных моделей разработаны методы автоматического извлечения признаков на основе результатов сравнительного анализа метагеномных образцов, экспертных метаданных и статистических тестов. Под признаками в данной работе понимаются связанные подграфы графа де Брейна. **Методы.** Первый метод *unique_kmers* применяется для извлечения из данных строк длины k (k -меров), которые встречаются только в образцах одного класса. Второй метод *stats_kmers* применяется для извлечения k -меров, частота встречаемости которых статистически отличается между классами образцов. Для извлечения интерпретируемых признаков разработан третий метод, в котором реализовано выделение подграфов из графов де Брейна на основе опорных вершин, получаемых в результате применения одного из первых двух методов. Анализ данных состоит из двух этапов: вначале применяется метод *unique_kmers* или *stats_kmers* для предварительной обработки данных, затем к полученным результатам применяется третий метод для получения интерпретируемых признаков. **Основные результаты.** Апробация методов проведена на четырех сгенерированных тестовых наборах данных, которые моделируют параметры реальных метагеномных сообществ, такие как наличие похожих видов (штаммов) или разницу в частоте встречаемости бактерий. Разработанные методы были применены для извлечения признаков, которые использовались для классификации образцов из тестовых наборов. Для сравнения в качестве признаков использовались результаты таксономической аннотации образцов с помощью программы Kraken2. Показано, что качество классификации образцов повысилось при применении в классификационных моделях признаков, полученных с помощью предложенных методов, по сравнению с классификационными моделями, обученными на таксономических признаках. **Обсуждение.** Разработанные методы полезны при сравнительном анализе данных метагеномного секвенирования и могут служить основой систем поддержки принятия решений, например, при диагностировании заболеваний людей на основе данных секвенирования микробиоты кишечника.

Ключевые слова

извлечение признаков, граф де Брейна, k -меры, классификация, метагеномика

Ссылка для цитирования: Иванов А.Б., Шалыто А.А., Ульяновцев В.И. Методы извлечения k -меров и признаков из наборов метагеномных графов де Брейна на основе информации о классах образцов // Научно-технический вестник информационных технологий, механики и оптики. 2025. Т. 25, № 3. С. 545–553. doi: 10.17586/2226-1494-2025-25-3-545-553

Feature extraction methods for metagenome de Bruijn graphs collections based on samples classification information

Artem B. Ivanov¹✉, Anatoly A. Shalyto², Vladimir I. Ulyantsev³

¹ Lopukhin FRCC PCM, Moscow, 119435, Russian Federation

^{1,2,3} ITMO University, Saint Petersburg, 197101, Russian Federation

¹ abivanov@itmo.ru✉, <https://orcid.org/0000-0002-7997-0637>

² anatology.shalyto@gmail.com, <https://orcid.org/0000-0002-2723-2077>

³ ulyantsev@itmo.ru, <https://orcid.org/0000-0003-0802-830X>

Abstract

The paper considers the comparative analysis of metagenomic samples collections using de Bruijn graphs. We propose methods for automatic feature extraction based on the results of comparative sample analysis, expert metadata, and statistical tests to improve the accuracy of classification models. In this paper features are connected subgraphs of the de Bruijn graph. The first method, named *unique k-mers*, is used to extract strings of length k (k -mers) that occur only in samples of the certain class. The second method, named *stats k-mers*, is used to extract k -mers whose frequency of occurrence statistically differs between sample classes. To extract interpretable features, a third method has been developed that implements the extraction of subgraphs from de Bruijn graphs based on the selected nodes obtained as a result of applying one of the first two methods. Data analysis consists of two stages: firstly, *unique k-mers* or *stats k-mers* method is applied for data preprocessing, secondly, the third method is applied to obtain interpretable features. The methods were tested on four generated datasets that model the properties of real metagenomic communities such as the presence of similar species (strains) or differences in the relative abundance of bacteria. The developed methods were used to extract features. Machine learning model was trained in extracted features to classify samples from the test datasets. For comparison, the results of taxonomic annotation of samples using the Kraken2 program were used as features. It was shown that the accuracy of samples classification increased when using features obtained using the proposed methods in classification models compared to classification models trained on taxonomic features. The developed methods are useful for comparative analysis of metagenomic sequencing data and can form the basis of decision support systems, for example, in human diseases diagnostics based on gut microbiota sequencing data.

Keywords

feature extraction, de Bruijn graph, k -mers, classification, metagenomics

For citation: Ivanov A.B., Shalyto A.A., Ulyantsev V.I. Feature extraction methods for metagenome de Bruijn graphs collections based on samples classification information. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2025, vol. 25, no. 3, pp. 545–553 (in Russian). doi: 10.17586/2226-1494-2025-25-3-545-553

Введение

Метагеномика — раздел вычислительной биологии, который изучает в совокупности сообщества микроорганизмов, населяющих определенные экологические ниши: почва [1], водоемы [2], кожные покровы и кишечник человека [3]. Анализ метагеномных данных включает в себя определение видового состава образца, установление функций и роли отдельных таксонов (бактерий, вирусов) во взаимодействиях внутри сообщества и с окружающей средой, а также сравнение образцов между собой. Сравнительная метагеномика исследует сходства и различия метагеномных образцов с целью выявления закономерностей между микробным составом и свойствами окружающей среды или хозяина. Например, анализ микробиоты кишечника человека играет важную роль в современных медицинских исследованиях: для диагностирования заболеваний, прогноза успешности лечения и подбора персонализированной терапии [4–10].

Одной группой методов анализа метагеномных данных является сборка полных геномов с их последующей аннотацией — определением видового состава образца. Для этих целей широкое применение нашли графы де Брейна [11–14]. Несмотря на наличие эффективных методов для построения графа де Брейна и разбиения его на компоненты [15–17], задача сборки является сложной в связи с объемом данных, их комплексностью (образец является смесью геномов многих

видов) и зашумленностью. Сборка генома на 90 % из метагеномного образца требует в среднем десятикратного покрытия при секвенировании [18], но большой объем получаемых данных трудно обрабатывать, а стоимость ультраглубокого секвенирования не позволяет использовать его повсеместно. Аннотация применяется с целью установить видовой состав микроорганизмов в образце, однако она ограничена неполнотой баз данных, существующей из-за неизученной части мира микробов и их быстрой естественной эволюции в настоящее время. При этом данные, полученные со стандартной глубиной секвенирования, позволяют обнаруживать только виды с относительной представленностью больше одного процента в образце, что приводит к потере большой доли информации о разнообразии слабо представленных видов [19].

Другой группой методов анализа метагеномных данных является обработка, фильтрация и классификация «сырых данных» — прочтений, поступающих из секвенатора. Это позволяет учитывать всю извлеченную из образца информацию и не зависеть от баз данных, однако требует разработки эффективных алгоритмов. В задачах сравнения метагеномных последовательностей используются методы на основе k -меров — подстроках длины k . Некоторые алгоритмы позволяют эффективно манипулировать k -мерами и применять их для оценки схожести исходных образцов [20–22]. Другие методы используют статистические тесты для отбора подмножества k -меров, которые являются клю-

чевыми для разбиения исходных образцов на категории [23, 24]. Однако все методы анализа «сырых данных» позволяют провести только сравнение метагеномных образцов.

Аннотация и интерпретация признаков в виде k -меров является сложной задачей в связи с неполнотой баз данных и короткой длиной k -меров (от 11 до 91 символа), что не позволяет точно установить их биологическую роль. С целью повышения точности интерпретации в данной работе в качестве признаков используются связанные подграфы графа де Брейна. Строковые последовательности для подграфов обладают достаточной длиной (от 100 до нескольких тысяч символов), что позволяет точнее установить их биологическую роль по сравнению с k -мерами. Методы извлечения подграфов используются в задаче сборки геномов. Заметим, что предположительно не существует методов извлечения подграфов с использованием результатов сравнительного анализа метагеномных образцов, экспертных метаданных и статистических тестов.

Сложность анализа метагеномных данных состоит также в их объеме. С одной стороны, каждый образец метагеномного секвенирования занимает 5–10 Гб, поэтому одновременная обработка даже 100 образцов для сравнительного анализа и извлечения признаков требует большого объема оперативной памяти и дискового пространства, как правило, доступных только на вычислительных серверах. С другой стороны, типичное число образцов в метагеномных исследованиях находится в диапазоне от 100 до 200 из-за сложности сбора данных, особенно если речь идет о работе с данными метагеномов людей, например пациентов с определенным заболеванием. Таким образом, возникает проблема большого числа признаков, которые могут быть извлечены из данных, и малого числа образцов, что является критической проблемой для методов машинного обучения. По этой причине актуальной является задача разработки методов извлечения признаков из графов де Брейна, в том числе и больших (10^8 – 10^{10} вершин), при наличии требования интерпретируемости (описание свойств признаков, например, установление биологической функции, соответствующей извлеченной последовательности).

В настоящей работе разработаны методы извлечения признаков из наборов метагеномных графов де Брейна, которые на основе результатов сравнительного анализа метагеномных образцов, экспертных метаданных и статистических тестов извлекают из указанных графов ветвящиеся связанные подграфы. Новизна предложенных методов состоит в использовании информации о классах образцов на этапе извлечения признаков и извлечении признаков достаточной для интерпретации длины. Существующие решения удовлетворяют только одному из двух указанных свойств: или они используют метаданные, но извлекают признаки в виде коротких неинтерпретируемых последовательностей [21–24]; или они извлекают длинные интерпретируемые признаки из графов сборки, которые, однако, не позволяют с высокой точностью классифицировать образцы на группы [15, 16]. Показано, что признаки, полученные с помощью предложенных методов, повышают точность

работы классификационных моделей по сравнению с моделями, обученными на признаках таксономической аннотации.

Для извлечения признаков в виде подграфов из наборов метагеномных графов де Брейна необходимо научиться выделять опорные k -меры. Опорными k -мерами будем называть подмножество всех k -меров из образцов, встречаемость которых отличается между образцами разных классов. Эти k -меры будут сопоставлены опорным вершинам графа, на основе которых будут строиться подграфы. Для извлечения опорных k -меров было разработано два метода *unique_kmers* и *stats_kmers*.

Метод *unique_kmers* для извлечения опорных k -меров на основе их уникальной представленности в данных

В настоящей работе разработан метод извлечения опорных k -меров на основе их уникальной представленности в данных. Этот метод назван *unique_kmers* и состоит из следующих этапов.

Этап 1. Каждый метагеномный образец представляется в виде множества k -меров. Полученные файлы с k -мерами группируются на основании экспертных метаданных, например классов метагеномных образцов.

Этап 2. Фиксируется один из классов, назовем его X . Для него выбирается пороговое значение G . Из всех образцов класса X отбираются k -меры, которые присутствуют не менее чем в G файлах данного класса. Полученные k -меры являются кандидатами на специфичность для класса X .

Этап 3. Производится фильтрация k -меров из множества кандидатов, полученных на этапе 2. k -мер исключается из множества, если он встречается хотя бы в одном образце не из класса X . Полученное множество образцов является уникальным для класса X .

Этап 4. Этапы 2 и 3 выполняются независимо по одному разу для каждого класса из набора данных.

В результате применения этого метода для каждого класса извлекается множество уникальных опорных k -меров. Он обеспечивает высокую точность и скорость на данных с сильными различиями между образцами разных классов, однако может приводить к ошибкам из-за вариабельности и зашумленности в данных.

Метод *stats_kmers* для извлечения опорных k -меров на основе статистических тестов

Для извлечения опорных k -меров из наборов метагеномных образцов недостаточно найти только уникальные k -меры, поскольку метагеномы обладают естественной вариацией и зашумленностью, а в данных содержатся ошибки. В рамках работы разработан метод извлечения опорных k -меров на основе статистических тестов. Этот метод назван *stats_kmers* и состоит из следующих этапов.

Этап 1. Каждый метагеномный образец представляется в виде множества k -меров. Полученные файлы с k -мерами группируются на основании экспертных метаданных, например классов метагеномных образцов.

Этап 2. Отбираются специфические k -меры, встречаемость которых статистически значимо различается между образцами разных классов. Для каждого k -мера формируется таблица из двух строк и M столбцов, где M — число классов образцов. В таблицу для каждого класса записывается число образцов, в которых встречается и не встречается k -мер. Затем с помощью критерия χ^2 [25] для таблиц сопряженности проверяется гипотеза о равенстве частоты встречаемости k -мера в классах (использована поправка Йейтса [26]). В результате остаются только те k -меры, для которых гипотеза отвергается при заданном уровне значимости.

Этап 3. Поскольку на этапе 2 используется только информация о присутствии k -меров в образцах, выполняется дополнительная фильтрация на основе представленности k -меров в классах образцов. В качестве выборки для каждого класса образцов рассматриваются частоты встречаемости k -меров в образцах этого класса. Проводится попарное сравнение выборок с использованием критерия Манна–Уитни [27], который проверяет гипотезу о том, что обе выборки получены из одного распределения. В результате удаляются k -меры, для которых ни одна гипотеза о равенстве распределений не отвергается: удаляются k -меры, присутствующие примерно одинаково во всех образцах всех классов. При наличии трех и более классов образцов используется поправка Бонферрони [28] для множественных сравнений.

Этап 4. Для прошедших все этапы фильтрации k -меров рассчитывается среднее значение его представленности в каждом из классов. k -мер признается специфичным для того класса, в которой этот показатель наибольший.

В результате применения этого метода для каждого класса извлекается множество специфичных опорных k -меров. Он показывает высокую точность при наличии в образцах разных классов одинаковых организмов (бактерий, вирусов) с разной частотой встречаемости.

Метод извлечения признаков в виде подграфов из графа де Брейна на основе опорных k -меров

Для удобства дальнейшей работы с извлеченными множествами опорных k -меров был предложен метод, который позволяет на их основе получать более длинные и интерпретируемые признаки. Недостатками использования k -меров в качестве признаков является их большое число и малая длина. Большое число признаков в виде k -меров затрудняет их применение в моделях машинного обучения случайного леса и линейной регрессии, особенно учитывая факт малого числа размеченных метагеномных образцов. Кроме того, малая длина k -меров существенно затрудняет их биологическую аннотацию, что ограничивает дальнейшую интерпретируемость модели и применимость в реальных задачах.

Для решения указанных проблем разработан метод, который на основе опорных k -меров позволяет извлекать признаки в виде подграфов графа де Брейна. Метод основывается на работе одного из двух методов извлечения опорных k -меров (*unique_kmers* или *stats_*

kmers) и позволяет получить независимые признаки. Метод извлечения опорных k -меров, на основе которых будут строиться подграфы, выбирается экспериментатором в зависимости от входных данных (результаты экспериментов показали, что метод *stats_kmers* работает дольше, но признаки, построенные на основе его опорных k -меров, позволяют точнее классифицировать более похожие метагеномные образцы). Предлагаемый метод извлечения признаков в виде подграфов из графа де Брейна запускается отдельно для каждого класса образцов и соответствующего множества опорных k -меров, и состоит из следующих этапов.

Этап 1. Построение общего графа де Брейна из k -меров всех образцов (вершины — k -меры, ребра — пересечения длины $k - 1$). Вершины графа де Брейна, соответствующие опорным k -мерам, полученным с помощью метода *unique_kmers* или *stats_kmers*, помечаются соответствующим образом.

Этап 2. В графе производится поиск связанных компонент на основе опорных вершин. Запускается алгоритм обхода в ширину, начиная от случайно выбранной опорной вершины. Алгоритм состоит из следующих этапов.

Этап 2.1. Если из текущего k -мера существует путь только в одну вершину в графе, то следующая вершина присоединяется к текущей компоненте.

Этап 2.2. Если на пути в графе встречается развилка, то алгоритм выполняет предпросмотр в глубину в каждой ветке до следующей развилки. В случае обнаружения в какой-либо из веток опорной вершины, вся ветка добавляется к текущей компоненте.

Этап 2.3. Далее этапы 2.1 и 2.2 повторяются для последней вершины каждой добавленной ветки.

Этап 2.4. Если для текущего k -мера нет исходящих путей или на развилке не найдено ни одной ветки с опорными k -мерами, то обход заканчивается.

Этап 3. Все просмотренные на этапе 2 вершины отмечаются как посещенные. Этап 2 повторяется для еще не просмотренных опорных вершин. В результате получается набор подграфов де Брейна, содержащий все опорные k -меры. Каждый подграф может рассматриваться как отдельный признак, специфичный для данного класса образцов.

Этап 4. Каждый подграф сохраняется как множество строк, соответствующих линейным путям в графе. Полученные строки могут использоваться для аннотации и интерпретации признаков.

Метод позволяет для каждого класса метагеномных образцов (например, образцы здоровых и больных пациентов) получить набор признаков, специфичных для каждого класса. Для возможности использования признаков в классификационных моделях необходимо сопоставить им численные значения для каждого образца. Для этого определяется покрытие извлеченных подграфов k -мерами каждого образца. Оно рассчитывается как отношение числа k -меров в образце, попавших в данный подграф, к общему числу вершин в подграфе. Полученные значения объединяются в таблицу признаков с числом строк, равным числу образцов, и числом столбцов, равным суммарному числу признаков во всех классах. Эта таблица признаков используется в дальнейшем в моделях искусственного интеллекта для

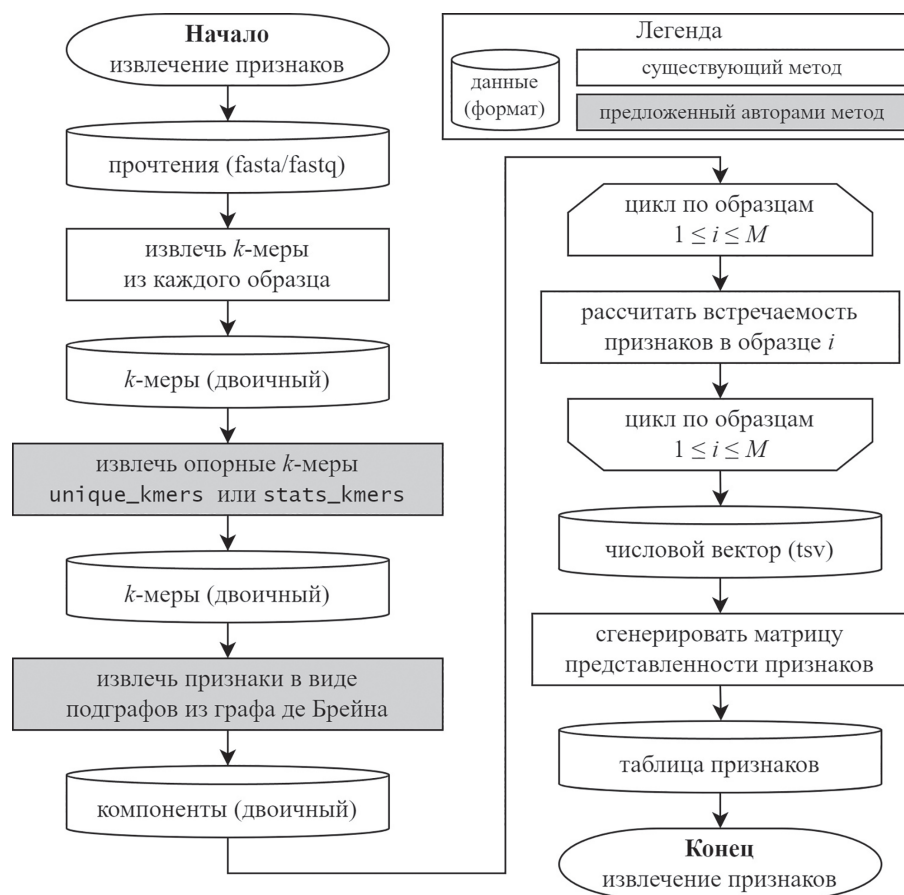


Рисунок. Схема алгоритма извлечения признаков из метагеномных данных
 Figure. Algorithmic pipeline for feature extraction from metagenomic data

обучения классификационных моделей и предсказании классов для новых неразмеченных метагеномных образцов. Кроме того, для возможности биологического анализа, аннотации и интерпретации извлеченных признаков они преобразовываются из внутреннего двоичного формата подграфов в набор строковых последовательностей. В дальнейшем производится их поиск в базах данных для установления видовой принадлежности и выполняемых функций. Схема алгоритма (рисунок) извлечения признаков из метагеномных данных содержит предложенные в данной работе методы и показывает этапы обработки метагеномных данных от файла с прочтениями до интерпретируемых признаков.

Вычислительные эксперименты

Для валидации разработанных методов проведены вычислительные эксперименты с использованием сгенерированных данных, моделирующих метагеномные сообщества различной степени сложности.

Наборы данных. Для тестирования предложенных методов извлечения признаков из наборов метагеномных образцов были сгенерированы наборы метагеномных данных с известным распределением бактерий и известными различиями между двумя классами образцов. Были смоделированы четыре набора данных, отличающихся похожестью содержащихся в них бактерий, каждый из которых состоял из 40 метагеном-

ных образцов — по 10 образцов в двух классах для тренировочной и тестовой выборки. Видовой состав бактерий, содержащихся в разных наборах, приведен в табл. 1. Отбор видов бактерий для образца проводился следующим образом:

- для каждого образца случайным образом выбираются 10 бактерий из фиксированного списка из 20;
- для наборов данных 2 и 4 в каждый образец добавляется общий штамм кишечной палочки *E. Coli* NZ_CP007265.1;
- для наборов данных 1 и 2 в образцы классов А и Б добавляются различные штаммы кишечной палочки: *E. Coli* NC_000913.3 в класс А и *E. Coli* NC_002695.2 в класс Б;
- для наборов данных 3 и 4 в образцы классов А и Б добавляется одинаковый штамм кишечной палочки *E. Coli* NC_000913.3, причем гарантируется, что относительная частота его встречаемости в образцах класса А находится в интервале $[0,01; 0,05]$, а в образцах класса Б в интервале $[0,05; 0,1]$.

Для генерации метагеномных прочтений на основе файлов с видовым составом образцов использовалась программа InSilicoSeq [29]. В качестве параметров запуска была выбрана модель прочтений `--model hiseq`, число прочтений `--n-reads 5 000 000` и модель относительной встречаемости видов `--abundance exponential`.

Методика эксперимента. Для отбора опорных k -меров из сгенерированных данных были использо-

Таблица 1. Параметры сгенерированных наборов данных

Table 1. Generated datasets properties

Параметр		Класс А		Класс Б	
Выборка		Тренировочная	Тестовая	Тренировочная	Тестовая
Число образцов		10	10	10	10
Номер набора данных	1	10 случайных бактерий из 20			
		<i>E. Coli</i> NC_000913.3		<i>E. Coli</i> NC_002695.2	
	2	10 случайных бактерий из 20 + <i>E. Coli</i> NZ_CP007265.1			
		<i>E. Coli</i> NC_000913.3		<i>E. Coli</i> NC_002695.2	
	3	10 случайных бактерий из 20			
		<i>E. Coli</i> NC_000913.3 встречаемость ∈ [0,01; 0,05]		<i>E. Coli</i> NC_000913.3 встречаемость ∈ [0,05; 0,1]	
	4	10 случайных бактерий из 20 + <i>E. Coli</i> NZ_CP007265.1			
		<i>E. Coli</i> NC_000913.3 встречаемость ∈ [0,01; 0,05]		<i>E. Coli</i> NC_000913.3 встречаемость ∈ [0,05; 0,1]	

ваны два разработанных метода *unique_kmers* и *stats_kmers*. Затем применен метод извлечения признаков в виде подграфов из графа де Брейна на основе опорных k -меров и построены таблицы численных признаков. Сравнение проводилось с признаками, полученными с помощью метода таксономической аннотации Kraken2 [30]. Метод таксономической аннотации также в качестве результата выполнения возвращает таблицу численных признаков представленности различных таксонов в образцах. Полученные признаки были использованы для построения классификатора на основе метода машинного обучения — случайный лес [31], который был обучен по тренировочным образцам. Выбор модели обусловлен работой с табличными признаками и хорошей интерпретируемостью результатов работы случайного леса (для каждого дерева решений можно извлечь информацию, на основе каких признаков был классифицирован образец). При реализации использовался класс RandomForestClassifier из библиотеки scikit-learn [32] версии 1.3.0 для языка программирования Python версии 3.9.5. Для параметра числа решающих деревьев было установлено значение *--n-estimators 100*, для остальных параметров использовались значения по умолчанию. Признаки были применены для разметки тестовых образцов.

Поскольку в сгенерированных данных размер классов сбалансирован, то, как метрики качества были выбраны Precision и Recall [33]. Точность (Precision) определяется как доля верно классифицированных образцов среди всех классифицированных образцов. Чувствительность (Recall) определяется как доля верно классифицированных образцов среди истинных меток. В случае двухклассовой задачи (две категории образцов) *истинный класс* определяется как класс объектов, которые должны обнаруживаться с помощью классификатора (например, при разработке классификатора для диагностирования заболевания, истинным классом будет образец с заболеванием). К недостаткам метрики Recall относится необходимость выбора истинного класса, а также сложность обобщения на многоклассовые задачи (если требуется распознавать более чем

два класса образцов). Однако эти метрики являются популярными для оценки качества классификации в моделях искусственного интеллекта.

Результаты. Результаты классификации приведены в табл. 2. Из полученных результатов видно, что для трех из четырех сгенерированных наборов данных, включая самый сложный — набор данных 4, с помощью разработанных методов извлечения признаков было получено более высокое качество классификации по сравнению с классическим методом таксономической аннотации. Для оставшегося набора данных метод *stats_kmers* показал максимально возможную точность, как и таксономическая аннотация. Метод *unique_kmers* допускал ошибки классификации в наборах данных 3 и 4, что объясняется отличиями между классами только в относительной представленности, в то время как этот метод направлен на поиск различных организмов. Полученные результаты подтверждают работоспособность и полезность разработанных методов. Показано повышение точности работы классификационных моделей при использовании извлеченных признаков по сравнению с моделями, обученными на признаках таксономической аннотации.

Обсуждение

Сравнение наборов метагеномных образцов, которые разделены на классы, является актуальной для решения прикладных биомедицинских задач. Примерами таких данных могут быть образцы, взятые при исследовании микробиоты кишечника здоровых людей и пациентов с заболеванием. Извлечение признаков может использоваться для выявления отдельных видов бактерий или продуцируемых ими метаболитов, которые связаны с развитием заболевания. Также на основании извлеченных признаков могут обучаться классификационные модели для ранней диагностики и скрининга людей на наличие заболеваний без необходимости проходить инвазивные процедуры.

В настоящей работе предложены методы для извлечения признаков из наборов метагеномных образцов.

Таблица 2. Результаты классификации наборов данных с помощью разработанных методов извлечения признаков и таксономической аннотации на сгенерированных данных

Table 2. Classification results based on developed feature extraction methods and taxonomic annotation for generated metagenomic datasets

Номер набора данных	Метрика	Метод извлечения признаков		
		Таксономия Kraken2	<i>unique_kmers</i>	<i>stats_kmers</i>
1	Precision	0,74	0,91	0,91
	Recall	0,70	1,00	1,00
2	Precision	0,55	1,00	1,00
	Recall	0,55	1,00	1,00
3	Precision	1,00	0,55	1,00
	Recall	1,00	0,60	1,00
4	Precision	0,45	0,67	0,83
	Recall	0,45	1,00	1,00

Примечание. Полужирным шрифтом выделено значение лучшего метода для каждой задачи и метрики.

Два метода позволяют извлекать из данных опорные k -меры, которые затем используются для построения графа де Брейна и извлечения интерпретируемых признаков. Третий метод извлекает признаки из наборов графов де Брейна, которые используются для построения классификационных моделей. Для вычислительных экспериментов были промоделированы данные, которые отражают специфику реальных метагеномных образцов. Часто отличия между классами образцов заключаются в штаммах одного вида бактерии, один из которых является нейтральным для человека, а второй — условно-патогенным и содержит мутации, которые коррелируют с наличием заболевания. Также отличия между классами образцов могут заключаться в

относительной представленности видов внутри образца, что было промоделировано в наборах данных 3 и 4.

Результаты показали, что использование признаков, полученных с помощью предложенных методов, повышает качество классификации образцов по сравнению с моделями, обученными с использованием классических признаков таксономической аннотации Kraken2. Разработанные методы могут быть применены к открытым данным метагеномного секвенирования для формулирования биологических гипотез о взаимосвязях между составом микробиоты кишечника человека и различными заболеваниями, которые затем могут быть экспериментально проверены.

Заключение

В работе предложены методы извлечения k -меров и признаков из наборов метагеномных данных, использующие информацию о классах образцов. Предложено три метода, два из которых отбирают опорные для наборов метагеномных образцов k -меры, а третий объединяет опорные k -меры в признаки в виде подграфов графа де Брейна. Использование информации о классах образцов на этапе построения признаков позволяет повысить точность работы классификационных моделей, обучаемых на извлекаемых признаках. Объединение k -меров в более длинные признаки и извлечение вет-

вящихся подграфов позволяет повысить интерпретируемость признаков.

Предложенные методы имеют важное значение — они поддерживают работу с наборами метагеномных данных с неограниченным числом классов. Метод извлечения признаков может быть использован для классификации метагеномных данных различной природы и быть основой методов поддержки принятия решений, например, при диагностировании заболеваний людей на основе данных секвенирования микробиоты кишечника.

Литература

1. Fierer N. Embracing the unknown: disentangling the complexities of the soil microbiome // *Nature Reviews Microbiology*. 2017. V.15. N 10. P. 579–590. <https://doi.org/10.1038/nrmicro.2017.87>
2. Garner R.E., Kraemer S.A., Onana V.E., Fradette M., Varin M.P., Huot Y., Walsh D.A. A genome catalogue of lake bacterial diversity and its drivers at continental scale // *Nature Microbiology*. 2023. V. 8. N 10. P. 1920–1934. <https://doi.org/10.1038/s41564-023-01435-6>
3. Huttenhower C., Gevers D., Knight R., et al. Structure, function and diversity of the healthy human microbiome // *Nature*. 2012. V. 486. N 7402. P. 207–214. <https://doi.org/10.1038/nature11234>

References

1. Fierer N. Embracing the unknown: disentangling the complexities of the soil microbiome. *Nature Reviews Microbiology*, 2017, vol. 15, no. 10, pp. 579–590. <https://doi.org/10.1038/nrmicro.2017.87>
2. Garner R.E., Kraemer S.A., Onana V.E., Fradette M., Varin M.P., Huot Y., Walsh D.A. A genome catalogue of lake bacterial diversity and its drivers at continental scale. *Nature Microbiology*, 2023, vol. 8, no. 10, pp. 1920–1934. <https://doi.org/10.1038/s41564-023-01435-6>
3. Huttenhower C., Gevers D., Knight R., et al. Structure, function and diversity of the healthy human microbiome. *Nature*, 2012, vol. 486, no. 7402, pp. 207–214. <https://doi.org/10.1038/nature11234>

4. Olekhnovich E., Ivanov A., Babkina A., Sokolov A., Ulyantsev V., Fedorov D., Ilina E. Consistent stool metagenomic biomarkers associated with the response to melanoma immunotherapy // *Msystems*. 2023. V. 8. N 2. <https://doi.org/10.1128/msystems.01023-22>
5. Ivanova V., Chernetskaya E., Vasiluev P., Ivanov A., Tolstogonov I., Shafranskaya D., Ulyantsev V., Korobeynikov A., Razin S., Beloborodova N., et al. Hi-C metagenomics in the ICU: exploring clinically relevant features of gut microbiome in chronically critically ill patients // *Frontiers in Microbiology*. 2022. V. 12. P. 770323. <https://doi.org/10.3389/fmicb.2021.770323>
6. Olekhnovich E., Ivanov A., Ulyantsev V., Ilina E. Separation of donor and recipient microbial diversity allows determination of taxonomic and functional features of gut microbiota restructuring following fecal transplantation // *Msystems*. 2021. V. 6. N 4. P. e00811-21. <https://doi.org/10.1128/msystems.00811-21>
7. Lloyd-Price J., Arze C., Ananthakrishnan A.N., Schirmer M., Avila-Pacheco J., Poon T.W., Andrews E., Ajami N.J., Bonham K.S., Brislawn C.J., et al. Multi-omics of the gut microbial ecosystem in inflammatory bowel diseases // *Nature*. 2019. V. 569. N 7758. P. 655–662. <https://doi.org/10.1038/s41586-019-1237-9>
8. Jie Z., Xia H., Zhong S.-L., Feng Q., Li S., Liang S., Zhong H., Liu Z., Gao Y., Zhao H., et al. The gut microbiome in atherosclerotic cardiovascular disease // *Nature Communications*. 2017. V. 8. P. 845. <https://doi.org/10.1038/s41467-017-00900-1>
9. Yu J., Feng Q., Wong S.H., Zhang D., Liang Q., Qin Y., Tang L., Zhao H., Stenvang J., Li Y., et al. Metagenomic analysis of faecal microbiome as a tool towards targeted non-invasive biomarkers for colorectal cancer // *Gut*. 2017. V. 66. N 1. P. 70–78. <https://doi.org/10.1136/gutjnl-2015-309800>
10. Qin J., Li Y., Cai Z., Li S., Zhu J., Zhang F., Liang S., Zhang W., Guan Y., Shen D., et al. A metagenome-wide association study of gut microbiota in type 2 diabetes // *Nature*. 2012. V. 490. N 7418. P. 55–60. <https://doi.org/10.1038/nature11450>
11. Idury R.M., Waterman M.S. A new algorithm for DNA sequence assembly // *Journal of Computational Biology*. 1995. V. 2. N 2. P. 291–306. <https://doi.org/10.1089/cmb.1995.2.291>
12. Pevzner P.A., Tang H., Waterman M.S. An Eulerian path approach to DNA fragment assembly // *Proceedings of the National Academy of Sciences of the United States of America*. 2001. V. 98. N 17. P. 9748–9753. <https://doi.org/10.1073/pnas.171285098>
13. Compeau P.E., Pevzner P.A., Tesler G. How to apply de Bruijn graphs to genome assembly // *Nature Biotechnology*. 2011. V. 29. N 11. P. 987–991. <https://doi.org/10.1038/nbt.2023>
14. Компо Ф., Певзнер П. Алгоритмы биоинформатики. Москва: ДМК Пресс, 2023. 680 с.
15. Nurk S., Meleshko D., Korobeynikov A., Pevzner P.A. metaSPAdes: new versatile metagenomic assembler // *Genome Research*. 2017. V. 27. N 5. P. 824–834. <https://doi.org/10.1101/gr.213959.116>
16. Kolmogorov M., Bickhart D.M., Behsaz B., Gurevich A., Rayko M., Shin S.B., Kuhn K., Yuan J., Polevikov E., Smith T.P., et al. metaFlye: scalable long-read metagenome assembly using repeat graphs // *Nature Methods*. 2020. V. 17. N 11. P. 103–1110. <https://doi.org/10.1038/s41592-020-00971-x>
17. Bankevich A., Bzikadze A.V., Kolmogorov M., Antipov D., Pevzner P.A. Multiplex de Bruijn graphs enable genome assembly from long, high-fidelity reads // *Nature Biotechnology*. 2022. V. 40. N 7. P. 1075–1081. <https://doi.org/10.1038/s41587-022-01220-6>
18. Meyer F., Fritz A., Deng Z.-L., Koslicki D., Lesker T.R., Gurevich A., Robertson G., Alser M., Antipov D., Beghini F., et al. Critical assessment of metagenome interpretation: the second round of challenges // *Nature Methods*. 2022. V. 19. N 4. P. 429–440. <https://doi.org/10.1038/s41592-022-01431-4>
19. Pereira-Marques J., Hout A., Ferreira R. M., Weber M., Pinto-Ribeiro I., Van Doorn L.-J., Knetsch C. W., Figueiredo C. Impact of host DNA and sequencing depth on the taxonomic resolution of whole metagenome sequencing for microbiome analysis // *Frontiers in Microbiology*. 2019. V. 10. P. 1277. <https://doi.org/10.3389/fmicb.2019.01277>
20. Marçais G., Kingsford C. A fast, lock-free approach for efficient parallel counting of occurrences of k -mers // *Bioinformatics*. 2011. V. 27. N 6. P. 764–770. <https://doi.org/10.1093/bioinformatics/btr011>
21. Ondov B.D., Treangen T.J., Melsted P., Mallonee A.B., Bergman N., Koren S., Phillippy A.M. Mash: fast genome and metagenome distance estimation using MinHash // *Genome Biology*. 2016. V. 17. P. 132. <https://doi.org/10.1186/s13059-016-0997-x>
4. Olekhnovich E., Ivanov A., Babkina A., Sokolov A., Ulyantsev V., Fedorov D., Ilina E. Consistent stool metagenomic biomarkers associated with the response to melanoma immunotherapy. *Msystems*, 2023, vol. 8, no. 2. <https://doi.org/10.1128/msystems.01023-22>
5. Ivanova V., Chernetskaya E., Vasiluev P., Ivanov A., Tolstogonov I., Shafranskaya D., Ulyantsev V., Korobeynikov A., Razin S., Beloborodova N., et al. Hi-C metagenomics in the ICU: exploring clinically relevant features of gut microbiome in chronically critically ill patients. *Frontiers in Microbiology*, 2022, vol. 12, pp. 770323. <https://doi.org/10.3389/fmicb.2021.770323>
6. Olekhnovich E., Ivanov A., Ulyantsev V., Ilina E. Separation of donor and recipient microbial diversity allows determination of taxonomic and functional features of gut microbiota restructuring following fecal transplantation. *Msystems*, 2021, vol. 6, no. 4. pp. e00811-21. <https://doi.org/10.1128/msystems.00811-21>
7. Lloyd-Price J., Arze C., Ananthakrishnan A.N., Schirmer M., Avila-Pacheco J., Poon T.W., Andrews E., Ajami N.J., Bonham K.S., Brislawn C.J., et al. Multi-omics of the gut microbial ecosystem in inflammatory bowel diseases. *Nature*, 2019, vol. 569, no. 7758, pp. 655–662. <https://doi.org/10.1038/s41586-019-1237-9>
8. Jie Z., Xia H., Zhong S.-L., Feng Q., Li S., Liang S., Zhong H., Liu Z., Gao Y., Zhao H., et al. The gut microbiome in atherosclerotic cardiovascular disease. *Nature Communications*, 2017, vol. 8, pp. 845. <https://doi.org/10.1038/s41467-017-00900-1>
9. Yu J., Feng Q., Wong S.H., Zhang D., Liang Q., Qin Y., Tang L., Zhao H., Stenvang J., Li Y., et al. Metagenomic analysis of faecal microbiome as a tool towards targeted non-invasive biomarkers for colorectal cancer. *Gut*, 2017, vol. 66, no. 1, pp. 70–78. <https://doi.org/10.1136/gutjnl-2015-309800>
10. Qin J., Li Y., Cai Z., Li S., Zhu J., Zhang F., Liang S., Zhang W., Guan Y., Shen D., et al. A metagenome-wide association study of gut microbiota in type 2 diabetes. *Nature*, 2012, vol. 490, no. 7418, pp. 55–60. <https://doi.org/10.1038/nature11450>
11. Idury R.M., Waterman M.S. A new algorithm for DNA sequence assembly. *Journal of Computational Biology*, 1995, vol. 2, no. 2, pp. 291–306. <https://doi.org/10.1089/cmb.1995.2.291>
12. Pevzner P.A., Tang H., Waterman M.S. An Eulerian path approach to DNA fragment assembly. *Proceedings of the National Academy of Sciences of the United States of America*, 2001, vol. 98, no. 17, pp. 9748–9753. <https://doi.org/10.1073/pnas.171285098>
13. Compeau P.E., Pevzner P.A., Tesler G. How to apply de Bruijn graphs to genome assembly. *Nature Biotechnology*, 2011, vol. 29, no. 11, pp. 987–991. <https://doi.org/10.1038/nbt.2023>
14. Compeau P., Pevzner P. *Bioinformatics Algorithms*. Active Learning Publishers, 2018, 728 p.
15. Nurk S., Meleshko D., Korobeynikov A., Pevzner P.A. metaSPAdes: new versatile metagenomic assembler. *Genome Research*, 2017, vol. 27, no. 5, pp. 824–834. <https://doi.org/10.1101/gr.213959.116>
16. Kolmogorov M., Bickhart D.M., Behsaz B., Gurevich A., Rayko M., Shin S.B., Kuhn K., Yuan J., Polevikov E., Smith T.P., et al. metaFlye: scalable long-read metagenome assembly using repeat graphs. *Nature Methods*, 2020, vol. 17, no. 11, pp. 103–1110. <https://doi.org/10.1038/s41592-020-00971-x>
17. Bankevich A., Bzikadze A.V., Kolmogorov M., Antipov D., Pevzner P.A. Multiplex de Bruijn graphs enable genome assembly from long, high-fidelity reads. *Nature Biotechnology*, 2022, vol. 40, no. 7, pp. 1075–1081. <https://doi.org/10.1038/s41587-022-01220-6>
18. Meyer F., Fritz A., Deng Z.-L., Koslicki D., Lesker T.R., Gurevich A., Robertson G., Alser M., Antipov D., Beghini F., et al. Critical assessment of metagenome interpretation: the second round of challenges. *Nature Methods*, 2022, vol. 19, no. 4, pp. 429–440. <https://doi.org/10.1038/s41592-022-01431-4>
19. Pereira-Marques J., Hout A., Ferreira R. M., Weber M., Pinto-Ribeiro I., Van Doorn L.-J., Knetsch C. W., Figueiredo C. Impact of host DNA and sequencing depth on the taxonomic resolution of whole metagenome sequencing for microbiome analysis. *Frontiers in Microbiology*, 2019, vol. 10, pp. 1277. <https://doi.org/10.3389/fmicb.2019.01277>
20. Marçais G., Kingsford C. A fast, lock-free approach for efficient parallel counting of occurrences of k -mers. *Bioinformatics*, 2011, vol. 27, no. 6, pp. 764–770. <https://doi.org/10.1093/bioinformatics/btr011>
21. Ondov B.D., Treangen T.J., Melsted P., Mallonee A.B., Bergman N., Koren S., Phillippy A.M. Mash: fast genome and metagenome distance estimation using MinHash. *Genome Biology*, 2016, vol. 17, pp. 132. <https://doi.org/10.1186/s13059-016-0997-x>

22. Maillat N., Collet G., Vannier T., Lavenier D., Peterlongo P. COMMET: comparing and combining multiple metagenomic datasets // *Proc. of the IEEE international conference on bioinformatics and biomedicine (BIBM)*. 2014. P. 94–98. <https://doi.org/10.1109/BIBM.2014.6999135>
23. Rahman A., Hallgrímsdóttir I., Eisen M., Pachter L. Association mapping from sequencing reads using k -mers // *Elife*. 2018. V. 7. P. e32920. <https://doi.org/10.7554/eLife.32920>
24. Wang Y., Chen Q., Deng C., Zheng Y., Sun F. KmerGO: a tool to identify group-specific sequences with k -mers // *Frontiers in Microbiology*. 2020. V. 11. P. 2067. <https://doi.org/10.3389/fmicb.2020.02067>
25. Greenwood P.E., Nikulin M.S. *A Guide to Chi-Squared Testing*. John Wiley & Sons, 1996. 304 p.
26. Крамер Г. Математические методы статистики. М.: Институт компьютерных исследований, 2019. 648 с.
27. Hettmansperger T.P., McKean J.W. *Robust nonparametric statistical methods*. CRC press, 2010. 554 p.
28. Dunn O.J. Multiple comparisons among means // *Journal of the American Statistical Association*. 1961. V. 56. N 293. P. 52–64. <https://doi.org/10.1080/01621459.1961.10482090>
29. Gourel H., Karlsson-Lindsjö O., Hayer J., Bongcam-Rudloff E. Simulating Illumina metagenomic data with InSilicoSeq // *Bioinformatics*. 2019. V. 35. N 3. P. 521–522. <https://doi.org/10.1093/bioinformatics/bty630>
30. Wood D.E., Lu J., Langmead B. Improved metagenomic analysis with Kraken 2 // *Genome Biology*. 2019. V. 20. N 1. P. 257. <https://doi.org/10.1186/s13059-019-1891-0>
31. Breiman L. Random forests // *Machine Learning*. 2001. V. 45. N 1. P. 5–32. <https://doi.org/10.1023/A:1010933404324>
32. Pedregosa F., Varoquaux G., Gramfort A., Michel V., et al. Scikit-learn: Machine learning in Python // *Journal of Machine Learning Research*. 2011. V. 12. P. 2825–2830.
33. Buckland M., Gey F. The relationship between recall and precision // *Journal of the American Society for Information Science*. 1994. V. 45. N 1. P. 12–19. [https://doi.org/10.1002/\(sici\)1097-4571\(199401\)45:1<12::aid-asi2>3.0.co;2-1](https://doi.org/10.1002/(sici)1097-4571(199401)45:1<12::aid-asi2>3.0.co;2-1)
22. Maillat N., Collet G., Vannier T., Lavenier D., Peterlongo P. COMMET: comparing and combining multiple metagenomic datasets // *Proc. of the IEEE international conference on bioinformatics and biomedicine (BIBM)*. 2014. pp. 94–98. <https://doi.org/10.1109/BIBM.2014.6999135>
23. Rahman A., Hallgrímsdóttir I., Eisen M., Pachter L. Association mapping from sequencing reads using k -mers. *Elife*, 2018, vol. 7, pp. e32920. <https://doi.org/10.7554/eLife.32920>
24. Wang Y., Chen Q., Deng C., Zheng Y., Sun F. KmerGO: a tool to identify group-specific sequences with k -mers. *Frontiers in Microbiology*, 2020, vol. 11, pp. 2067. <https://doi.org/10.3389/fmicb.2020.02067>
25. Greenwood P.E., Nikulin M.S. *A Guide to Chi-Squared Testing*. John Wiley & Sons, 1996, 304 p.
26. Cramér H. *Mathematical Methods of Statistics*. Princeton University Press, 2019, 575 p.
27. Hettmansperger T.P., McKean J.W. *Robust nonparametric statistical methods*. CRC press, 2010, 554 p.
28. Dunn O.J. Multiple comparisons among means. *Journal of the American Statistical Association*, 1961, vol. 56, no. 293, pp. 52–64. <https://doi.org/10.1080/01621459.1961.10482090>
29. Gourel H., Karlsson-Lindsjö O., Hayer J., Bongcam-Rudloff E. Simulating Illumina metagenomic data with InSilicoSeq. *Bioinformatics*, 2019. vol. 35, no. 3, pp. 521–522. <https://doi.org/10.1093/bioinformatics/bty630>
30. Wood D.E., Lu J., Langmead B. Improved metagenomic analysis with Kraken 2. *Genome Biology*, 2019, vol. 20, no. 1, pp. 257. <https://doi.org/10.1186/s13059-019-1891-0>
31. Breiman L. Random forests. *Machine Learning*, 2001, vol. 45, no. 1, pp. 5–32. <https://doi.org/10.1023/A:1010933404324>
32. Pedregosa F., Varoquaux G., Gramfort A., Michel V., et al. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 2011, vol. 12, pp. 2825–2830.
33. Buckland M., Gey F. The relationship between recall and precision. *Journal of the American Society for Information Science*, 1994, vol. 45, no. 1, pp. 12–19. [https://doi.org/10.1002/\(sici\)1097-4571\(199401\)45:1<12::aid-asi2>3.0.co;2-1](https://doi.org/10.1002/(sici)1097-4571(199401)45:1<12::aid-asi2>3.0.co;2-1)

Авторы

Иванов Артем Борисович — младший научный сотрудник, Федеральный научно-клинический центр физико-химической медицины им. академика Ю. М. Лопухина Федерального медико-биологического агентства, Москва, 119435, Российская Федерация; аспирант, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, [sc 57222438932](https://orcid.org/0000-0002-7997-0637), <https://orcid.org/0000-0002-7997-0637>, abivanov@itmo.ru

Шалыто Анатолий Абрамович — доктор технических наук, профессор, главный научный сотрудник, профессор, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, [sc 56131789500](https://orcid.org/0000-0002-2723-2077), <https://orcid.org/0000-0002-2723-2077>, anatoly.shalyto@gmail.com

Ульянцев Владимир Игоревич — кандидат технических наук, доцент, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, [sc 55062303000](https://orcid.org/0000-0003-0802-830X), <https://orcid.org/0000-0003-0802-830X>, ulyantsev@itmo.ru

Authors

Artem B. Ivanov — Junior Researcher, Lopukhin FRCC PCM, Moscow, 119435, Russian Federation; PhD Student, ITMO University, Saint Petersburg, 197101, Russian Federation, [sc 57222438932](https://orcid.org/0000-0002-7997-0637), <https://orcid.org/0000-0002-7997-0637>, abivanov@itmo.ru

Anatoly A. Shalyto — D.Sc., Chief Researcher, Full Professor, ITMO University, Saint Petersburg, 197101, Russian Federation, [sc 56131789500](https://orcid.org/0000-0002-2723-2077), <https://orcid.org/0000-0002-2723-2077>, anatoly.shalyto@gmail.com

Vladimir I. Ulyantsev — PhD, Associate Professor, ITMO University, Saint Petersburg, 197101, Russian Federation, [sc 55062303000](https://orcid.org/0000-0003-0802-830X), <https://orcid.org/0000-0003-0802-830X>, ulyantsev@itmo.ru

Статья поступила в редакцию 18.03.2025
Одобрена после рецензирования 26.04.2025
Принята к печати 27.05.2025

Received 18.03.2025
Approved after reviewing 26.04.2025
Accepted 27.05.2025



Работа доступна по лицензии
Creative Commons
«Attribution-NonCommercial»

doi: 10.17586/2226-1494-2025-25-3-554-564

Automatic calibration of the receiving line of information and control systems in real time

Nguyen Trong Nhan¹✉, Xuan Luong Nguyen², Phung Bao Nguyen³

¹ Air Defense - Air Force Technical Institute, Hanoi, 11400, Vietnam

² VNU University of Science, Hanoi, 100000, Vietnam

³ Le Quy Don Technical University, Hanoi, 11917, Vietnam

¹ 10th20th30th@gmail.com✉, <https://orcid.org/0000-0001-6626-893X>

² nguyenxuanluong_sdh22@hus.edu.vn, <https://orcid.org/0009-0000-6988-4738>

³ nguyenphungbao@lqdtu.edu.vn

Abstract

In this paper, the novel methodology for real-time automatic calibration of digital transceiver modules in the receiving path of information and control systems is presented. This methodology is grounded in the formation of calibration coefficients through a comparison between the complex signal amplitude at the output of the receiving path of the “virtual” reference module and the complex signal amplitude at the output of the receiving path following signal accumulation. The calibration value for each receiving path output complex signal amplitude is determined by multiplying the output complex signal amplitude by its corresponding calibration coefficient. The gain pattern of the information and control system is synthesized by calculating the weighted sum of the calibrated output complex signal amplitudes across all receiving paths, thereby maximizing the peak gain and minimizing side lobe levels. Simulations and experimental analyses were performed on an information and control system operating in the L-band to validate the proposed methodology. The results indicated a reduction in amplitude errors to 3.79 dB and a decrease in phase errors to 5°40'12". The proposed methodology meets the requirements for synthesizing a self-calibrating subsystem model employing a soft configuration approach.

Keywords

information and control system, radiation pattern, gain pattern, reference module, calibration coefficient

For citation: Nhan N.T., Nguyen X.L., Nguyen P.B. Automatic calibration of the receiving line of information and control systems in real time. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2025, vol. 25, no. 3, pp. 554–564. doi: 10.17586/2226-1494-2025-25-3-554-564

УДК 004.85

Автоматическая калибровка приемного тракта информационно-управляющих систем в режиме реального времени

Нгуен Чонг Нхан¹✉, Суан Лыонг Нгуен², Фунг Бао Нгуен³

¹ Технический институт ПВО, Ханой, 11400, Вьетнам

² Ханойский государственный университет, Ханой, 100000, Вьетнам

³ Вьетнамский государственный технический университет имени Ле Куй Дона, Ханой, 11917, Вьетнам

¹ 10th20th30th@gmail.com✉, <https://orcid.org/0000-0001-6626-893X>

² nguyenxuanluong_sdh22@hus.edu.vn, <https://orcid.org/0009-0000-6988-4738>

³ nguyenphungbao@lqdtu.edu.vn

Аннотация

Введение. Представлена новая методика автоматической калибровки приемного тракта цифровых приемопередающих модулей в реальном времени. **Метод.** Методика калибровки основана на формировании калибровочных коэффициентов путем сравнения комплексной амплитуды сигнала на выходе приемного тракта «виртуального» эталонного модуля и комплексной амплитуды сигнала на выходе приемного тракта

© Nhan N.T., Nguyen X.L., Nguyen P.B., 2025

после накопления сигнала. Калибровочное значение амплитуды комплексного сигнала на выходе каждого приемного тракта определяется с помощью умножения амплитуды комплексного сигнала на соответствующий калибровочный коэффициент. Диаграмма усиления информационно-управляющей системы синтезируется путем вычисления взвешенной суммы калиброванных амплитуд выходных комплексных сигналов по всем приемным трактам, что позволяет максимизировать пиковое усиление и минимизировать уровень боковых лепестков. **Основные результаты.** Для проверки предложенной методики проведены моделирование и экспериментальный анализ информационно-управляющей системы, работающей в L-диапазоне. Результаты показали снижение амплитудных ошибок до 3,79 дБ и фазовых ошибок до $5^{\circ}40'12''$. **Обсуждение.** Предложенная методика удовлетворяет требованиям к синтезу самокалибрующейся модели подсистемы с использованием подхода мягкой конфигурации.

Ключевые слова

информационно-управляющая система, диаграмма направленности излучения, коэффициент усиления, опорный модуль, калибровочный коэффициент

Ссылка для цитирования: Нхан Ч.Н., Нгуен С.Л., Нгуен Ф.Б. Автоматическая калибровка приемного тракта информационно-управляющих систем в режиме реального времени // Научно-технический вестник информационных технологий, механики и оптики. 2025. Т. 25, № 3. С. 554–564 (на англ. яз.). doi: 10.17586/2226-1494-2025-25-3-554-564

Introduction

In the current phase of technological development in multifunctional radar sensor systems, the proportion of Information and Control System (ICS) has grown significantly relative to conventional antenna types [1]. This expansion is driven by the need to achieve superior technical performance in such systems. ICS are composed of numerous transceiver modules which typically consist of functional units designed and manufactured using various active and passive components. Over time, due to factors, such as manufacturing variability, temperature fluctuations, aging, and other influences, the electrical parameters of these modules deviate from uniformity. Consequently, the amplitude-phase distribution of the electromagnetic field across the antenna aperture becomes non-uniform, resulting in a reduction in antenna gain and an increase in the side-lobe levels of the Radiation Pattern RP_{Σ} and Gain Patterns GP_{Σ} .

To ensure that the ICS operates with parameters aligned with theoretical calculations, it is essential to perform calibration not only at regular intervals but also in real time during operation. However, the calibration methods and techniques employed for these antenna arrays exhibit unique characteristics that differ markedly from those used for conventional antennas. This distinction is particularly relevant in the case of ICS where Digital Transceiver Modules (DTM) are synthesized using quadrature modulation and demodulation algorithms [2, 3]. This distinction arises from the process of synthesis and formation of the transmitting signal which is directly generated at each digital transceiver module. Unlike traditional active ICS, models, where the Radio Frequency (RF) transmitting signal is produced by a central source and distributed to system modules via a RF splitters and transmission cables, the ICS relies solely on digital data. This data encodes the parameters of the signal to be formed, eliminating the need for an intermediary RF splitters and transmission cables system. In conventional phased array antenna systems, the RF splitters and transmission cables itself introduces errors in the amplitude and phase of the RF signals at the input of the DTM [4–7]. By bypassing this system, the ICS architecture significantly reduces such errors.

The aforementioned factors can lead to imbalances between I and Q channels within the transmitting and receiving paths of DTM. To address these challenges, extensive research has focused on mitigating I/Q imbalances through diverse technical approaches. These studies commonly explore contributing factors, such as the operating frequency range of the radar sensor, the structural configuration of the receiving path, and the use of calibration signals. The ultimate goal is to develop and synthesize methods and models for automatic calibration using various approaches [8–15].

However, upon examining the aforementioned published results and other related studies, it can be seen that a small point has not been adequately addressed. Specifically, except for the external noise with high intensity, the power level of the reflected signals from the target objects of the ICS is often much smaller than the internal noise power level, as presented in [16, 17]. This characteristic necessitates careful consideration of power parameter synthesis for calibration signals. Moreover, the reliance on a rigidly synthesized reference object model may introduce additional errors in both measurement and automatic calibration processes. These observations form the foundation for the research presented in this paper.

Building on the preliminary analysis discussed above, the second section of this paper introduces a novel methodology for addressing the challenge of real-time automatic calibration of the receiving path in DTM of ICS grounded in the proposed mathematical model. This section also includes the synthesis of the structural diagram of the subsystem designed to implement this approach. The third section presents simulation and experimental results, demonstrating the correctness of the proposed methodology.

Theory and simulation of the proposed methodology

The novel methodology for solving the problem of real-time automatic calibration of the receiving path in DTM of ICS

According to [2, 3], the primary objective of the calibration process for any antenna system is to minimize the error between the measured gain pattern parameters and

their theoretically calculated counterparts. The approach presented in this study for addressing the automatic calibration problem aligns with the overarching goal. Specifically, the formation of the calibration coefficients proposed here is based on the comparison of the receiving path output signal of each DTM with the receiving path output signal of the reference DTM using a calibration signal as the input. From a physical perspective, this method is not fundamentally novel. The novelty here is expressed as follows.

According to [2, 16, 17], the reflected signals received from targets via the antenna element and reaching the input of each receiving path are typically significantly weaker than the internal noise level, $P_{int.noise}$, of the receiving path. Consequently, the design and implementation of a calibration signal must be carefully considered. An effective approach involves utilizing a RF calibration signal encoded with a Binary Phase-Shift Keying (BPSK), scheme featuring $M_{pos.}$ positions. The signal duration $t_{sig.}$ is designed to be $M_{pos.}$ times the width of an individual sub-pulse, $\tau_{sub.pul.}$ (Fig. 1). Furthermore, the power of the sub-pulse, $P_{\tau_{sub.pul.}}$ at the input of the receiving path is deliberately adjusted to remain significantly lower than the $P_{int.noise}$ of the receiving path to ensure accurate calibration. That is, the Signal-To-Noise Ratio (SNR) of the sub-pulse at the receiving path input $SNR_{\tau_{sub.pul.}}$ is within the limit (10^{-6} – 10^{-3}) for the common radar operating frequency bands as presented in [18, 19].

Application of the approach Software-Defined Radio (SDR) technology utilizes a “virtual” reference DTM model. This model incorporates data on the input and output signal parameters for both the transmitting and receiving channels which are initially determined through theoretical calculations. These parameters are subsequently will be refined, supplemented, and calibrated using data obtained from the calibration process conducted at the manufacturing facility as well as during field deployment and operational phases. So, this approach effectively eliminates the potential errors of physical reference DTM during actual operation. To ensure accurate estimation of the calibration coefficient parameters for each receiving path, their output signals must be coherently accumulated with the required accumulation coefficient. For example, it

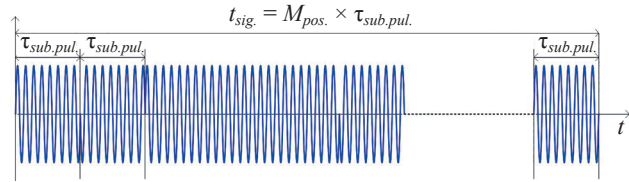


Fig. 1. The structure of the calibration signal

is possible to accumulate according to $M_{pos.}$, or accumulate according to the number of receiving paths corresponding to the number $M_{col} \times N_{row}$ of DTM, in which, M_{col} and N_{row} are the column and row indices and columns of the ICS, respectively, where DTM are located at the intersection points of the row and column. The calibration value of the output complex signal amplitude for each DTM is determined by multiplying the output complex signal amplitude with the corresponding calibration coefficient. So, the GP_{Σ} of the ICS is synthesized by computing the weighted sum of the calibrated output complex signal amplitudes across the receiving channels. This process will ensure the largest possible value for the total gain pattern GP_{Σ} .

A mathematical model of the receiving paths during automatic calibration according to the proposed methodology

Consider an ICS comprising $M_{col} \times N_{row}$ DTM in which its architecture synthesized based on the principles of quadrature modulation and demodulation. The calibration signal $S(j\omega)$ is constructed using a BPSK code with $M_{pos.}$ positions [19, 20]. The use of a rectangular windowed Fourier transform is instrumental in the analysis of complex signal images. This signal consists of RF sub-pulses generated by a Controlled Signal Generator (CSG). The generated signal passes through a directional coupler element and is subsequently routed via a circulator element, functioning as a transceiver switch, to the input of the receiving path of the i -th DTM, where $i \in [1, M_{col} \times N_{row}]$ as illustrated in Fig. 2. In there, $i = 0$ represents the “virtual” reference DTM.

Consequently, to ensure the SNR reaches a sufficiently high level required for subsequent processing, the digital data will have to be coherently accumulated by required to

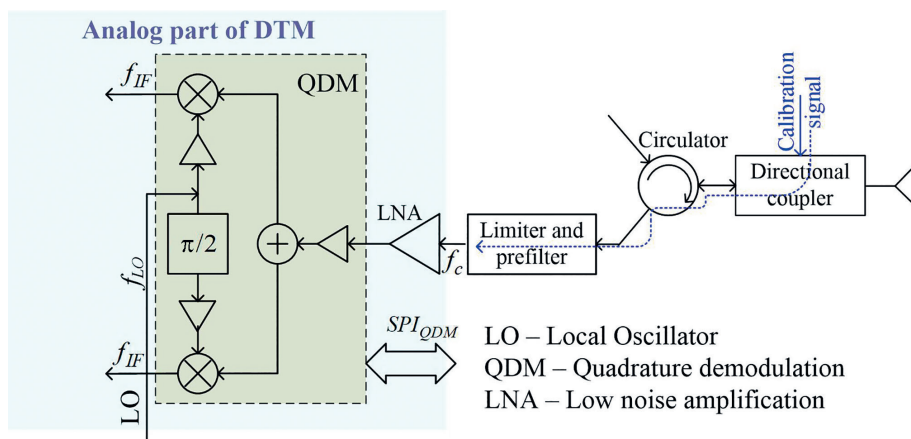


Fig. 2. The route of the calibration signal to the input of the receiving path: f_{IF} — intermediate frequency; f_{LO} — frequency of LO; SPI_{QDM} — serial peripheral interface of QDM; and f_c — carrier frequency

form a complex signal $S_i(j\omega)$. This signal comprises two components, real and imaginary, structured in the form:

$$\text{Re}S_i(j\omega) = \sum_m^M \text{Re}S_{m,i}(j\omega); \text{Im}S_i(j\omega) = \sum_m^M \text{Im}S_{m,i}(j\omega), \quad (1)$$

where, M is the number of coherent accumulations, and $m = 1, 2, \dots, M$. If we consider $M \cong M_{pos}$, the process of coherent accumulation can be interpreted as the sub-pulses being accumulated in accordance with M_{pos} . Furthermore, additionally coherent accumulation with $M_{col} \times N_{row}$, the expression (1) will be in the form:

$$\begin{aligned} \text{Re}S_i(j\omega) &= \sum_p^{M_{col} \times N_{row}} \sum_m^{M_{pos}} \text{Re}S_{p,m,i}(j\omega); \\ \text{Im}S_i(j\omega) &= \sum_p^{M_{col} \times N_{row}} \sum_m^{M_{pos}} \text{Im}S_{p,m,i}(j\omega). \end{aligned}$$

From this point, the complex calibration coefficient can be defined for each i -th receiving path:

$$K_i(j\omega) = \frac{S_0(j\omega)}{S_i(j\omega)} = \frac{|S_0(j\omega)| \times \exp(j\varphi_0)}{|S_i(j\omega)| \times \exp(j\varphi_i)}. \quad (2)$$

Where $|S_0(j\omega)|$ and φ_0 denote the amplitude and phase, respectively, of the accumulated complex signal corresponding to the receiving path of the reference module. And $|S_i(j\omega)|$ and φ_i represent the amplitude and phase, respectively, of the accumulated complex signal associated with the receiving path of the i -th module undergoing calibration. The values of $|S_{\Sigma,i}(j\omega)|$ and φ_i are determined by the expression: $|S_i(j\omega)| = \sqrt{(\text{Re}S_i(j\omega))^2 + (\text{Im}S_i(j\omega))^2}$; $\varphi_i = \arctg(\text{Im}S_i(j\omega)/\text{Re}S_i(j\omega))$.

During the process of coherent accumulation, whether following M_{pos} or using $M_{col} \times N_{row}$, the absolute value of the complex amplitude $|S_{\Sigma,0}(j\omega)|$ for the reference channel and $|S_{\Sigma,i}(j\omega)|$ for the channel undergoing calibration can be expressed in the following form: $|S_{\Sigma,0}(j\omega)| = M_{pos} \times |S_0(j\omega)|$ or $|S_{\Sigma,0}(j\omega)| = M_{pos} \times (M_{col} \times N_{row}) \times |S_0(j\omega)|$; $|S_{\Sigma,i}(j\omega)| = M_{pos} \times |S_i(j\omega)|$ or $|S_{\Sigma,i}(j\omega)| = M_{pos} \times (M_{col} \times N_{row}) \times |S_i(j\omega)|$.

The complex calibration coefficient for each receiving path in expression (2) can be rewritten as $K_i(j\omega) = |K_i(j\omega)| \exp(j\Delta\varphi_i)$.

Where, $|K_i(j\omega)|$ is defined as the absolute value of the complex calibration coefficient for the i -th receiving path, and the $\Delta\varphi_i = \varphi_0 - \varphi_i$ represents the magnitude of the phase error that requires correction. It is possible to write that: $|K_i(j\omega)| = |S_0(j\omega)|/|S_i(j\omega)|$. This equation is expressed in the form of a trigonometric function: $K_i(j\omega) = |K_i(j\omega)| \times [\cos(\Delta\varphi_i) + j\sin(\Delta\varphi_i)]$. And, with two real and imaginary parts: $\text{Re}K_i(j\omega) = |K_i(j\omega)| \times \cos(\Delta\varphi_i)$; $\text{Im}K_i(j\omega) = |K_i(j\omega)| \times \sin(\Delta\varphi_i)$.

The complex amplitude of the signal at the output of the receiving path, $S_{out,i}(j\omega)$, expressed in both exponential and trigonometric forms, is represented as follows: $S_{out,i}(j\omega) = |S_{out,i}(j\omega)| \times \exp(j\varphi_i)$; $S_{out,i}(j\omega) = |S_{out,i}(j\omega)| \times [\cos(\varphi_i) + j\sin(\varphi_i)]$.

The product of the complex amplitude of the output signal from the i -th receiving path and the corresponding complex calibration coefficient for that path can be represented in the following form: $S_{out,i}(j\omega) \times K_i(j\omega) = |S_{out,i}(j\omega)| \times |K_i(j\omega)| [\cos(\varphi_i + \Delta\varphi_i) + j\sin(\varphi_i + \Delta\varphi_i)]$. We have:

$|S_{out,0}(j\omega)| = |S_0(j\omega)|$; $|S_{out,i}(j\omega)| = |S_i(j\omega)|$, and $S_{out,i}(j\omega) \times K_i(j\omega) = |S_0(j\omega)| \times [\cos(\varphi_0) + j\sin(\varphi_0)]$. This equation represents the calibrated complex amplitude of the input signal for the i -th receiving path. And, it can also be expressed either in exponential form or as two separate components, namely, the real and imaginary parts, as follows:

$$\begin{aligned} S_{i,cal}(j\omega) &= |S_0(j\omega)| \times \exp(j\varphi_0) \\ \text{and } \text{Re}S_{i,cal}(j\omega) &= |S_0(j\omega)| \times \cos(\varphi_0); \\ \text{Im}S_{i,cal}(j\omega) &= |S_0(j\omega)| \times \sin(\varphi_0). \end{aligned}$$

These components are precisely equal to the quadrature components of the complex amplitude of the output signal from the reference receiving path. Thus, after calibration, the output signals of all paths will be identical in both amplitude and phase.

Synthesis of the subsystem diagram for real-time automatic calibration of the receiving path in DTM of the ICS

Based on the results derived from the proposed mathematical model, the process for automatic calibration of the receiving path in the DTM can be synthesized with the following key considerations:

- Calibration Signal Timing: The calibration signal is supplied in to the input of the receiving path during the time interval corresponding to the receiving period of the radar sensor reflected signals (Fig. 3).
- Signal Blocking: To minimize errors caused by interference, the calibration signal is isolated by a detection unit that filters out reflected signals and significant external noise (Fig. 4).

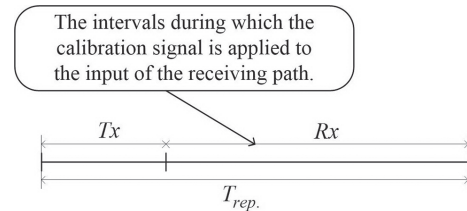


Fig. 3. The receiving period Rx: Tx is the transmission interval of the signal.

T_{rep} — period repeats

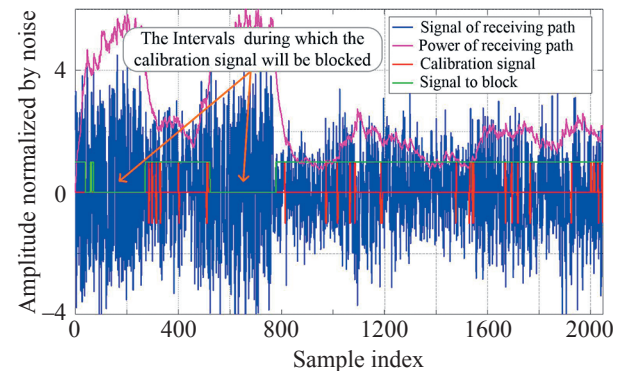


Fig. 4. The intervals during which the calibration signal will be blocked

- “Virtual” reference DTM structure: The “virtual” reference DTM consists of two primary components: the first component is integrated within the Central Computing and Processing Unit (CCPU), and generates control data containing information about the structure of $S_0(j\omega)$. The second component is a CSG which is structurally independent and operates in isolation within the digital part of the transmitting part in the DTM to form $S_0(j\omega)$.

Fig. 5 presents the structural diagram of the automatic calibration subsystem for the receiving path within the DTM of the ICS. Accordingly, in each DTM, the calibration signal is generated by the CSG based on input data provided by the CCPU which includes information regarding the formation of the signal and periodic updates. The CSG produces the calibration signal $S_i(j\omega)$, with parameters $t_{sig.}$, $\tau_{sub.pul.}$ and $P_{\tau_{sub.pul.}} \ll P_{int.noise.}$ which is transmitted to the receiving paths upon receiving the P_x command. An instantaneous power meter monitors the power level of the received signal which includes both reflected signals and external noise. If the measured power level significantly exceeds the internal noise power, the blocking command will be formed and sent to the CCPU, enabling it to block the supply of calibration signal data for each CSG.

Following single-pulse filtering, the output of the each i -th DTM produces two quadrature components of the calibration signal, $ReS_i(j\omega)$ and $ImS_i(j\omega)$. These components are transmitted to the Calculation and Automatic Calibration Unit (C&ACU) where accumulation operations, specifically digital coherent accumulation processing, are performed based on the indices $M_{pos.}$ or $M_{pos.}$ and $M_{col} \times N_{row}$, as required. Subsequently, the processed data are compared with the corresponding data of the reference signal to estimate the calibration coefficients

$K_i(j\omega)$ and to form the data of $S_{i,n}(j\omega)$, for generating the necessary $S_i(j\omega)$. Data of $S_{i,n}(j\omega)$ is transmitted to the CSG_i of the i -th DTM according to the specified address, and forwarded to the CCPU. Additionally, the digital data of $ReS_{i,cal}(j\omega)$ and $ImS_{i,cal}(j\omega)$ with $i \in (M_{col} \times N_{row})$ are transmitted to the GP_Σ formation unit.

It is easy to see that the functional units involved in the real-time automatic calibration of the receiving path within the DTM are integral components of both the basic architecture of the DTM and the overall ICS. Notably, the calibration process necessitates only a minimal addition of hardware components. Consequently, the use of the term “integrated subsystem” to explain in this study is entirely justified.

Simulation of the proposed methodology

The simulation conducted to validate the theoretical research findings incorporates the components illustrated in Fig. 6. These include four receiving paths where structural improvements aimed at adaptive dynamic range control have been implemented as described in [17]; a Local Oscillator voltage generation and Distribution Unit (LO&DU); an Automatic Calculation, Calibration, and Control Unit (CAC&CU); and a computer serving as the central processing and display interface. The signal parameters utilized during the simulation are: $f_0 = 1560\text{--}1590$ MHz; $f_{LO} = 1470\text{--}1500$ MHz; $f_{IF} = 90$ MHz; $M_{pos.} = 256$. Fig. 7 presents the simulation results of signals from the four receiver paths ($M_{rec.p.} = 4$), both prior to and following the automatic calibration process.

The cumulative processing is conducted with $M = M_{pos.} \times M_{rec.p.} = 256 \times 4 = 1024$, corresponding to the number of samples. For the sake of simplifying the simulation process, it is assumed that the receiving path No 1 serves as the reference path. Table 1 presents the results

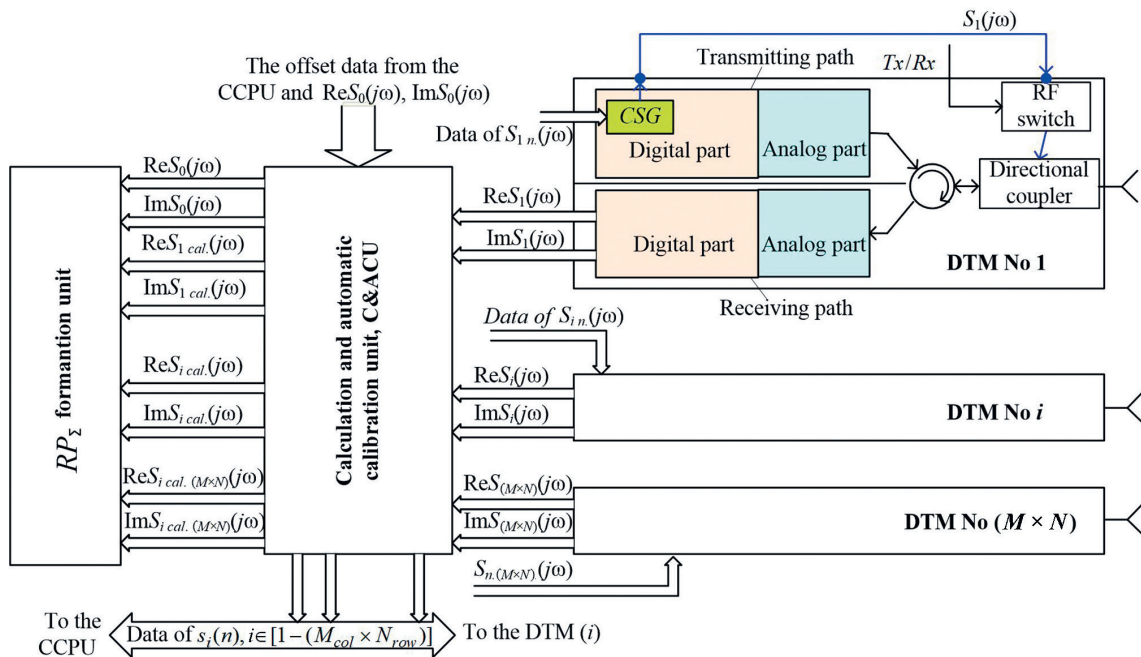


Fig. 5. The general structural diagram of the automatic calibration subsystem for the receiving path within the DTM of the ICS: RP_Σ — Radiation Pattern after beamforming; S_{1cal} — the first calibration signal; S_{ical} — the i -th calibration signal; $S_{ical(M \times N)}$ — the i -th calibration signal at M_{col} and N_{row} ; S_{1n} — the reference of the first signal; S_{in} — the reference of the i -th signal; $S_{i(M \times N)}$ — the reference of the i -th signal at M_{col} and N_{row}

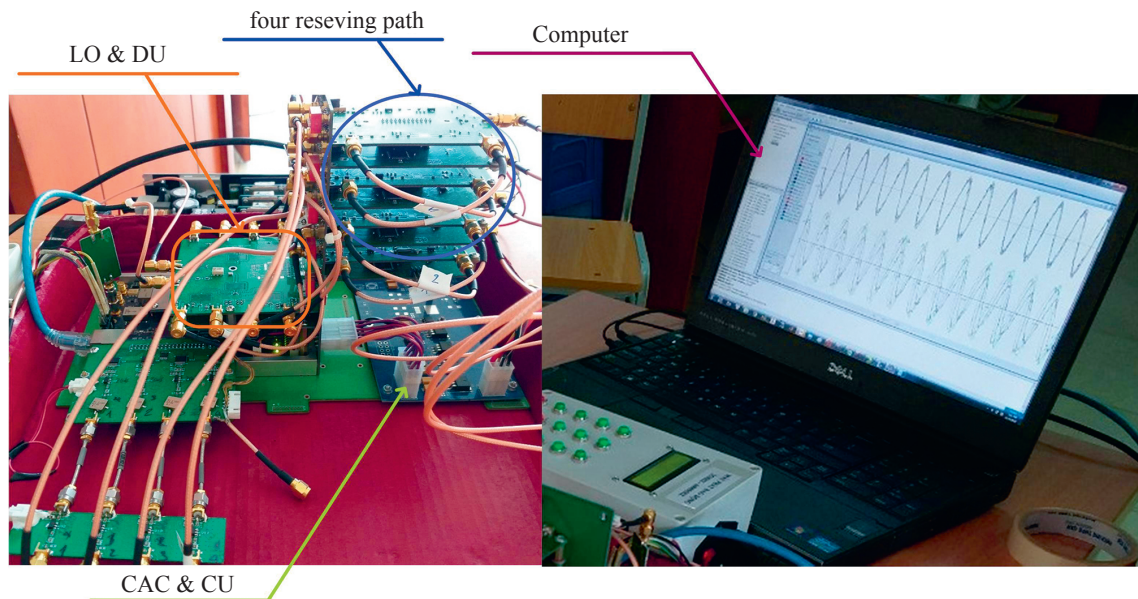


Fig. 6. Diagram illustrating the implementation of the simulation for the subsystem responsible for the automatic calibration of receiving paths (four paths)

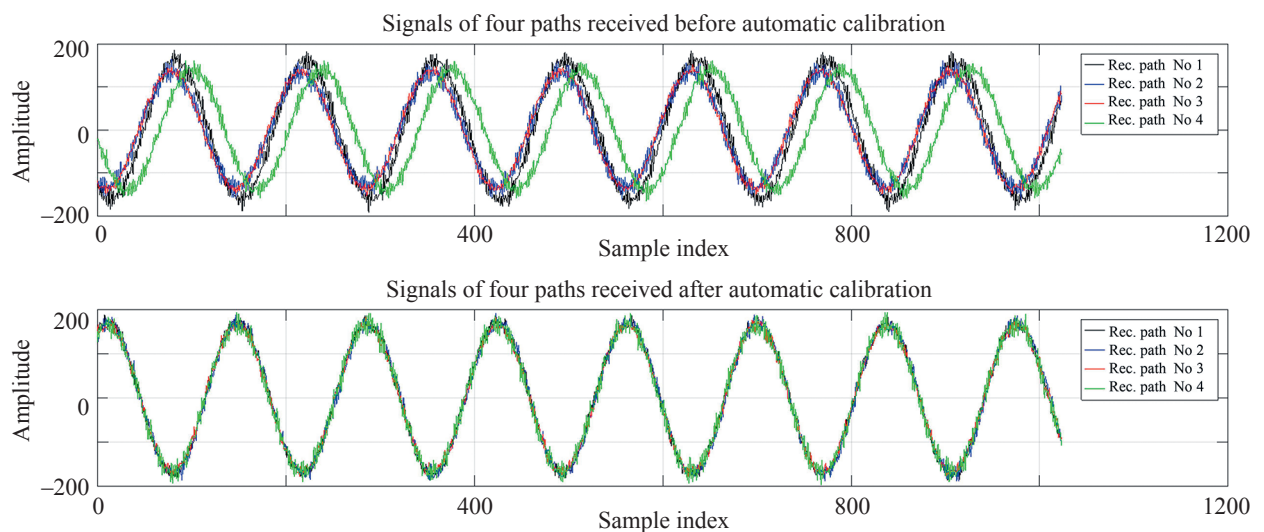


Fig. 7. Simulation results of signals from the four receiver paths, both prior to and following the automatic calibration process

Table 1. Values of the signal amplitude and phase errors, without incorporating improvements to the receiving path model

State of the receiving path	Path No 2	Path No 3	Path No 4
Signal amplitude error (*), dB			
Before calibration	2.20	1.90	2.60
After calibration	0.25	0.34	0.41
Signal phase error (*)			
Before calibration	6°18'	5°30'	7°12'
After calibration	1°36'	2°00'	1°06'

Note. (*) Compared to the reference receiving path.

of the estimation of amplitude and phase errors of the signal, both before and after calibration. These estimations were performed without incorporating improvements to the receiving path model, which was synthesized as a model of

a fully matched filter, and without implementing adaptive dynamic range control.

Table 2 presents the results of the estimation of amplitude and phase errors of the signal, both before and

Table 2. Values of the signal amplitude and phase errors, when incorporating improvements to the receiving path model

State of the receiving path	Path No 2	Path No 3	Path No 4
Signal amplitude error (*), dB			
Before calibration	3.10	2.30	2.80
After calibration	0.14	0.13	0.16
Signal phase error (*)			
Before calibration	4°30'	6°18'	7°12'
After calibration	0°48'	1°06'	0°42'

Note. (*) Compared to the reference receiving path.

after calibration. These estimations were performed when incorporating improvements to the receiving path model which was synthesized as a model of a fully matched filter, and with implementing adaptive dynamic range control.

The methodology for identifying amplitude and phase errors is comprehensively described in [18, 19, 21]. The simulation results validate the correctness of the proposed methodology. Specifically, as presented in Table 1, the magnitude of amplitude errors is reduced, ranging from 1.95 dB to 2.19 dB and phase errors decrease to values between 3°30' and 6°06'. Furthermore, implementing the solutions to enhance the receiving path yields further reductions in error magnitudes, for both the amplitude and phase of the signal. This improvement significantly enhances the formation of GP_{Σ} as evidenced by parameters such as peak gain and the suppression of the sub-lobes, as illustrated in Fig. 9.

Results and Discussion

The validity of the proposed methodology is further corroborated through experimental verification. The experimental setup is depicted in Fig. 8 and consists of a radar sensor system equipped with four receiving paths, each comprising a DTM connected to an individual antenna element (Fig. 8, a), all mounted on a turntable (Fig. 8, c). Additionally, the system includes a CSG and its associated irradiation antenna (Fig. 8, b), positioned at a specified

distance D from the radar probe system to ensure the incident waves can be approximated as plane waves. A computer serves as the CCPU and display. The catalogue of key components employed in the experiment is presented in Table 3.

For a frequency f_0 in the range of 1560–1590 MHz, the far-field region is considered to be at distances greater than 2 m. The experiment was conducted with a separation r of 7.5 m between the radiation antenna and the antenna system of ICS. To consistent with the simulation procedure, the experiment utilized the first receiving path as the reference path.

The experimental results, summarized in Tables 4, and 5, present the errors in signal amplitude and phase both before and after calibration. Additionally, the results also illustrate the effects of enhancements to the receiving path, implemented based on the fully matched filter model and the dynamic range adaptive control methodology.

Experimental results were also obtained by synthesizing GP_{Σ} for different observation angles, utilizing a turntable to rotate the ICS. These results were compared with those generated through simulation after calibration. Fig. 9 presents the synthesized GP_{Σ} obtained via the simulation method post-calibration. Fig. 10 illustrates the results derived from measuring and processing the experimental data.

Both GP_{Σ} synthesis results were obtained under conditions in which the receiving paths were improved

Table 3. Catalogue of the key components used in the experiment

Brands	Part Number	Applications
Macom	CG2H40010	Power amplifier
Macom	CMPA0527005	Power preamplifier
MAX	MAX2021	Direct Up/Down conversion Quadrature Mod/Demod
MAX	MAX9491 or equivalent	VCO and Synth
TI	DAC3283 or equivalent	16-Bits DAC
TI	ADC3683 or equivalent	High-speed ADCs
UMS	CHA3801-99F	L-Band LNA
UMS	GVA-63+	IF Amplifier
AD	ADL5531 or equivalent	Voltage-controlled amplifier/attenuator Operating frequency
AD	AD9914S or equivalent	3.5 GSPS Direct Digital Synthesizer with 12-Bit DAC
AD	AD9833 EP	CSG (Programmable Waveform Generator)
Atmel	AVR2560/V	8-Bit Atmel Microcontroller with

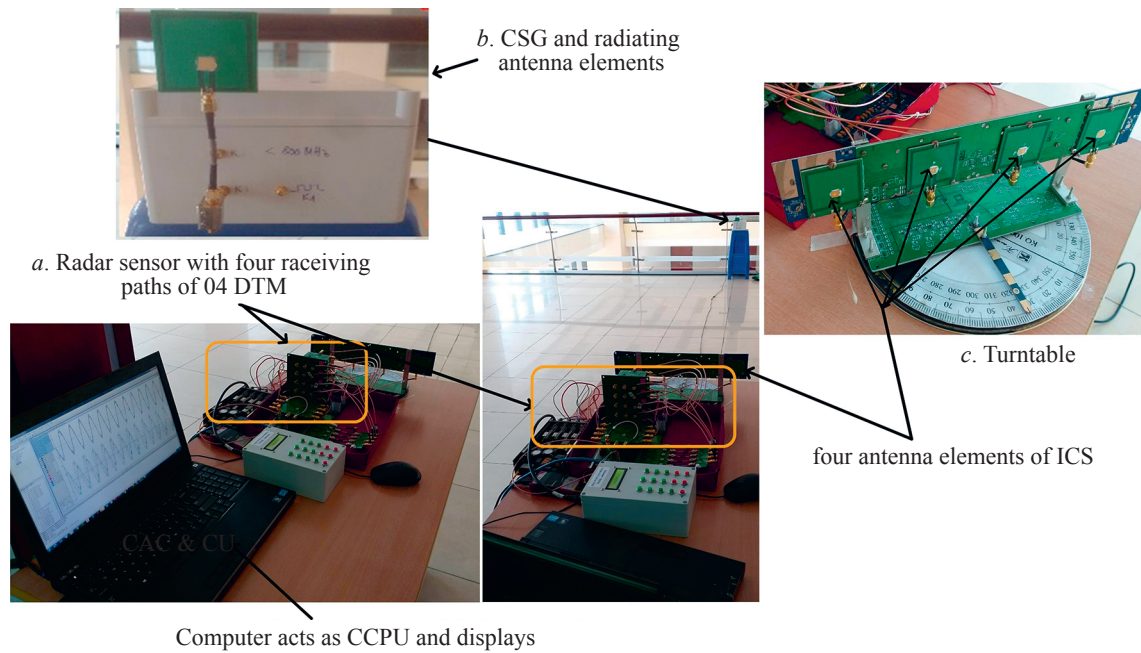


Fig. 8. Experimental setup for validating the proposed methodology

Table 4. Experimental results determine the signal amplitude and phase errors before improving the receiving path

State of the receiving path	Path No 2	Path No 3	Path No 4
Signal amplitude error (*), dB			
Before calibration	5.50	6.30	7.10
After calibration	0.91	1.34	1.53
Signal phase error (*)			
Before calibration	8°06'	9°06'	6°54'
After calibration	2°12'	2°42'	1°48'

Note. (*) Compared to the reference receiving path.

Table 5. Experimental results determine the signal amplitude and phase errors after improving the receiving path

State of the receiving path	Path No 2	Path No 3	Path No 4
Signal amplitude error (*), dB			
Before calibration	6.10	6.30	5.50
After calibration	0.41	0.34	0.25
Signal phase error (*)			
Before calibration	7°36'00"	6°48'00"	7°18'00"
After calibration	0°14'24"	0°12'36"	0°18'36"

Note. (*) Compared to the reference receiving path.

through the application of a fully matched filter model and dynamic range adaptive control. Based on the simulation and experimental results presented above, several observations can be made as follows:

- The proposed methodology for automatic measurement and calibration of the receiving paths in DTM has been validated both theoretically and experimentally. While experimental data exhibit some phase and amplitude errors compared to theoretical calculations and simulations; these discrepancies can be attributed to factors, such as coefficient estimation errors,

quantization noise, and hardware and software imperfections. These findings provide valuable insights for further improvement in research, design, and production processes.

- Accurate determination of errors in the calculation and establishment of calibration coefficients can be achieved through adaptive adjustments to the number of coherent accumulations across all receiving paths, as outlined in this study.
- In principle, the calibration process model can be characterized as “automatic, real-time, and integrated”.

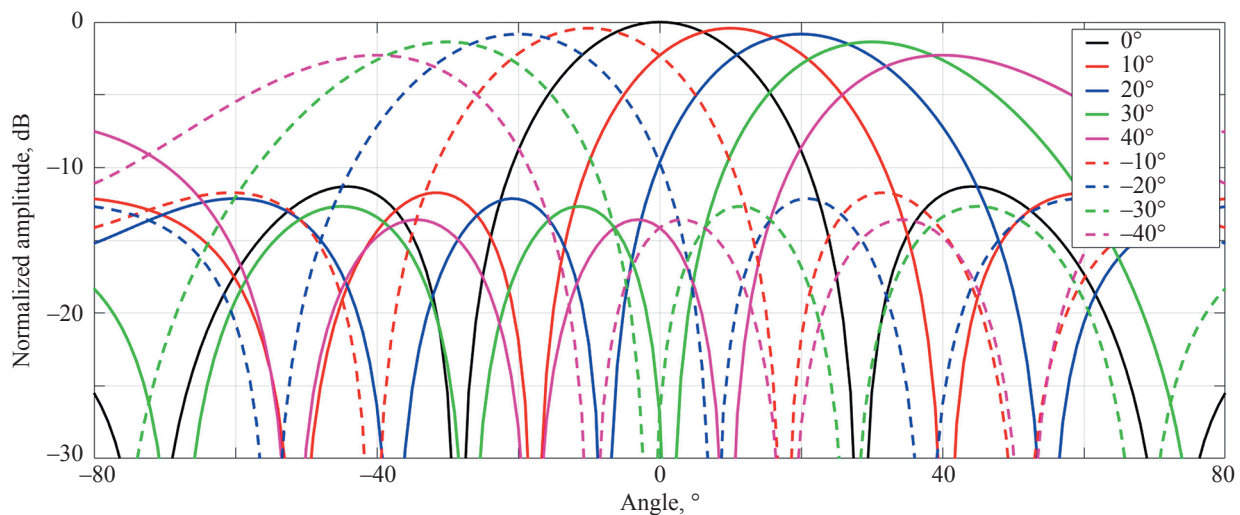


Fig. 9. The GP_{Σ} obtained through simulation after calibration

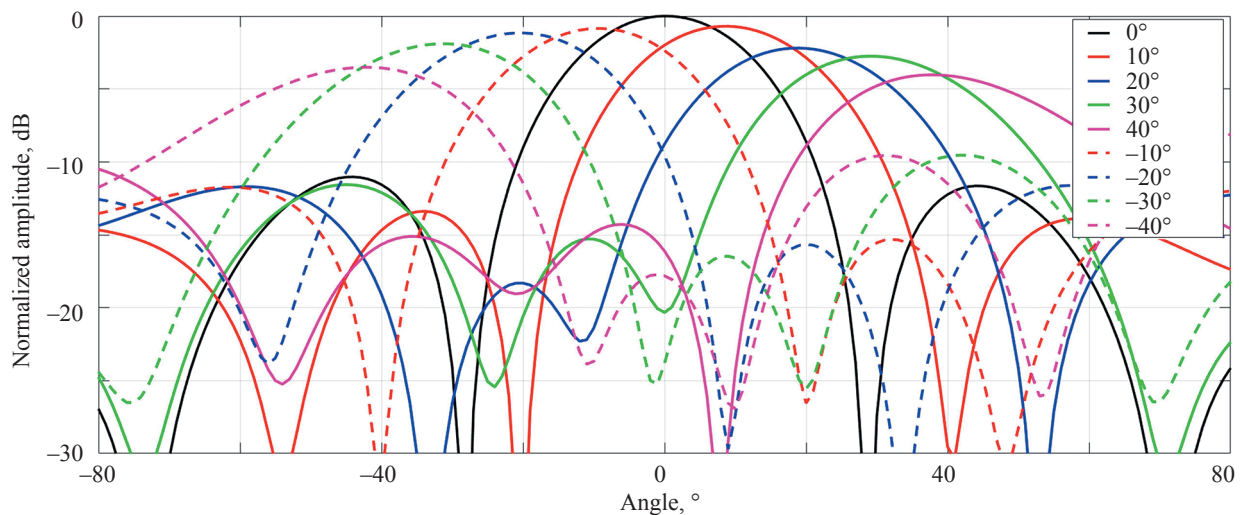


Fig. 10. The GP_{Σ} is synthesized after calibration and according to the measurement data: the GP_{Σ} are synthesized at angles: 0° , $\pm 10^\circ$, $\pm 20^\circ$, $\pm 30^\circ$ and $\pm 40^\circ$

This is because the measurement and estimation of errors are conducted entirely automatically and concurrently with the system operational processes. Furthermore, the hardware and software components are seamlessly integrated and implemented within the essential elements required for ICS in general and for each DTM specifically.

Conclusions

The primary focus of the paper is to propose a novel methodology for real-time automatic calibration of the Information and Control System (ICS), specifically targeting the receiving path in the Digital Transceiver Module (DTM). The innovation of this approach stems from leveraging the difference in power levels of reflected signals observed at the receiver input of a radar sensor using a conventional antenna and the input of the receiving path in the DTM of the ICS. To achieve accurate calibration, it is critical to ensure that the power level of the calibration signal aligns with the minimal power level of the reflected

signal. Accurate estimation of amplitude and phase errors in the calibration signal as it propagates through the receiving path requires coherent accumulation using appropriate accumulation coefficients. This process enhances precision in error quantification. The methodology introduces a virtual reference DTM model, synthesized using a “soft configuration” approach, to serve as a reference receiving path. This virtual model effectively replaces the physical reference DTM, mitigating potential errors that could arise during operation. Calibration coefficients are derived by comparing the complex signal amplitude of the receiving path in the virtual reference DTM with the complex signal amplitudes of other receiving paths after coherent accumulation. The calibration values for each module output are computed by multiplying the complex signal amplitude by the corresponding calibration coefficient. The synthesis for the gain patterns of the ICS are weighted summation of the calibrated complex signal amplitudes of the receiving paths. This synthesis ensures optimal performance by maximizing peak gain and minimizing side lobe levels in the system.

References

1. Kedar A. Phased array antenna for radar application. *Handbook of Metrology and Applications*, 2023, pp. 1443–1469. https://doi.org/10.1007/978-981-99-2074-7_81
2. Ttofis C., Papadopoulos A., Theocharides T., Michael M.K., Doumenis D. An MPSoC-based QAM modulation architecture with run-time load-balancing. *Eurasip Journal on Embedded Systems*, 2011, no. 1, pp. 790265. <https://doi.org/10.1155/2011/790265>
3. Fischer W. Basic principles of digital modulation. *Signals and Communication Technology*, 2010, no. 3 rd, pp. 219–260. https://doi.org/10.1007/978-3-642-11612-4_13
4. Lu Guoming, Zakharov P.N., Korolev A.F. Digital phased antenna array transceiver with multibeam radiation pattern. *Bulletin of the Russian Academy of Sciences: Physics*, 2023, vol. 87, no. 1, pp. 51–54. <https://doi.org/10.3103/S1062873822700125>
5. Babur G., Manokhin G.O., Monastirev E.A., Geltser A.A., Shibelgut A.A. Simple calibration technique for phased array radar systems. *Progress in Electromagnetics Research M*, 2017, vol. 55, pp. 109–119. <https://doi.org/10.2528/PIERM16101203>
6. Agrawal A., Jablon A. A calibration technique for active phased array antennas. *Proc. of the IEEE International Symposium on Phased Array Systems and Technology*, 2003, pp. 223–228. <https://doi.org/10.1109/PAST.2003.1256985>
7. Pan C., Ba X., Tang Y., Zhang F., Zhang Y., Wang Z., Fan W. Phased array antenna calibration method experimental validation and comparison. *Electronics*, 2023, vol. 12, no. 3, pp. 489. <https://doi.org/10.3390/electronics12030489>
8. Wang R., Gao P., Liu J., Wang Z., Wang C., Yu F. A hybrid scheme for TX I/Q imbalance self-calibration in a direct-conversion transceiver. *Electronics*, 2024, vol. 13, no. 9, pp. 1653. <https://doi.org/10.3390/electronics13091653>
9. Peng X., Wang Z., Mo J., Wang C., Liu J., Yu F. A blind calibration model for I/Q imbalances of wideband zero-IF receivers. *Electronics*, 2020, vol. 9, no. 11, pp. 1868. <https://doi.org/10.3390/electronics9111868>
10. Djigan V.I., Kurganov V.V. Antenna array calibration algorithm without access to channel signals. *Radioelectronics and Communications Systems*, 2020, vol. 63, no. 1, pp. 1–14. <https://doi.org/10.3103/S073527272001001X>
11. Lim A.G.K.C., Sreeram V., Wang G.-Q. Digital compensation in IQ modulators using adaptive FIR filters. *IEEE Transactions on Vehicular Technology*, 2004, vol. 53, no. 6, pp. 1809–1817. <https://doi.org/10.1109/TVT.2004.836934>
12. Tuthill J., Cantoni A. Efficient compensation for frequency-dependent errors in analog reconstruction filters used in IQ modulators. *IEEE Transactions on Communications*, 2005, vol. 53, no. 3, pp. 489–496. <https://doi.org/10.1109/tcomm.2005.843455>
13. He G., Gao X., Zhang R. Impact analysis and calibration methods of excitation errors for phased array antennas. *IEEE Access*, 2021, vol. 9, pp. 59010–59026. <https://doi.org/10.1109/ACCESS.2021.3073222>
14. Anttila L., Valkama M., Renfors M. Circularity-Based I/Q imbalance compensation in wideband direct-conversion receivers. *IEEE Transactions on Vehicular Technology*, 2008, vol. 57, no. 4, pp. 2099–2113. <https://doi.org/10.1109/TVT.2007.909269>
15. Li H., Liu A., Yang Q., Yu C., Lyv Z. Antenna pattern calibration method for phased array of high-frequency surface wave radar based on first-order sea clutter. *Remote Sensing*, 2023, vol. 15, no. 24, pp. 5789. <https://doi.org/10.3390/rs15245789>
16. Viet H.T., Minh T.H. A real-time internal calibration method for radar systems using digital phase array antennas. *Lecture Notes of the Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering*, 2021, vol. 379, pp. 88–103. https://doi.org/10.1007/978-3-030-77424-0_8
17. Nguyen X.L., Thi T.T.D., Nguyen P.B., Tran V.H. Receiving paths improvement of digital phased array antennas using adaptive dynamic range. *Electronics*, 2024, vol. 13, no. 21, pp. 4161. <https://doi.org/10.3390/electronics13214161>
18. Peter D. Hybrid beamforming receiver dynamic range theory to practice. *Microwave Product Digest & Technologies*. 2022.
19. Chen Y., Ming C., Xie K., Gao S., Jiang Q., Liu Z., Yao H., Dong K. All-in-One BPSK/QPSK switchable transmission and reception for adaptive free-space optical communication links. *Photonics*, 2024, vol. 11, no. 4, pp. 326. <https://doi.org/10.3390/photonics11040326>
20. Qi C., Wu L. PLL demodulation technique for M-ray Position Phase Shift Keying. *Journal of Electronics (China)*, 2009, vol. 26, no. 3, pp. 289–295. <https://doi.org/10.1007/s11767-008-0021-z>

Литература

1. Kedar A. Phased array antenna for radar application // *Handbook of Metrology and Applications*. 2023. P. 1443–1469. https://doi.org/10.1007/978-981-99-2074-7_81
2. Ttofis C., Papadopoulos A., Theocharides T., Michael M.K., Doumenis D. An MPSoC-based QAM modulation architecture with run-time load-balancing // *Eurasip Journal on Embedded Systems*. 2011. N 1. P. 790265. <https://doi.org/10.1155/2011/790265>
3. Fischer W. Basic principles of digital modulation // *Signals and Communication Technology*. 2010. N 3 rd. P. 219–260. https://doi.org/10.1007/978-3-642-11612-4_13
4. Lu Guoming, Zakharov P.N., Korolev A.F. Digital phased antenna array transceiver with multibeam radiation pattern // *Bulletin of the Russian Academy of Sciences: Physics*. 2023. V. 87. N 1. P. 51–54. <https://doi.org/10.3103/S1062873822700125>
5. Babur G., Manokhin G.O., Monastirev E.A., Geltser A.A., Shibelgut A.A. Simple calibration technique for phased array radar systems // *Progress in Electromagnetics Research M*. 2017. V. 55. P. 109–119. <https://doi.org/10.2528/PIERM16101203>
6. Agrawal A., Jablon A. A calibration technique for active phased array antennas // *Proc. of the IEEE International Symposium on Phased Array Systems and Technology*. 2003. P. 223–228. <https://doi.org/10.1109/PAST.2003.1256985>
7. Pan C., Ba X., Tang Y., Zhang F., Zhang Y., Wang Z., Fan W. Phased array antenna calibration method experimental validation and comparison // *Electronics*. 2023. V. 12. N 3. P. 489. <https://doi.org/10.3390/electronics12030489>
8. Wang R., Gao P., Liu J., Wang Z., Wang C., Yu F. A hybrid scheme for TX I/Q imbalance self-calibration in a direct-conversion transceiver // *Electronics*. 2024. V. 13. N 9. P. 1653. <https://doi.org/10.3390/electronics13091653>
9. Peng X., Wang Z., Mo J., Wang C., Liu J., Yu F. A blind calibration model for I/Q imbalances of wideband zero-IF receivers // *Electronics*. 2020. V. 9. N 11. P. 1868. <https://doi.org/10.3390/electronics9111868>
10. Djigan V.I., Kurganov V.V. Antenna array calibration algorithm without access to channel signals // *Radioelectronics and Communications Systems*. 2020. V. 63. N 1. P. 1–14. <https://doi.org/10.3103/S073527272001001X>
11. Lim A.G.K.C., Sreeram V., Wang G.-Q. Digital compensation in IQ modulators using adaptive FIR filters // *IEEE Transactions on Vehicular Technology*. 2004. V. 53. N 6. P. 1809–1817. <https://doi.org/10.1109/TVT.2004.836934>
12. Tuthill J., Cantoni A. Efficient compensation for frequency-dependent errors in analog reconstruction filters used in IQ modulators // *IEEE Transactions on Communications*. 2005. V. 53. N 3. P. 489–496. <https://doi.org/10.1109/tcomm.2005.843455>
13. He G., Gao X., Zhang R. Impact analysis and calibration methods of excitation errors for phased array antennas // *IEEE Access*. 2021. V. 9. P. 59010–59026. <https://doi.org/10.1109/ACCESS.2021.3073222>
14. Anttila L., Valkama M., Renfors M. Circularity-Based I/Q imbalance compensation in wideband direct-conversion receivers // *IEEE Transactions on Vehicular Technology*. 2008. V. 57. N 4. P. 2099–2113. <https://doi.org/10.1109/TVT.2007.909269>
15. Li H., Liu A., Yang Q., Yu C., Lyv Z. Antenna pattern calibration method for phased array of high-frequency surface wave radar based on first-order sea clutter // *Remote Sensing*. 2023. V. 15. N 24. P. 5789. <https://doi.org/10.3390/rs15245789>
16. Viet H.T., Minh T.H. A real-time internal calibration method for radar systems using digital phase array antennas // *Lecture Notes of the Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering*. 2021. V. 379. P. 88–103. https://doi.org/10.1007/978-3-030-77424-0_8
17. Nguyen X.L., Thi T.T.D., Nguyen P.B., Tran V.H. Receiving paths improvement of digital phased array antennas using adaptive dynamic range // *Electronics*. 2024. V. 13. N 21. P. 4161. <https://doi.org/10.3390/electronics13214161>
18. Peter D. Hybrid beamforming receiver dynamic range theory to practice // *Microwave Product Digest & Technologies*. 2022.
19. Chen Y., Ming C., Xie K., Gao S., Jiang Q., Liu Z., Yao H., Dong K. All-in-One BPSK/QPSK switchable transmission and reception for adaptive free-space optical communication links // *Photonics*. 2024. V. 11. N 4. P. 326. <https://doi.org/10.3390/photonics11040326>
20. Qi C., Wu L. PLL demodulation technique for M-ray Position Phase Shift Keying // *Journal of Electronics (China)*. 2009. V. 26. N 3. P. 289–295. <https://doi.org/10.1007/s11767-008-0021-z>

21. Duong V.M., Vesely J., Hubacek P., Janu P., Phan N.G. Detection and parameter estimation analysis of binary shift keying signals in high noise environments. *Sensors*, 2022, vol. 22, no. 9, pp. 3203. <https://doi.org/10.3390/s22093203>

21. Duong V.M., Vesely J., Hubacek P., Janu P., Phan N.G. Detection and parameter estimation analysis of binary shift keying signals in high noise environments // *Sensors*. 2022. V. 22. N 9. P. 3203. <https://doi.org/10.3390/s22093203>

Authors

Nguyen Trong Nhan — PhD, Researcher, Air Defense - Air Force Technical Institute, Hanoi, 11400, Vietnam, [sc 57222239467](https://orcid.org/0000-0001-6626-893X), <https://orcid.org/0000-0001-6626-893X>, 10th20th30th@gmail.com

Xuan Luong Nguyen — PhD Student, VNU University of Science, Hanoi, 100000, Vietnam, <https://orcid.org/0009-0000-6988-4738>, nguyenxuanluong_sdh22@hus.edu.vn

Phung Bao Nguyen — PhD, Associate Professor, Le Quy Don Technical University, Hanoi, 11917, Vietnam nguyenphungbao@lqdtu.edu.vn

Авторы

Нхан Нгуен Чонг — кандидат технических наук, исследователь, Технический институт ПВО, Ханой, 11400, Вьетнам, [sc 57222239467](https://orcid.org/0000-0001-6626-893X), <https://orcid.org/0000-0001-6626-893X>, 10th20th30th@gmail.com

Нгуен Суан Лыонг — аспирант, Ханойский государственный университет, Ханой, 100000, Вьетнам, <https://orcid.org/0009-0000-6988-4738>, nguyenxuanluong_sdh22@hus.edu.vn

Нгуен Фунг Бао — кандидат технических наук, доцент, преподаватель, Вьетнамский государственный технический университет имени Ле Куй Дона, Ханой, 11917, Вьетнам, nguyenphungbao@lqdtu.edu.vn

Received 22.01.2025

Approved after reviewing 07.03.2025

Accepted 24.05.2025

Статья поступила в редакцию 22.01.2025

Одобрена после рецензирования 07.03.2025

Принята к печати 24.05.2025



Работа доступна по лицензии
Creative Commons
«Attribution-NonCommercial»

doi: 10.17586/2226-1494-2025-25-3-565-573

УДК 004.95

Модель хранения пространственных данных тензорных геофизических полей

Гульнара Равилевна Воробьева^{1✉}, Андрей Владимирович Воробьев²,
Глеб Олегович Орлов³

^{1,2,3} Уфимский университет науки и технологий, Уфа, 450008, Российская Федерация

² Геофизический центр РАН, Москва, 119296, Российская Федерация

¹ Gulnara.vorobeva@gmail.com[✉], <https://orcid.org/0000-0001-7878-9724>

² geomagnet@list.ru, <https://orcid.org/0000-0002-9680-5609>

³ orlovgleb99@mail.ru, <https://orcid.org/0009-0003-5123-3859>

Аннотация

Введение. Известно, что геофизические поля (геомагнитное, гравитационное и электромагнитное) при их регистрации или моделировании представляют совокупность нескольких векторных составляющих, характеризующих изменение соответствующих параметров в пространстве и во времени. Хранение данных геофизических полей в настоящее время реализуется на основе известных моделей данных, представляющих собой, как правило, реляционную структуру. Анализ известных исследований показал избыточность и неэффективность такого подхода. Это выражается в низкой скорости получения искоемых данных при использовании комплексных мультипредикатных запросов. Непрерывно растущий объем и сложность рассматриваемых данных требуют новых подходов к организации их хранения для повышения производительности информационных систем, применяемых для поддержки принятия решений на основе данных геофизических полей. **Метод.** В данной работе предлагается и исследуется модель представления и хранения данных геофизических полей, обеспечивающая повышение производительности информационных систем. Представлен анализ специфических особенностей геофизических полей, обусловленных их тензорным характером. Рассмотрены основные компоненты данных, определены перспективные варианты комбинирования известных моделей данных для получения наилучшего результата для повышения производительности соответствующих баз данных. **Основные результаты.** Предложена мультиосевая модель данных геофизических полей, учитывающая тензорную многокомпонентную структуру полей и сочетающая особенности иерархической организации данных и элементоцентричной разметки информации. Отличительной особенностью предложенной модели является введение статичной и динамической осей. Такой подход обеспечивает представления метаданных, оперативных и архивных данных, взаимодействие между ними на уровне фоновых процессов с участием программных триггеров с темпоральными предикатами. **Обсуждение.** На примере данных геомагнитного поля и его вариаций продемонстрировано увеличение скорости выполнения одно- и мультипредикатных запросов на выборку данных, вставку новых записей в хранилище. Вычислительные эксперименты по сравнению предлагаемого и известных подходов к организации и хранению данных геофизических полей на различных наборах и объемах данных показали, что реализация мультиосевой модели данных дает возможность повысить скорость выполнения однопредикативных запросов на 25,7 %, мультипредикативных запросов на 20,1 %, запросов на вставку новых записей на 21,3 %. Это позволяет сделать вывод о целесообразности применения предложенного решения.

Ключевые слова

данные геофизических полей, модель данных, иерархическая модель данных, тензорное исчисление, геопространственные данные

Благодарности

Работа выполнена при поддержке Российского научного фонда (проект № 21-77-30010-П).

Ссылка для цитирования: Воробьева Г.Р., Воробьев А.В., Орлов Г.О. Модель хранения пространственных данных тензорных геофизических полей // Научно-технический вестник информационных технологий, механики и оптики. 2025. Т. 25, № 3. С. 565–573. doi: 10.17586/2226-1494-2025-25-3-565-573

Model for storing spatial data of tensor geophysical fields

Gulnara R. Vorobeva¹✉, Andrei V. Vorobev², Gleb O. Orlov³

^{1,2,3} Ufa University of Science and Technology, Ufa, 450008, Russian Federation

² Geophysical Center of the RAS, Moscow, 119296, Russian Federation

¹ gulnara.vorobeva@gmail.com✉, <https://orcid.org/0000-0001-7878-9724>

² geomagnet@list.ru, <https://orcid.org/0000-0002-9680-5609>

³ orlovgleb99@mail.ru, <https://orcid.org/0009-0003-5123-3859>

Abstract

It is known that geophysical fields (geomagnetic, gravitational and electromagnetic), when recorded or modeled, represent a set of several vector components characterizing the change in the corresponding parameters in space and time. Geophysical field data is currently stored based on known data models which usually have a relational structure. Analysis of known studies has shown the redundancy and inefficiency of this approach. This is reflected in the low speed of obtaining the desired data when using complex multi-predicate queries. The continuously growing volume and complexity of the data under consideration require new approaches to organizing their storage to improve the performance of information systems used to support decision-making based on geophysical field data. This paper proposes and examines a model for representing and storing geophysical field data that ensures increased performance of information systems. An analysis of specific features of geophysical fields due to their tensor nature is presented. The main data components are considered, promising options for combining known data models are determined to obtain the best result to improve the performance of the corresponding databases. A multi-axis model of geophysical field data is proposed that takes into account the tensor multi-component structure of the fields and combines the features of the hierarchical organization of data and element-centric information markup. A distinctive feature of the proposed model is the introduction of static and dynamic axes. This approach ensures the presentation of metadata, operational and archived data, and the interaction between them at the level of background processes with the participation of software triggers with temporal predicates. Using the example of geomagnetic field data and its variations, an increase in the speed of executing single- and multi-predicate queries for data selection and insertion of new records into the storage is demonstrated. Computational experiments comparing the proposed and known approaches to the organization and storage of geophysical field data on various sets and volumes of data showed that the implementation of the multi-axis data model allows increasing the speed of executing single-predicate queries by 25.7 %, multi-predicate queries by 20.1 %, and queries for inserting new records by 21.3 %. This allows us to conclude that the proposed solution is appropriate.

Keywords

geophysical field data, data model, hierarchical data model, tensor calculus, geospatial data

Acknowledgements

This work was funded by the Russian Science Foundation (project No. 21-77-30010-P).

For citation: Vorobeva G.R., Vorobev A.V., Orlov G.O. Model for storing spatial data of tensor geophysical fields. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2025, vol. 25, no. 3, pp. 565–573 (in Russian). doi: 10.17586/2226-1494-2025-25-3-565-573

Введение

В настоящее время результаты регистрации и моделирования параметров геофизических полей являются важным источником информации о сложных процессах, протекающих в недрах Земли, на ее поверхности и в околоземном пространстве. Такая информация используется как при решении научных задач (например, при исследовании сопутствующих процессов естественной и техногенной природы происхождения), так и для поддержки принятия решений в узкопрофильных прикладных областях (в тех случаях, когда воздействие геофизических полей значимо для функционирования объектов и систем техносферы).

Доступ к обозначенным данным, как правило, реализуется силами соответствующих владельцев геофизической информации, в качестве которых выступают различной ведомственной принадлежности научные организации и учреждения, располагающие необходимым для регистрации искомым параметров полей информационно-измерительным оборудованием. Так, к примеру, регистрация параметров геомагнитного поля на протяжении многих десятилетий осуществляется магнитными обсерваториями и вариационными стан-

циями, распределенными по всей земной поверхности (хотя и неравномерно) [1, 2]. Для предоставления потребителям возможности получения данных используются специализированные веб-ресурсы, которые в соответствии с predetermined протоколами доступа предоставляют по параметризованным (в основном) запросам искомые геофизические данные.

Отметим, что не всегда искомые данные являются результатом регистрации и наблюдения параметров реально протекающих геофизических процессов. В некоторых случаях требуется моделирование «идеального» геофизического поля в соответствующей геопространственной локации, что необходимо, в частности, для сопоставления полученных значений с зарегистрированными и оценки имеющего при этом места отклонения, что также служит важным источником информации при решении прикладных и/или научно-исследовательских задач. Например, параметры вектора главного (невозмущенного) геомагнитного поля могут быть рассчитаны с использованием различным моделей: WMM [3], IGRF [1, 4] и др.

Сложная организация геофизических данных и их значительный объем накладывают определенные ограничения на формат их представления таким образом,

чтобы не терять важной информации о поле, с одной стороны, и обеспечить высокую вычислительную скорость выполнения запросов к соответствующему хранилищу данных, с другой. В этой связи и с учетом известных форматов представления данных необходимо разработать модель представления данных геофизических полей, учитывающую их тензорный характер и обеспечивающую при программной реализации высокую производительность выполнения запросов к ним в условиях необходимости оперативного принятия решений на их основе (в том числе, в критических ситуациях, связанных, например, с вопросами обеспечения техносферной безопасности).

Характеристика исходных данных

Параметры геофизических полей, регистрируемые с заданным шагом дискретизации различными информационно-измерительными устройствами, аккумулируются в наборы геофизических данных, предоставляемых по запросу потребителям соответствующими поставщиками информации. Здесь и далее в качестве примера рассматриваются параметры геомагнитного поля и его вариаций, регистрируемые наземными магнитными обсерваториями и вариационными станциями международного проекта SuperMAG [5, 6].

Отличительной особенностью рассматриваемых полей является их тензорный характер, который в упрощенном виде может быть описан некоторой исходной (чаще пространственной) точкой с исходящими из нее векторами, в свою очередь, задающими определенные атрибутивные параметры рассматриваемого процесса. Фактически тензорное поле при этом представляет собой задаваемую функцией поверхность, каждой точке которой ставится в соответствие тензор, привязанный к соответствующей системе координат и имеющий начало в заданной пространственной точке [7–10].

Количество исходящих векторов определяет ранг тензора [11]. Данные о состоянии электромагнитного поля Земли в каждой точке трехмерного пространства в системе координат, связанной с центром масс Земли, представляют собой прямое произведение тензора электромагнитного поля на радиус-вектор исследуемой точки. В связи с этим геопространственные данные о состоянии магнитного поля Земли представляют собой прямое произведение тензора второго порядка на тензор первого порядка, т. е. описываются при помощи тензора третьего порядка [12].

Визуализация рассматриваемого тензора (тензорного поля), как правило, выполняется посредством декомпозиции его векторов на скалярные величины, каждой из которых ставится в соответствие свой графический примитив. Так, например, в случае использования комплекса геоинформационных средств и технологий для визуализации геомагнитного поля в качестве геопространственных примитивов, применяются пространственные точки с соответствующими атрибутивными значениями. Множество формируемых таким образом геопространственных точек формируют в своей совокупности геопространственный слой, который в дальнейшем визуализируется на некоторой

картографической подложке [2, 4]. Возможна одновременная визуализация многих накладываемых друг на друга слоев, имитирующих различные вектора тензора геомагнитного поля, что в подавляющем большинстве случаев приводит к повышению сложности визуального восприятия итогового геопространственного изображения, а также накладывает определенные ограничения на совместное использование для рассматриваемых данных геоинформационных инструментов.

Целесообразно учитывать специфику представления данных тензорных полей, чтобы для оптимизации хранения и вариативности представления искомых параметров поля применять и осевую структуру соответствующего тензора, который для совокупности пространственных точек также демонстрирует картину пространственно-временного распределения соответствующих параметров.

Мультиосевая модель хранения данных тензорных полей

Применим комплексную структуру параметров тензорного поля в качестве основы для создания модели представления и хранения соответствующих данных. В этой связи предлагается структура, условно названная мультиосевой. Центральным ее элементом является источник геопространственной/геофизической информации, для которого осуществляется хранение регистрируемой информации. При этом в качестве обозначенного центрального компонента мультиосевой модели может выступать как источник, обеспечивающий регистрацию искомых параметров в режиме реального или квазиреального времени, так и некоторая цифровая модель, выполняющая расчет тех же атрибутивных параметров. Например, в случае геомагнитных данных, в качестве центрального элемента модели может выступать магнитная обсерватория/вариационная станция, в режиме реального времени осуществляющая регистрацию параметров магнитного поля Земли и его вариаций. С другой стороны, в случае физического отсутствия магнитометра в той или иной пространственной точке, в качестве рассматриваемого компонента модели данных выступает некоторая модель или цифровой двойник магнитной обсерватории/вариационной станции, моделирующей расчет аналогичных параметров в рассматриваемой пространственной точке.

Для каждого центрального звена модели данных поля формируются три оси данных, представляющие собой независимые пары вида «атрибут–значение».

Количество пар данных при этом неограничено. С учетом специфики данных рассматриваемых тензорных полей предлагается введение трех основных осей, логически связанных с обозначенным центральным элементом/звеном модели данных (рис. 1):

— статическая ось метаданных, предназначенная для хранения информации об источнике данных поля. Предположим, если рассматривается мультиосевая модель геомагнитных данных, в которой в качестве центрального элемента выступает магнитная обсерватория/вариационная станция, то на оси метаданных размещается информация: название обсервато-

- рии, ее географические координаты, ведомственная принадлежность, используемое оборудование, шаг дискретизации, единицы измерения, принятые допущения и пр.;
- динамическая ось оперативных данных, обеспечивающая структуру для хранения данных, актуальных на текущий момент времени в соответствии с шагом дискретизации источника, который определен на статичной оси метаданных. Для текущего момента времени результаты регистрации или расчета параметров геофизического поля декомпозируются на отдельные независимые пары вида «атрибут–значение». При этом значение атрибута может отсутствовать (если по тем или иным причинам регистрация или расчет соответствующего параметра поля оказались невозможны). Параметры, характеризующие текущий момент времени, для которого определены соответствующие атрибуты с их значениями, задаются в узле типа «Временная метка», который является родительским для всех узлов, представляющих атрибуты поля и их значения;
- статичная ось архивных данных, предназначенная для создания структуры хранения данных в моменты времени, следующие за текущим. По истечении заданного в метаданных шага дискретизации ветка динамической оси оперативных данных перемещается в статичную ось архивных данных, становясь уже новой веткой последней. Таким образом, на оси архивных данных размещается множество последо-

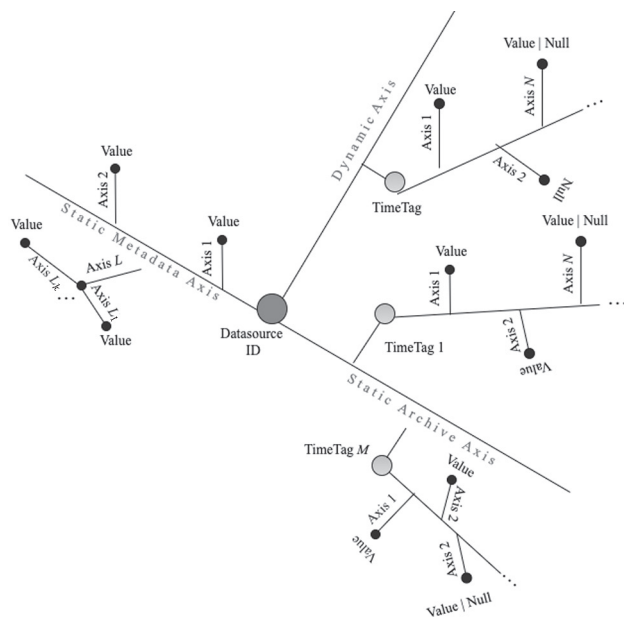


Рис. 1. Обобщенная схема мультиосевой модели хранения данных тензорных геофизических полей: L — количество атрибутов-метаданных; L_1, \dots, L_k — значения каждого из k атрибутов L -й оси метаданных; N — количество атрибутов поля; M — количество зарегистрированных моментов времени для параметров поля

Fig. 1. Generalized scheme of a multilevel model for storing data of tensor geophysical fields: L — number of metadata attributes; L_1, \dots, L_k — values of each of the k attributes of the L -th metadata axis; N — number of field attributes; M — number of registered time points for field parameters

вательных узлов вида «Временная метка», каждый из которых содержит свою ветку дочерних узлов с соответствующими параметрами поля и их значениями.

При этом отметим, что предлагаемая структура предполагает атомарную структуру значения атрибута. В случае комплексного параметра, предполагающего множество значений, необходима декомпозиция атрибута на отдельные составляющие с атомарными значениями.

Статичная и динамическая оси данных предполагают формирование вложенной иерархии в соответствии с рангом геофизическим полем, значения параметров которого необходимо представить. Каждому параметру ставится в соответствие отдельная ось данных, являющаяся дочерней по отношению к родительской оси, в качестве которой выступает непосредственно ось, соответствующая временной метке. Так, в частности, разложение тензора третьего порядка на пары «атрибут–значение» происходит следующим образом. Пусть тензор третьего порядка представляет собой множество данных, где каждый элемент определяется тремя индексами. Например, для геомагнитного поля могут быть использованы компоненты поля по трем осям (X, Y, Z), временные производные этих компонент или взаимные корреляции между ними (рис. 2). Для формирования предлагаемой структуры для каждого момента времени создается узел вида «Временная метка», который служит родительским элементом. Далее каждый компонент тензора становится отдельным атрибутом, а значение компонента становится значением атрибута. Выделенные атрибуты группируются под родительским узлом «Временная метка». При этом в контексте геомагнитного поля отметим, что:

- каждый атрибут имеет четкое физическое или математическое значение;
- значения могут быть как числовыми, так и отсутствовать (null);
- структура остается неизменной для каждого момента времени.

Все атрибуты привязаны к конкретной временной метке. После завершения текущего момента времени вся ветка с атрибутами перемещается в статичную ось архивных данных, где становится одним из множества исторических узлов.

Данные статичной оси не меняются или меняются крайне редко. В связи с этим физическое представление

```

Временная метка (2025-03-17 16:24:42)
├─ X_component: 25672 нТл
├─ Y_component: 12345 нТл
├─ Z_component: -4321 нТл
├─ dX/dt: 0.5 нТл/с
├─ dY/dt: -0.3 нТл/с
├─ dZ/dt: 0.1 нТл/с
├─ X_Y_correlation: 0.75
├─ X_Z_correlation: -0.2
└─ Y_Z_correlation: 0.6
    
```

Рис. 2. Пример структуры данных геомагнитного поля для заданного момента времени

Fig. 2. An example of the data structure for a geomagnetic field at a given moment in time

этих данных должно быть организовано таким образом, чтобы максимальная вычислительная скорость достигалась при извлечении данных в большей степени, чем при внесении новых или редактировании уже внесенных значений параметров метаданных.

В противоположность статичной динамическая ось подразумевает оперативное изменение и/или добавление хранимых данных. При регистрации/расчете искомым атрибутивных значений они должны с минимальной задержкой быть представлены на динамической оси таким образом, чтобы участвовать в дальнейшем в оперативном принятии решений на этой основе. Соответственно, физическая реализация рассматриваемой оси предполагает необходимость обеспечения максимально возможной вычислительной скорости на добавление и непосредственно извлечение искомым значений атрибутивных параметров.

Рассмотрим структуру связей между узлами и их привязку к осям на примере данных геомагнитного поля, для некоторого момента времени (рис. 3). Здесь центральный элемент является корневым узлом, от которого берут начало три основные оси: статичная ось метаданных (не меняется во времени, отражает постоянную информацию об источнике данных, технические характеристики оборудования, параметры сбора данных); динамическая ось оперативных данных (отражает текущие измерения; для текущего момента имеет корневой узел вида «Временная метка», от него отходят узлы с параметрами поля; ось обновляется в соответствии с шагом дискретизации); статичная ось архивных данных (отражает исторические данные; содержит некоторую последовательность узлов вида «Временная метка» по аналогии с динамической осью, однако добавление данных осуществляется по мере истечения текущего периода).

```
graph TD
    A[Центральный элемент] --> B[Статичная ось метаданных]
    A --> C[Динамическая ось оперативных данных]
    A --> D[Статичная ось архивных данных]

    B --> B1[Название обсерватории]
    B --> B2[Географические координаты]
    B --> B3[Ведомственная принадлежность]
    B --> B4[Используемое оборудование]
    B --> B5[Шаг дискретизации]
    B --> B6[Единицы измерения]

    C --> C1["Временная метка (текущая)"]
    C1 --> C2["X_component: 25672 нТл"]
    C1 --> C3["Y_component: 12345 нТл"]
    C1 --> C4["Z_component: -4321 нТл"]
    C1 --> C5["dX/dt: 0.5 нТл/с"]

    D --> D1["Временная метка (t-1)"]
    D1 --> D2["X_component: 25668 нТл"]
    D1 --> D3["Y_component: 12342 нТл"]
    D1 --> D4["Z_component: -4323 нТл"]
    D --> D5["Временная метка (t-2)"]
    D5 --> D6["X_component: 25670 нТл"]
```

Рис. 3. Пример распределения данных геомагнитного поля по статичной и динамической осям данных и оси метаданных (в общем виде)

Fig. 3. An example of the distribution of geomagnetic field data along the static and dynamic data axes and the metadata axis (in general)

В рамках предлагаемой модели предполагается взаимодействие между статичной осью архивных данных и динамической осью оперативных данных, организованное таким образом, что данные последней перемещаются в рассматриваемую ось по умолчанию без каких-либо изменений (рис. 4). При этом при необходимости возможно введение дополнительного узла в рассматриваемую модель, обеспечивающую обработку и трансформацию данных динамической оси в данные статичной.

Перемещение узла динамической оси в статичную осуществляется в автоматическом режиме с использованием специализированного триггерного механизма. Последний является параметризованным и настраиваемым, поскольку в своем функционировании опирается на значение шага дискретизации, который, в свою очередь, должен быть либо задан на статичной оси метаданных соответствующего узла, либо определен программными настройками триггера, определяющими «время жизни» узла динамической оси оперативных данных. Отметим, что в каждый момент времени при этом на динамической оси оперативных данных возможно присутствие одного и только одного узла с

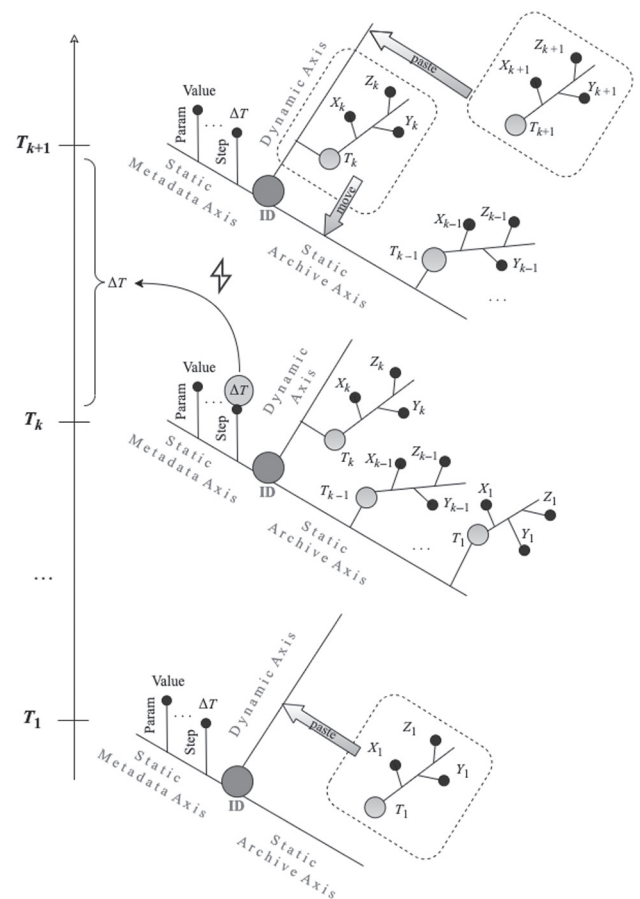


Рис. 4. Схема взаимодействия между статичной и динамической осями модели данных: T_i — i -й момент времени; k — порядковый номер момента времени; ΔT — временной интервал; ID — идентификатор узла

Fig. 4. Diagram of interaction between static and dynamic axes of the data model: T_i — i -th moment in time; k — serial number of the moment in time; ΔT — time interval; ID — node identifier

заданной временной меткой. При этом на статичной оси архивных данных количество узлов не ограничено ни верхней, ни нижней планкой. Единственное условие, которое обязательно должно быть соблюдено при реализации модели данных, заключается в том, чтобы обеспечить отсутствие дубликатов узлов с одноименными временными метками на одной статичной оси для заданного экземпляра центрального элемента соответствующей модели.

Рассмотрим функционирование параметризованного триггерного механизма на примере системы мониторинга геомагнитного поля с шагом дискретизации 1 ч. В начальный момент времени $t_0 = 14:00$ формируется базовая конфигурация системы, включающая центральный элемент, связанный с метаданными и динамической осью, содержащей, в свою очередь, узел с временной меткой t_0 с компонентами геомагнитного поля (X, Y, Z). В общем виде триггерный механизм базируется на некоторой функции текущего времени $T(t)$, в которой временной параметр определяется шагом дискретизации Δt , т. е. для n -го шага имеет место соотношение вида: $t_n = t_0 + n\Delta t$. Активация триггерного механизма происходит при достижении условия $T(t) = t_n$. В продолжение примера в момент времени $t_1 = 15:00$ происходит последовательная активация триггерного механизма, включающая проверку условия $t_1 = t_0 + \Delta t$, архивацию узла с меткой t_0 в статичную ось и создание нового узла t_1 в динамической оси. После срабатывания триггера структура данных трансформируется, образуя иерархическую систему, где центральный элемент связан как с динамической осью, содержащей актуальный узел t_1 , так и со статичной осью, включающей архивный узел t_0 . При последующем достижении момента времени $t_2 = 16:00$ процесс повторяется, что приводит к дальнейшему расширению статичной оси за счет добавления узла t_1 и обновлению динамической оси новым узлом t_2 .

Статичная ось метаданных предназначена для представления основных данных об источнике и меняется крайне редко. Отметим, что данные статичной оси метаданных являются основным источником геопространственной информации, характеризующей соответствующую геопространственную привязку источника (в виде пары географических координат — широты и долготы) или параметров поля (значения параметров поля обязательно должны рассматриваться в контексте соответствующей геопространственной точки или области).

Таким образом, функционирование системы обеспечивает детерминированное перемещение данных между осями, поддержание целостности временных рядов, автоматическое масштабирование хранения информации при сохранении геопространственной привязки и обеспечении уникальности временных меток.

Особенности программной реализации мультиосевой модели данных

В настоящее время данные геофизических полей традиционно представлены в хранилищах, имеющих реляционную структуру [13–15]. Вместе с тем тензорный характер данных, а также практически статичный

набор метаданных предполагают необходимость введения иного представления информации, учитывающего перечисленные особенности. Применение реляционной структуры представляет избыточные метаданные соответствующих таблиц и связей между ними, что увеличивает расходы дискового пространства. Кроме того, искомая информация распределена по многим нормализованным таблицам соответствующей реляционной базы, что в общем случае приводит к значительному увеличению сложности запросов, поскольку данные необходимо собрать из множества разных таблиц. При этом необходимые для выполнения запросов операции соединения таблиц требуют больших затрат вычислительных ресурсов, что в целом также может негативно сказаться на производительности.

Представляется целесообразным для реализации предложенной модели использовать формат, учитывающий древовидную структуру данных, а также геопространственную и временную привязки рассматриваемых при этом параметров поля. В настоящей работе для представления и обработки различных данных из прикладных областей был выбран формат XML (eXtensible Markup Language, расширяемый язык разметки) [16–18].

Каждый источник данных с уникальным (в том числе суррогатным идентификатором), либо геопространственной привязкой (также для удобства сопровождаемая некоторым идентификатором) является корнем соответствующего дерева, а оси метаданных и данных (статичная и динамическая) — дочерними узлами. Все пары данных и метаданных формата «атрибут–значение» являются соседними между собой и относятся к одному и тому же родительскому узлу первого уровня, соответствующему статичной или динамической оси данных или метаданных соответственно.

Для формализации структуры соответствующего XML-документа выполнено ее описание в виде XML-схемы формата XSD [19]. Применение данной схемы позволило контролировать процесс формирования документа таким образом, чтобы не допустить внесения избыточных изменений в итоговую структуру, и не изменить элементоцентричную архитектуру документа. Проверка документа на валидность (соответствие XSD-схеме) должна выполняться каждый раз при формировании новой структуры (появлении нового источника, регистрации новых данных, изменении узлов статичной оси архивных данных и пр.).

Запрос к данным полей по пространственно-временным параметрам предполагает перемещение поискового курсора по осям соответствующей модели данных. С учетом XML-структуры документа, реализующего рассматриваемую модель применительно к геофизическим полям, запросы целесообразно реализовывать в виде конструкций на основе спецификаций XPath [20, 21], учитывающих древовидную организацию данных в XML-документе и соответствующей модели.

Оценка эффективности предложенного решения

Оценка эффективности предложенного решения по организации хранения данных геофизических полей

на основе мультиосевой модели была выполнена посредством сравнения выделенных метрик качества с иными практикуемыми в настоящее время подходами к хранению аналогичных данных. Для проведения вычислительных экспериментов был сформирован набор геомагнитных данных, зарегистрированных с шагом дискретизации в 1 с магнитными обсерваториями и вариационными станциями проекта SuperMAG, насчитывающем в общей сложности более 350 источников данных.

Для сравнительной оценки в качестве основных метрик качества были определены скорость получения данных из хранилища по одному и нескольким пространственно-временным параметрам (предикатам) соответственно (однопредикативные и мультипредикативные запросы), а также скорость внесения новых геомагнитных данных в те же хранилища. Вычислительные эксперименты выполнялись на примере предложенного решения, а также реляционного представления геомагнитных данных (совокупности связанных таблиц, построенных с использованием системы управления базами данных (СУБД) PostgreSQL) и представления данных в виде совокупности текстовых CSV-файлов, доступ к которым предоставляется посредством специализированного прикладного программного интерфейса (Application programming interface, API) (сервис проекта SuperMAG для языка программирования Python). Мультиосевая модель данных геофизических полей была реализована в XML-формате под управлением СУБД Sedna [22, 23].

Для проведения вычислительных экспериментов был разработан исследовательский прототип веб-ориентированного приложения, обеспечивающего внесение геомагнитных данных в хранилище заданного типа и извлечение данных по заданным параметрам. Поскольку задача визуализации данных при этом не рассматривалась, результат выполнения запросов был представлен простым вариантом загрузки сформированного файла KML/KMZ [24, 25] на клиентской стороне. Исследования проводились на клиентской стороне

с применением персонального компьютера (CPU Intel Core i5 10300H ГГц, оперативная память 4 Гб, скорость интернет-соединения ~52,4 Мбит/с и на серверной стороне – на базе веб-сервера с процессором 72* Intel(R) Xeon(R) Gold 6140 CPU @ 2.30GHz).

В результате проведения вычислительных экспериментов на различных наборах и объемах данных геофизических полей было установлено, что применение мультиосевой модели данных для хранения рассматриваемых данных позволит повысить скорость выполнения однопредикативных запросов на 25,7 %, мультипредикативных запросов на 20,1 %, запросов на вставку новых записей в среднем на 21,3 %, что подтверждает целесообразность реализации предложенного решения.

Заключение

Данные геофизических полей в настоящее время являются основным источником информации о сложных процессах, протекающих на земной поверхности и за ее пределами, а также широко используются в задачах информационной поддержки принятия решений. В этой связи актуальна задача совершенствования известных подходов к организации и хранению данных геофизических полей таким образом, чтобы обеспечить повышение производительности различных запросов к соответствующим базам данных.

В работе предложена мультиосевая модель данных геофизических полей, основанная на иерархических структурах данных и элементоцентричной разметке информации, отличающаяся тем, что в соответствии с тензорной многокомпонентной структурой полей вводятся статическая и динамическая оси, которые берут начало в центральном узле, соответствующем источнику данных, и предназначены для представления метаданных, оперативных и архивных данных, взаимодействие между которыми осуществляется на уровне фоновых процессов с участием программных триггеров с темпоральными предикатами.

Литература

1. Reay S.J., Herzog D.C., Alex S., Kharin E.P., McLean S., Nosé M., Sergeyeva N.A. Magnetic observatory data and metadata: types and availability // *Geomagnetic Observations and Models*. 2011. V. 5. P. 149–181. https://doi.org/10.1007/978-90-481-9858-0_7
2. St-Louis B. *Intermagnet Technical Reference Manual*. Version 4.6. Edinburgh: INTERMAGNET, BGS, 2012. 100 p.
3. Califf S., Alken P., Chulliat A., Anderson B., Rock K., Vines S., Barnes R., Liou K. Investigation of geomagnetic reference models based on the Iridium® constellation // *Earth, Planets and Space*. 2022. V. 74. N 1. P. 37. <https://doi.org/10.1186/s40623-022-01574-w>
4. Lazzeri C., Samsonov A., Forsyth C., Branduardi-Raymont G., Bogdanova Y. A statistical study of the properties of, and geomagnetic responses to, large, rapid southward turnings of the interplanetary magnetic field // *Journal of Geophysical Research: Space Physics*. 2024. V. 129. N 9. P. e2023JA032160. <https://doi.org/10.1029/2023JA032160>
5. Gjerloev J.W. The SuperMAG data processing technique // *Journal of Geophysical Research: Space Physics*. 2012. V. 117. N 9. P. A09213. <https://doi.org/10.1029/2012JA017683>
6. Janzhura A.S., Troshichev O.A. Determination of the running quiet daily geomagnetic variation // *Journal of Atmospheric and Solar-*

References

1. Reay S.J., Herzog D.C., Alex S., Kharin E.P., McLean S., Nosé M., Sergeyeva N.A. Magnetic observatory data and metadata: types and availability. *Geomagnetic Observations and Models*, 2011, vol. 5, pp. 149–181. https://doi.org/10.1007/978-90-481-9858-0_7
2. St-Louis B. *Intermagnet Technical Reference Manual*. Version 4.6. Edinburgh: INTERMAGNET, BGS, 2012, 100 p.
3. Califf S., Alken P., Chulliat A., Anderson B., Rock K., Vines S., Barnes R., Liou K. Investigation of geomagnetic reference models based on the Iridium® constellation. *Earth, Planets and Space*, 2022, vol. 74, no. 1, pp. 37. <https://doi.org/10.1186/s40623-022-01574-w>
4. Lazzeri C., Samsonov A., Forsyth C., Branduardi-Raymont G., Bogdanova Y. A statistical study of the properties of, and geomagnetic responses to, large, rapid southward turnings of the interplanetary magnetic field. *Journal of Geophysical Research: Space Physics*, 2024, vol. 129, no. 9, pp. e2023JA032160. <https://doi.org/10.1029/2023JA032160>
5. Gjerloev J.W. The SuperMAG data processing technique. *Journal of Geophysical Research: Space Physics*, 2012, vol. 117, no. 9, pp. A09213. <https://doi.org/10.1029/2012JA017683>
6. Janzhura A.S., Troshichev O.A. Determination of the running quiet daily geomagnetic variation. *Journal of Atmospheric and Solar-*

- Terrestrial Physics, 2008. V. 70. N 7. P. 962–972. <https://doi.org/10.1016/j.jastp.2007.11.004>
7. Peng Z., Laramée S. Higher dimensional vector field visualization. a survey // *Proc. of the Theory and Practice of Computer Graphics (TPCG '09)*, 2009. P. 149–163.
8. Hergl C., Blecha C., Kretzschmar V., Raith F., Gunther F., Stommel M., Jankowai J., Hotz I., Nagel T., Scheuermann G. Visualization of tensor fields in mechanics // *Computer Graphics Forum*, 2021. V. 40. N 6. P. 135–161. <https://doi.org/10.1111/cgf.14209>
9. Jilesh V., Pournami A. On a generalization of Laplace distribution with applications // *International Journal of Data Science and Analytics*, 2025. in press. <https://doi.org/10.1007/s41060-024-00706-7>
10. He Z., Hu X., Teng Y., Zhang X., Shen X. Data agreement analysis and correction of comparative geomagnetic vector observations // *Earth, Planets and Space*, 2022. V. 74. N 1. P. 29. <https://doi.org/10.1186/s40623-022-01583-9>
11. Huang Y., Wu L., Li D. Theoretical research on full attitude determination using geomagnetic gradient tensor // *The Journal of Navigation*, 2015. V. 68. N 5. P. 951–961. <https://doi.org/10.1017/S0373463315000259>
12. Oliva P., Xu C. On the Herbrand functional interpretation // *Mathematical Logic Quarterly*, 2020. V. 66. N 1. P. 91–98. <https://doi.org/10.1002/malq.201900067>
13. Воробьев А.В., Воробьева Г.Р. Индуктивный метод восстановления временных рядов геомагнитных данных // *Труды СПИИРАН*, 2018. № 2(57). С. 104–133. <https://doi.org/10.15622/sp.57.5>
14. Воробьев А.В., Воробьева Г.Р. Оценка влияния геомагнитной активности на метрологические характеристики инклинометрических информационно-измерительных систем // *Измерительная техника*, 2017. № 6. С. 21–24.
15. Воробьев А.В., Воробьева Г.Р., Юсупова Н.И. Концепция единого пространства геомагнитных данных // *Труды СПИИРАН*, 2019. Т. 18. № 2. С. 390–415. <https://doi.org/10.15622/sp.18.2.390-415>
16. Muhammad A., Minhaj K. Enhancing XML data parsing and querying performance on multi-core architectures // *Journal of Statistics, Computing and Interdisciplinary Research*, 2024. V. 6. N 1. P. 75–89. <https://doi.org/10.52700/scir.v6i1.158>
17. Воробьев А.В., Воробьева Г.Р. Подход к обнаружению и устранению артефактов пространственных изолиний в приложениях Веб-ГИС // *Компьютерная оптика*, 2023. Т. 47. № 1. С. 126–136. <https://doi.org/10.18287/2412-6179-CO-1127>
18. Воробьев А.В., Воробьева Г.Р. Подход к динамической визуализации разнородных геопро пространственных векторных изображений // *Компьютерная оптика*, 2024. Т. 48. № 1. С. 123–138. <https://doi.org/10.18287/2412-6179-CO-1279>
19. Shene J., Chnoor R., Mzhda A. XML schema validation using Java API for XML processing // *UKH Journal of Science and Engineering*, 2022. V. 6. N 1. P. 33–41. <https://doi.org/10.25079/ukhjse.v6n1y2022>
20. Chen R., Wang Z., Su H., Xie S., Wang Z. Parallel XPath query based on cost optimization // *The Journal of Supercomputing*, 2022. V. 78. N 4. P. 5420–5449. <https://doi.org/10.1007/s11227-021-04074-y>
21. Areces C., Fervari R. Axiomatizing Hybrid XPath with data // *Logical Methods in Computer Science*, 2021. V. 17. N 3. P. 5. [https://doi.org/10.46298/lmcs-17\(3:5\)2021](https://doi.org/10.46298/lmcs-17(3:5)2021)
22. Chernenkiy V.M., Gapanyuk Y.E., Kaganov Y.T., Dunin I.V., Lyaskovsky M.A., Larionov V.S. Storing metagraph model in relational, document-oriented, and graph databases // *CEUR Workshop Proceedings*, 2018. V. 2277. P. 82–89.
23. Terekhov V., Gapanyuk Y., Kanev A. Metagraph representation for overcoming limitations of existing knowledge bases // *Proc. of the 28th Conference of Open Innovations Association (FRUCT)*, 2021. P. 458–464. <https://doi.org/10.23919/fruct50888.2021.9347601>
24. Zichar M. Geovisualization based upon KML // *Journal of Agricultural Informatics*, 2012. V. 3. N 1. P. 19–26. <https://doi.org/10.17700/jai.2012.3.1.49>
25. Ortiz-Alvarez A., Ruiz-Suárez L.G., Luis G., Ortega E., García-Reynoso A., Peralta O., López-Gaona A., Castro, T., Martínez-Arroyo A. Emission inventory point source visualization on Google Earth and integrated with HYSPLIT model // *Atmósfera*, 2021. V. 34. N 2. P. 143–156. <https://doi.org/10.20937/ATM.52834>
- Terrestrial Physics*, 2008, vol. 70, no. 7, pp. 962–972. <https://doi.org/10.1016/j.jastp.2007.11.004>
7. Peng Z., Laramée S. Higher dimensional vector field visualization. a survey. *Proc. of the Theory and Practice of Computer Graphics (TPCG '09)*, 2009, pp. 149–163.
8. Hergl C., Blecha C., Kretzschmar V., Raith F., Gunther F., Stommel M., Jankowai J., Hotz I., Nagel T., Scheuermann G. Visualization of tensor fields in mechanics. *Computer Graphics Forum*, 2021, vol. 40, no. 6, pp. 135–161. <https://doi.org/10.1111/cgf.14209>
9. Jilesh V., Pournami A. On a generalization of Laplace distribution with applications. *International Journal of Data Science and Analytics*, 2025. in press. <https://doi.org/10.1007/s41060-024-00706-7>
10. He Z., Hu X., Teng Y., Zhang X., Shen X. Data agreement analysis and correction of comparative geomagnetic vector observations. *Earth, Planets and Space*, 2022, vol. 74, no. 1, pp. 29. <https://doi.org/10.1186/s40623-022-01583-9>
11. Huang Y., Wu L., Li D. Theoretical research on full attitude determination using geomagnetic gradient tensor. *The Journal of Navigation*, 2015, vol. 68, no. 5, pp. 951–961. <https://doi.org/10.1017/S0373463315000259>
12. Oliva P., Xu C. On the Herbrand functional interpretation. *Mathematical Logic Quarterly*, 2020, vol. 66, no. 1, pp. 91–98. <https://doi.org/10.1002/malq.201900067>
13. Vorobev A., Vorobeva G. Inductive method of geomagnetic data time series recovering. *SPIIRAS Proceedings*, 2018, vol. 2, no. 57, pp. 104–133. (in Russian). <https://doi.org/10.15622/sp.57.5>
14. Vorob'ev, A.V., Vorob'eva, G.R. Evaluation of the influence of geomagnetic activity on metrological characteristics of inclinometric information measuring systems. *Measurement Techniques*, 2017, vol. 60, no. 6, pp. 546–551. <https://doi.org/10.1007/s11018-017-1232-1>
15. Vorobev A., Vorobeva G., Yusupova N. Conception of geomagnetic data integrated space. *SPIIRAS Proceedings*, 2019, vol. 18, no. 2, pp. 390–415. (in Russian). <https://doi.org/10.15622/sp.18.2.390-415>
16. Muhammad A., Minhaj K. Enhancing XML data parsing and querying performance on multi-core architectures. *Journal of Statistics, Computing and Interdisciplinary Research*, 2024, vol. 6, no. 1, pp. 75–89. <https://doi.org/10.52700/scir.v6i1.158>
17. Vorobev A.V., Vorobeva G.R. An approach to detecting and eliminating spatial contour artifacts in Web GIS applications. *Computer Optics*, 2023, vol. 47, no. 1, pp. 126–136. (in Russian). <https://doi.org/10.18287/2412-6179-co-1127>
18. Vorobev A.V., Vorobeva G.R. An approach to dynamic visualization of heterogeneous geospatial vector images. *Computer Optics*, 2024, vol. 48, no. 1, pp. 123–138. (in Russian). <https://doi.org/10.18287/2412-6179-CO-1279>
19. Shene J., Chnoor R., Mzhda A. XML schema validation using Java API for XML processing. *UKH Journal of Science and Engineering*, 2022, vol. 6, no. 1, pp. 33–41. <https://doi.org/10.25079/ukhjse.v6n1y2022>
20. Chen R., Wang Z., Su H., Xie S., Wang Z. Parallel XPath query based on cost optimization. *The Journal of Supercomputing*, 2022, vol. 78, no. 4, pp. 5420–5449. <https://doi.org/10.1007/s11227-021-04074-y>
21. Areces C., Fervari R. Axiomatizing Hybrid XPath with data. *Logical Methods in Computer Science*, 2021, vol. 17, no. 3, pp. 5. [https://doi.org/10.46298/lmcs-17\(3:5\)2021](https://doi.org/10.46298/lmcs-17(3:5)2021)
22. Chernenkiy V.M., Gapanyuk Y.E., Kaganov Y.T., Dunin I.V., Lyaskovsky M.A., Larionov V.S. Storing metagraph model in relational, document-oriented, and graph databases. *CEUR Workshop Proceedings*, 2018, vol. 2277, pp. 82–89.
23. Terekhov V., Gapanyuk Y., Kanev A. Metagraph representation for overcoming limitations of existing knowledge bases. *Proc. of the 28th Conference of Open Innovations Association (FRUCT)*, 2021, pp. 458–464. <https://doi.org/10.23919/fruct50888.2021.9347601>
24. Zichar M. Geovisualization based upon KML. *Journal of Agricultural Informatics*, 2012, vol. 3, no. 1, pp. 19–26. <https://doi.org/10.17700/jai.2012.3.1.49>
25. Ortiz-Alvarez A., Ruiz-Suárez L.G., Luis G., Ortega E., García-Reynoso A., Peralta O., López-Gaona A., Castro, T., Martínez-Arroyo A. Emission inventory point source visualization on Google Earth and integrated with HYSPLIT model. *Atmósfera*, 2021, vol. 34, no. 2, P. 143–156. <https://doi.org/10.20937/ATM.52834>

Авторы

Воробьева Гульнара Равилевна — доктор технических наук, доцент, профессор, Уфимский университет науки и технологий, Уфа, 450008, Российская Федерация, [sc 57189275619](https://orcid.org/0000-0001-7878-9724), <https://orcid.org/0000-0001-7878-9724>, Gulnara.vorobeve@gmail.com

Воробьев Андрей Владимирович — доктор технических наук, доцент, заведующий кафедрой, профессор, Уфимский университет науки и технологий, Уфа, 450008, Российская Федерация; старший научный сотрудник, Геофизический центр РАН, Москва, 119296, Российская Федерация, [sc 56767909700](https://orcid.org/0000-0002-9680-5609), <https://orcid.org/0000-0002-9680-5609>, geomagnet@list.ru

Орлов Глеб Олегович — аспирант, Уфимский университет науки и технологий, Уфа, 450008, Российская Федерация, [sc 58970339400](https://orcid.org/0009-0003-5123-3859), <https://orcid.org/0009-0003-5123-3859>, orlovgleb99@mail.ru

Authors

Gulnara R. Vorobeve — D.Sc., Associate Professor, Professor, Ufa University of Science and Technology, Ufa, 450008, Russian Federation, [sc 57189275619](https://orcid.org/0000-0001-7878-9724), <https://orcid.org/0000-0001-7878-9724>, gulnara.vorobeve@gmail.com

Andrei V. Vorobev — D.Sc., Associate Professor, Head of Department, Professor, Ufa University of Science and Technology, Ufa, 450008, Russian Federation; Senior Researcher, Geophysical Center of the RAS, Moscow, 119296, Russian Federation, [sc 56767909700](https://orcid.org/0000-0002-9680-5609), <https://orcid.org/0000-0002-9680-5609>, geomagnet@list.ru

Gleb O. Orlov — PhD Student, Ufa University of Science and Technology, Ufa, 450008, Russian Federation, [sc 58970339400](https://orcid.org/0009-0003-5123-3859), <https://orcid.org/0009-0003-5123-3859>, orlovgleb99@mail.ru

*Статья поступила в редакцию 22.04.2025
Одобрена после рецензирования 15.05.2025
Принята к печати 25.05.2025*

*Received 22.04.2025
Approved after reviewing 15.05.2025
Accepted 25.05.2025*



Работа доступна по лицензии
Creative Commons
«Attribution-NonCommercial»

doi: 10.17586/2226-1494-2025-25-3-574-583

УДК 004.7

Граничная оценка надежности кластерных систем на основе декомпозиции марковской модели при ограниченном восстановлении узлов с накоплением отказов

Владимир Анатольевич Богатырев¹✉, Станислав Владимирович Богатырев²,
Анатолий Владимирович Богатырев³

¹ Санкт-Петербургский государственный университет аэрокосмического приборостроения, Санкт-Петербург, 190000, Российская Федерация

^{1,2} Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация

^{2,3} ООО «Центр разработки облачных хранилищ Ядро», Санкт-Петербург, 195027, Российская Федерация

¹ vladimir.bogatyrev@gmail.com✉, <https://orcid.org/0000-0003-0213-0223>

² realloc@gmail.com, <https://orcid.org/0000-0003-0836-8515>

³ gangleon@gmail.com, <https://orcid.org/0000-0001-5447-7275>

Аннотация

Введение. Исследованы возможности граничной оценки надежности кластера, состоящего из множества узлов, каждый из которых может находиться в значительном числе состояний, различающихся производительностью выполнения требуемых функций и средним временем восстановления до исправного состояния узла. Оценка надежности такой кластерной системы на основе марковских процессов затруднена на этапе построения диаграммы состояний и переходов из-за большой ее размерности. Трудность построения модели особенно возрастает при ограниченном восстановлении узлов, приводящем к очереди узлов, требующих восстановления. Преодолеть указанную трудность позволяет предлагаемый подход. Представленный подход предусматривает декомпозицию марковской модели кластера и поэтапное последовательное уточнение верхней и нижней граничных оценок надежности кластера с учетом влияния на замедление восстановления каждого узла кластера других его узлов. **Метод.** Особенность предлагаемого подхода заключается в декомпозиции модели с выделением некоторого отдельного узла кластера и построении его марковской модели с введением состояний ожидания начала восстановлений узла из-за обслуживания очереди на восстановление других ранее отказавших узлов кластера. Определив на марковской модели выделенного узла вероятности всех его состояний, учитывая идентичность всех узлов кластера, на основе формулы перебора гипотез определяются средние задержки до восстановления исправного состояния остальных узлов кластера, имеющих ранее возникшие отказы. Вычисленные средние задержки используются на следующем этапе расчета марковской модели узла с уточнением задержки начала восстановления выделенного узла из-за влияния очереди восстановления остальных узлов кластера. **Основные результаты.** На основе предлагаемой модели дана оценка коэффициента готовности кластера, комплектуемого из значительного числа структурно сложных узлов, характеризующихся множеством состояний разной производительности и времени восстановления узла до исходного исправного состояния. **Обсуждение.** Представленная в результате декомпозиции модель позволяет преодолеть проблему лавинообразного возрастания сложности модели кластера при увеличении числа его узлов и числа их состояний. Выполненные расчеты показали сходимость предлагаемой граничной оценки надежности кластера из структурно сложных узлов. Полученные результаты могут быть использованы при оценке надежности и обосновании выбора структуры кластеров, а также дисциплин их обслуживания и восстановления при накоплении отказов с учетом ограниченных ресурсов восстановления, приводящих к образованию очередей отказавших элементов, подлежащих восстановлению. Предложенная модель может быть применена при анализе влияния накопления отказов в разных узлах кластера на задержки обслуживания поступающего потока запросов.

Ключевые слова

граничная оценка, надежность, кластер, узлы с множеством состояний, ограниченное восстановление, коэффициент готовности, марковская модель, декомпозиция, задержка восстановления

Ссылка для цитирования: Богатырев В.А., Богатырев С.В., Богатырев А.В. Граничная оценка надежности кластерных систем на основе декомпозиции марковской модели при ограниченном восстановлении узлов с накоплением отказов // Научно-технический вестник информационных технологий, механики и оптики. 2025. Т. 25, № 3. С. 574–583. doi: 10.17586/2226-1494-2025-25-3-574-583

Boundary estimation of the reliability of cluster systems based on the decomposition of the Markov model with limited recovery of nodes with accumulated failures

Vladimir A. Bogatyrev¹✉, Stanislav V. Bogatyrev², Anatoly V. Bogatyrev³

¹ Saint Petersburg State University of Aerospace Instrumentation (SUAI), Saint Petersburg, 190000, Russian Federation

^{1,2} ITMO University, Saint Petersburg, 197101, Russian Federation

^{2,3} Yadro Cloud Storage Development Center, Saint Petersburg, 195027, Russian Federation

¹ vladimir.bogatyrev@gmail.com✉, <https://orcid.org/0000-0003-0213-0223>

² realloc@gmail.com, <https://orcid.org/0000-0003-0836-8515>

³ gangleon@gmail.com, <https://orcid.org/0000-0001-5447-7275>

Abstract

The possibilities of a boundary assessment of the reliability of a cluster consisting of many nodes, each of which can be in a significant number of states, differing in the performance of the required functions and the average recovery time to a healthy node, are being investigated. Estimating the reliability of such a cluster system based on Markov processes is difficult at the stage of constructing a diagram of states and transitions due to its large dimension. The difficulty of building a model increases especially with limited node recovery, leading to a queue of nodes requiring recovery. The proposed approach allows us to overcome this difficulty. The differences between the proposed approaches are that it provides for the decomposition of the Markov cluster model and a step-by-step sequential refinement of the upper and lower boundary estimates of cluster reliability, taking into account the impact on slowing down the recovery of each cluster node of its other nodes. The peculiarity of the proposed approach is the decomposition of the model with the allocation of a certain individual cluster node and the construction of its Markov model with the introduction of waiting states for node recoveries due to queue maintenance for the restoration of other previously failed cluster nodes. Having determined the probabilities of all its states on the Markov model of the selected node, taking into account the identity of all cluster nodes, the average delays until the restoration of the serviceable state of the remaining cluster nodes with previous failures are determined based on the hypothesis enumeration formula. The calculated average delays are used in the next stage of calculating the Markov node model, specifying the delay in starting recovery of the allocated node due to the influence of the recovery queue of the remaining nodes in the cluster. Based on the proposed model, the availability coefficient of a cluster is estimated, consisting of a significant number of structurally complex nodes characterized by a variety of states of different performance and recovery time of the node to its initial working condition. As a result of decomposition, the proposed model makes it possible to overcome the problem of an avalanche-like increase in the complexity of the cluster model with an increase in the number of its nodes and the number of their states. The calculations performed have shown the convergence of the proposed boundary estimate of the reliability of a cluster of a significant number of structurally complex nodes. The results obtained can be used to assess the reliability and justify the choice of cluster structure as well as the disciplines of their maintenance and recovery when failures accumulate, taking into account limited recovery resources leading to the formation of queues of failed elements to be restored. The proposed model can be used to analyze the impact of the accumulation of failures in different cluster nodes on the delays in servicing the incoming request stream.

Keywords

boundary estimation, reliability, cluster, nodes with multiple states, limited recovery, availability factor, Markov model, decomposition, recovery delay

For citation: Bogatyrev V.A., Bogatyrev S.V., Bogatyrev A.V. Boundary estimation of the reliability of cluster systems based on the decomposition of the Markov model with limited recovery of nodes with accumulated failures. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2025, vol. 25, no. 3, pp. 574–583 (in Russian). doi: 10.17586/2226-1494-2025-25-3-574-583

Введение

К современным распределенным системам обработки и хранения данных предъявляются высокие требования по надежности [1–3], отказоустойчивости [4–6] и малым задержкам выполнения функциональных задач [7, 8]. В ряде случаев для систем реального времени дополнительно требуются обеспечение непрерывности и своевременности вычислительного процесса, когда запросы должны выполняться за ограниченное предельно допустимое время [7]. Меры по поддержке

надежности и снижению задержек реализуются на всех уровнях и стадиях функционирования инфокоммуникационной системы, в том числе при межмашинном обмене и распределении запросов (нагрузки) через сеть и их выполнении в центрах обработки и хранения данных [8–10]. Высокая надежность при малых задержках обслуживания запросов может достигаться при консолидации ресурсов хранения и обработки данных в кластеры с многопутевым доступом к ним через структурно избыточную сеть [11–13]. Особенно остро задача обеспечения высокой надежности отказоустойчивости

при ультрамалых задержках передачи данных стоит для распределенных облачных систем [14–17] и систем интернет вещей реального времени [18–20].

Современные резервированные компьютерные системы, в том числе узлы кластеров хранения и обработки данных, как правило представляют собой структурно сложные устройства, имеющие множество возможных состояний различной эффективности и запаса надежности функционирования [21–23]. Такая ситуация усложняет построение моделей оценки (прогнозирование) надежности кластеров и оптимизацию их структурно параметрической организации, а также обоснование выбора дисциплин обслуживания и восстановления после отказов. Трудность оценки надежности кластеров усугубляется характерной для них ограниченностью восстановления, при которой в каждый момент времени восстанавливается только один элемент и возможны очереди на восстановление отказавших элементов разных узлов [1–4]. Преодоление указанной сложности оценки надежности кластеров возможно при ее граничной верхней и нижней оценке. При этом нижнее приближение используется для оценки надежности, а верхнее для вычисления ее погрешности.

Приближенная верхняя оценка надежности восстанавливаемых кластерных систем может основываться на предположении неограниченного восстановления, при котором ремонт элемента начинается сразу после его отказа. Однако такое упрощение модели связано с необходимостью оценки ее погрешности. В предположении неограниченного восстановления расчет надежности можно свести к параллельно-последовательной структуре соединения элементов, но это происходит не всегда [1–4]. Для систем, модель которых не сводится к параллельно-последовательной структуре, расчет надежности может базироваться на граничных оценках Эзари–Прошана, Литвака–Ушакова или других известных методов расчета структурно сложных систем [1–4]. Однако предположение неограниченного восстановления может не гарантировать, что оценка по минимальным сечениям Эзари–Прошану приведет к нижней оценке надежности. То же относится к оценке по минимальным непересекающимся путям Литвака–Ушакова. Риск вместо искомого нижнего приближения надежности получить его верхнее приближение существует, так как оценка вероятности работоспособности (коэффициент готовности) элемента при неограниченном его восстановлении является завышенной. Действительно, в предположении неограниченного обслуживания задержка начала восстановления элемента из-за очереди на ремонт других ранее отказавших элементов не учитывается, что может привести к недопустимо завышенной оценке надежности структуры даже при использовании известных методов, ориентированных на ее нижнюю граничную оценку.

При рассмотрении узлов компьютерной системы в двух состояниях — работоспособном и отказавшем [23] — модель надежности ограниченного восстановления может быть сведена к известной из теории массового обслуживания модели размножения и гибели [24]. Недостаток такого подхода заключается в достаточной погрешности расчета при структурно сложных узлах

кластера, которые реально могут находиться в нескольких работоспособных состояниях с разным качеством обслуживания поступающих в кластер запросов и разным временем восстановления [23, 24].

Рассчитать надежность несложных систем при незначительном числе состояний даже при ограниченном восстановлении возможно на основе марковских процессов. Так известны марковские модели компьютерных систем с дублированием вычислителей и устройств памяти [25, 26], и систем с организацией памяти на базе Redundant Array of Independent Disks массивов [27–30]. Построения полумарковских моделей двухузловых систем, позволяющих учесть не только простейшие потоки отказов и экспоненциальное восстановление, предложены в работе [31]. Однако в случае кластера из достаточно большого числа структурно сложных узлов размерность диаграммы состояний и переходов (в том числе марковской модели) резко возрастает [32], что затрудняет оценку надежности кластеров.

Целью настоящей работы является граничная оценка надежности кластера с ограниченным восстановлением из значительного числа узлов, каждый из которых имеет множество состояний, характеризующихся различным запасом надежности, временем восстановления до исправного состояния и производительностью по обслуживанию трафика.

Метод граничной оценки надежности кластеров

Предлагаемый метод предусматривает декомпозицию марковской модели и поэтапное последовательное уточнение верхней и нижней граничных оценок надежности кластера с учетом влияния на замедление восстановления каждого узла кластера очередей на восстановление ранее отказавших элементов остальных его узлов.

Кластеры из структурно сложных узлов. Примеры кластеров из структурно сложных узлов, которые могут находиться в состояниях с разным временем восстановления, представлены на рис. 1. Отметим, что возможна модификация структуры кластеров, при которой устройства памяти заменяются на серверы.

Верхняя грубая оценка коэффициента готовности кластера с предположением неограниченного восстановления элементов и узлов. Верхнюю оценку надежности кластера с узлами, характеризующимися множеством работоспособных состояний, можно грубо рассчитать в предположении неограниченного восстановления каждого элемента. При условии, что сразу после отказа кластера начинается его восстановление без учета, что могут возникать очереди среди других ожидающих восстановления элементов в том же или иных узлах кластера.

Оценку стационарного коэффициента готовности проведем для структуры по рис. 1, б. При неограниченном восстановлении элементов, в предположении что кластер сохраняет работоспособность, если работоспособен хотя бы один его узел, получаем наиболее грубую верхнюю оценку (h), определяемую как

$$K_{h0} = 1 - (1 - k)^n,$$

где k — коэффициент готовности узла кластера

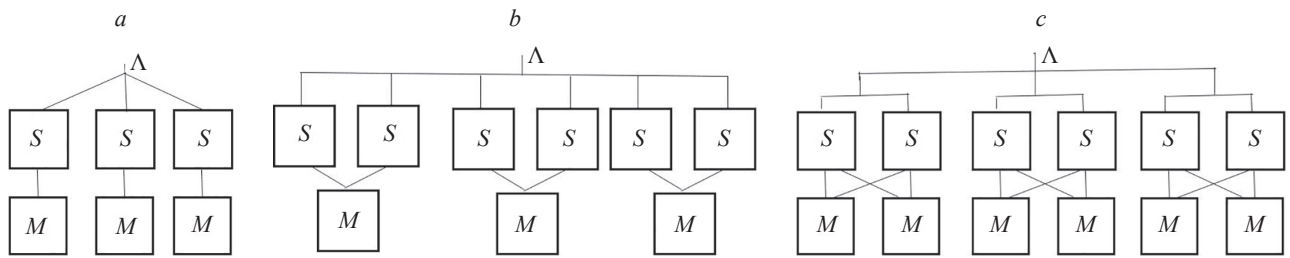


Рис. 1. Кластеры из структурно сложных узлов.

Узел кластера содержит: один сервер (S) и одно устройство памяти (M), при этом Λ — интенсивность входного потока (a); два сервера и одно устройство памяти (b); два сервера и два устройства памяти (c)

Fig. 1. Clusters of structurally complex nodes. The cluster node contains: one server and one memory device (a); two servers and one memory device (b); two servers and two memory devices (c)

$$k = k_2[1 - (1 - k_1)^2].$$

При этом для структур (рис. 1, b) k_1 и k_2 — коэффициенты готовности серверов S и устройств хранения M .

$$k_1 = \mu_1(\lambda_1 + \mu_1)^{-1}, \quad k_2 = \mu_2(\lambda_2 + \mu_2)^{-1},$$

где $\lambda_1, \lambda_2, \mu_1, \mu_2$ — интенсивности отказов и восстановлений серверов S и устройств хранения M .

Расчеты выполним при $\lambda_1 = 2 \cdot 10^{-4} \text{ ч}^{-1}$, $\lambda_2 = 2 \cdot 10^{-4} \text{ ч}^{-1}$, $\mu_1 = 5 \cdot 10^{-4} \text{ ч}^{-1}$, $\mu_2 = 4 \cdot 10^{-4} \text{ ч}^{-1}$. В этом случае коэффициент готовности кластера $K_{h0} = 0,9999999993795285$.

Проведем уточненную верхнюю оценку коэффициента готовности кластера в предположении ограниченного обслуживания каждого его узла, когда задержка восстановления учитывается только из-за отказов элементов внутри узла, а задержки из-за образования очереди восстановления отказов в других узлах кластера не учитываются. Такое предположение приводит к завышенной оценке надежности кластера.

Предлагаемый подход характеризуется поэтапным последовательным приближением верхней и нижней граничных оценок надежности кластера с уточнением на каждом этапе влияния на замедление восстановления каждого узла кластера очередей на восстановления ранее отказавших элементов других его узлов.

Оценка выполняется в два этапа.

Этап 1. Построение марковской модели узла без учета влияния на задержку его восстановления других узлов кластера, что приводит к завышенной оценке надежности узла кластера.

Этап 2. По вычисленным на марковской модели вероятностям возможных работоспособных состояний узла, по формуле перебора гипотез находим надежность кластера целиком. Полученная оценка надежности кластера будет завышенной, но погрешность будет меньше, чем для первого рассматриваемого приближения, полученного на этапе 1.

Для узла кластера по рис. 1, b без учета влияния на задержки восстановления других узлов кластера диаграмма состояний и переходов представлена на рис. 2. Каждое возможное состояние узла кластера представлено узлом (вершиной) диаграммы, в поле которого верхней строкой обозначены состояния работоспособности двух серверов, а нижней — устройства памяти,

при этом работоспособные состояния соответствующих элементов узла кластера обозначены как «1», а отказавшие — «0». Рядом с вершинами (узлами) диаграммы показаны вероятности соответствующих состояний P_0, P_1, P_2, P_3, P_4 . К работоспособным состояниям узла кластера относятся состояния при исправности хотя бы одного сервера и устройства памяти (вероятности этих состояний P_0 или P_1).

По диаграмме (рис. 2) составим систему алгебраических уравнений:

$$\begin{cases} -(2\lambda_1 + \lambda_2)P_0 + \mu_2P_3 + \mu_1P_1 = 0, \\ -\mu_1P_2 + \lambda_1P_1 = 0, \\ -\mu_2P_3 + \lambda_2P_0 = 0, \\ -\mu_2P_4 + \lambda_2P_1 = 0, \\ P_0 + P_1 + P_2 + P_3 + P_4 = 1. \end{cases}$$

По представленной системе уравнений можно найти вероятности всех состояний узла кластера P_0, P_1, P_2, P_3, P_4 .

Стационарный коэффициент готовности определим как вероятность того, что в произвольный момент времени работоспособен хотя бы один из n узлов кластера вне зависимости числа исправных в нем серверов S

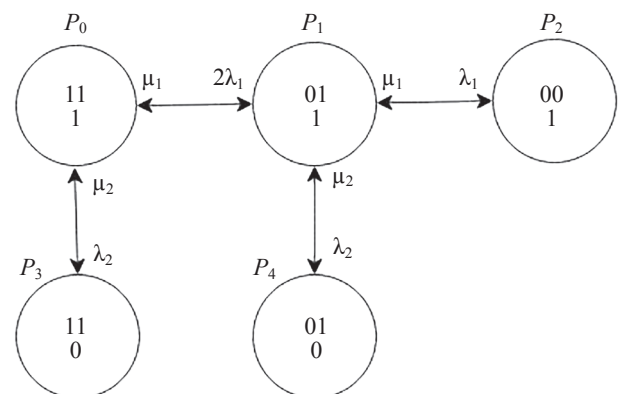


Рис. 2. Диаграмма состояний и переходов узла кластера при неограниченном его восстановлении без учета задержек восстановления из-за отказов других узлов кластера.

P_0, P_1, P_2, P_3, P_4 — вероятности состояний

Fig. 2. Diagram of the states and transitions of a cluster node with unlimited recovery without taking into account recovery delays due to failures of other cluster nodes

$$K_{h1} = \sum_{i=1}^n C_n^i (P_0 + P_1)^i (1 - P_0 - P_1)^{n-i}. \quad (1)$$

Если условие работоспособности кластера заключается в работоспособности не менее g его узлов, то:

$$K_{h1} = \sum_{i=g}^n C_n^i (P_0 + P_1)^i (1 - P_0 - P_1)^{n-i}.$$

Если следует разграничивать число работоспособных узлов с исправностью одного или двух серверов, то коэффициент готовности определяется как:

$$K_{h11} = \sum_{i=0}^n \sum_{j=0}^{n-i} \delta_{ij} C_n^i C_{n-i}^j P_0^i P_1^j (1 - P_0 - P_1)^{n-i-j}, \quad (2)$$

$$\delta_{ij} = \begin{cases} 1, & \text{if } i+j \geq 1, \\ 0, & \text{if } i+j < 1. \end{cases} \quad (3)$$

Разграничение работоспособных состояний по числу сохраненных серверов позволяет вычислить такие показатели системы как коэффициент сохранения эффективности или средние задержки обслуживания запросов с накопления отказов в узлах кластера. Так, коэффициент сохранения эффективности в зависимости от общего числа работоспособных серверов может быть вычислен по формуле

$$K_{сэ} = \sum_{i=0}^n \sum_{j=0}^{n-i} \delta_{ij} (2i+j) C_n^i C_{n-i}^j P_0^i P_1^j (1 - P_0 - P_1)^{n-i-j}.$$

Заметим, что результаты расчета коэффициента готовности по формулам (1) и (2) совпадают и равны

$$K_{h1} = 0,999999993717164.$$

Расчет выполнен для исходных данных, заданных при оценке K_{h0} .

Таким образом, точность оценки коэффициента готовности повысилась сравнительно с наиболее грубым верхним приближением K_{h0} .

Оценив вероятности возможных состояний каждого узла кластера P_0, P_1, P_2, P_3, P_4 , выделим некоторый узел кластера и среди остальных $n-1$ узлов кластера вычислим математическое ожидание числа узлов, находящихся в каждом (i -ом) из возможных состояний, с разным временем восстановления до исправного состояния узла. Искомое математическое ожидание найдем по формуле

$$n_i = P_i(n-1). \quad (4)$$

Таким образом, математическое ожидание задержки τ начала восстановления некоторого выделенного узла из-за очереди восстановления ранее отказавших элементов в остальных $n-1$ узлах кластера имеет вид

$$\tau = \frac{n_1}{\mu_1} + \frac{2n_2}{\mu_1} + \frac{n_3}{\mu_2} + n_4 \left(\frac{1}{\mu_1} + \frac{1}{\mu_2} \right), \quad (5)$$

и соответственно $\gamma = \tau^{-1}$, где γ — интенсивность перехода к началу восстановления узла.

После соответствующих предварительных расчетов перейдем к дальнейшему уточнению верхней оценки коэффициента готовности.

Уточнение верхней оценки коэффициента готовности кластера с учетом задержек из-за восстановления отказавших элементов всех его узлов.

Особенность предлагаемого уточнения заключается в декомпозиции, предполагающей выделение некоторого отдельного узла кластера и построении его марковской модели с введением состояний ожидания начала восстановления узла из-за обслуживания очереди на восстановление других ранее отказавших узлов кластера.

На данном этапе построения марковской модели учтем задержку начала обслуживания в некотором выделенном узле из-за очереди на восстановление элементов, отказавших в других узлах кластера.

Построение модели реализуется в два этапа.

Этап 1. Формируем марковскую модель некоторого узла с учетом возможных задержек начала его восстановления из-за обслуживания очереди ранее отказавших элементов других узлов кластера.

Этап 2. С учетом вычисленных на этапе 1 вероятностей различных состояний узлов кластера по формуле перебора гипотез находим коэффициент готовности всего кластера.

Диаграмма состояний и переходов марковской модели некоторого узла кластера с учетом задержек начала восстановления $\tau = 1/\gamma$ приведена на рис. 3. Если ресурсы узла позволяют вычислять требуемые функции, то восстановление может быть совмещено с обслуживанием поступающих функциональных запросов, что возможно связано с дополнительными отказами и переходами из состояния с вероятностью P_8 , в состояния, вероятности, которых P_6 или P_7 (рис. 3, а). Если восстановление отказавших элементов узла кластера проводится при отключенном электропитании, что исключает отказы работоспособных элементов узла, то диаграмма состояний и переходов узла преобразуется к виду на рис. 3, б.

Математическое ожидание задержки τ начала восстановления некоторого выделенного узла из-за очереди ранее отказавших элементов в остальных $n-1$ узлах кластера вычислим по формуле (5). В результате интенсивность перехода к началу восстановления выделенного узла будет равна $\gamma = \tau^{-1}$.

Оценка τ проведена с учетом среднего числа узлов n_1, n_2, n_3, n_4 , определяемых по формуле (4) с различным набором отказавших элементов (по рис. 2), и времени, требуемого для восстановления исправности этих узлов.

Для определения вероятностей состояний по диаграмме (рис. 3) сформируем систему алгебраических уравнений:

$$\begin{cases} -(2\lambda_1 + \lambda_2)P_0 + \mu_2 P_5 + \mu_1 P_8 = 0, \\ -(\lambda_1 + \lambda_2 + \gamma)P_1 + \lambda_2 P_0 = 0, \\ -\gamma P_2 + \lambda_2 P_1 = 0, \\ -\gamma P_3 + \lambda_2 P_0 = 0, \\ -\gamma P_4 + \lambda_1 P_1 = 0, \\ -\mu_2 P_5 + \gamma P_1 = 0, \\ -\mu_2 P_6 + \gamma P_3 = 0, \\ -\mu_1 P_7 + \gamma P_4 + \lambda_1 P_8 = 0, \\ P_0 + P_1 + P_2 + P_3 + P_4 + P_5 + P_6 + P_7 + P_8 = 1. \end{cases} \quad (6)$$

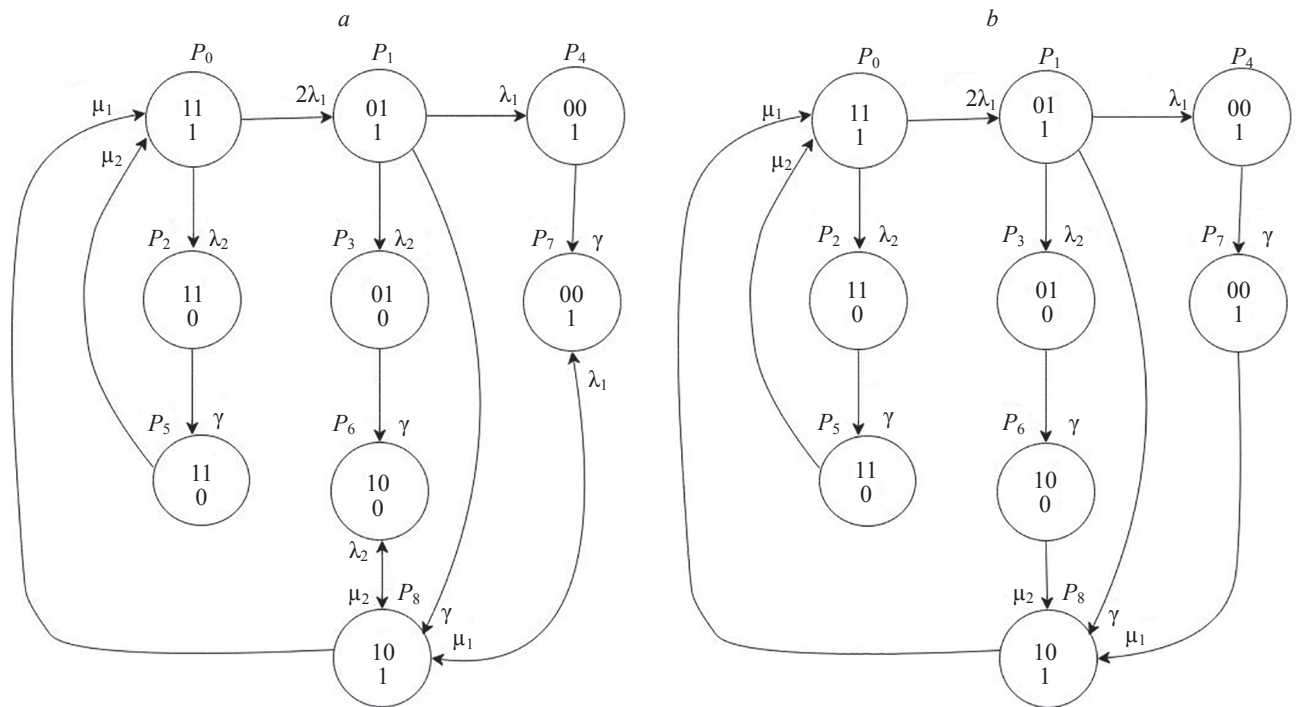


Рис. 3. Диаграммы состояний и переходов марковской модели узла кластера с учетом задержек начала восстановления без отключения электропитания во время восстановления (а) и с его отключением до возобновления исправного состояния узла (б).

P_0 и P_1 — вероятности состояний с выполнением требуемых функций при исправности двух и одного сервера; P_2, P_3 и P_4 — вероятности неработоспособных состояний с ожиданием начала восстановления; P_5-P_8 — вероятности состояний, в которых производится восстановление узла

Fig. 3. Diagram of the states and transitions of the Markov cluster node model, taking into account delays in the start of recovery without power outage during recovery (a) and with its shutdown until the resumption of the serviceable state of the node (b)

Стационарный коэффициент готовности кластера вычислим по формуле

$$K_{h2} = \sum_{i=1}^n C_n^i (P_0 + P_1 + P_8)^i (1 - P_0 - P_1 - P_8)^{n-i}. \quad (7)$$

Если следует разграничивать число работоспособных узлов с исправностью одного или двух серверов, то коэффициент готовности получим в виде:

$$K_{h21} = \sum_{i=0}^n \sum_{j=0}^{n-i} \delta_{ij} C_n^i C_{n-i}^j P_0^i (P_1 + P_8)^j (1 - P_0 - P_1 - P_8)^{n-i-j},$$

где δ_{ij} задается по формуле (3).

Уточненной верхней оценке коэффициента готовности кластера соответствует

$$K_{h2} = 0,9999999993065137.$$

Для прогнозирования надежности технических систем более важна ее нижняя оценка. Останемся на поэтапном итерационном приближении нижней оценки. Результаты верхней оценки могут быть использованы при расчете погрешности нижней оценки.

Нижняя оценка надежности кластера. При грубом нижнем приближении выделим некоторый узел и будем считать, что в каждом из остальных $n-1$ узлов кластера отказывают элементы, восстановление совокупности которых до исправного состояния узла требует наибольшее время. Диаграмма состояний и пере-

ходов марковской модели каждого из n узлов кластера соответствует рис. 3. Система уравнений для поиска вероятностей состояний выделенного узла имеет вид (6), причем интенсивность перехода к началу восстановления узла находим для наихудшего случая

$$\gamma = \mu_2 / 2(n-1).$$

Коэффициент готовности кластера K_{n1} (7) для нижней оценки (n) вычисляется по формуле перебора гипотез.

На основе уравнений (6), вычислив вероятности всех состояний каждого узла, уточненное время ожидания начала восстановления некоторого узла оценим как

$$\tau = (1 - P_0) \frac{n-1}{P_5 + P_6 + P_7 + P_8} \left[\frac{2P_7}{\mu_1} + \frac{P_8}{\mu_1} + \frac{P_5}{\mu_2} + P_6 \left(\frac{1}{\mu_1} + \frac{1}{\mu_2} \right) \right].$$

Вычислив уточненное значение интенсивности $\gamma = \tau^{-1}$ на основе системы уравнений (6), находим уточненные значения вероятностей состояний узла кластера, после чего по формуле (7) вычисляем коэффициент готовности кластера K_{n2} .

Для заданных исходных данных имеем

$$K_{n0} = 0,9999984687564836, \\ K_{n1} = 0,9999999981600136.$$

Таблица. Поэтапные уточнения верхней и нижней оценок надежности кластера
Table. Step-by-step refinement of the upper and lower estimates of cluster reliability

Верхняя оценка		Нижняя оценка	
этап уточнения	значение	этап уточнения	значение
K_{h0}	0,999999993795285	K_{n0}	0,9999984687564836
K_{h1}	0,999999993717164	K_{n1}	0,999999981600136
K_{h2}	0,999999993065137	K_{n2}	0,999999992045664

В результате уточнения нижней оценки в соответствии с предложенным методом был получен результат

$$K_{n2} = 0,999999992045664.$$

Основные результаты

Основной результат заключается в том, что предложена верхняя и нижняя оценки стационарного коэффициента готовности кластера, из структурно сложных узлов, каждый из которого в результате различных комбинаций отказов элементов может обладать различным запасом надежности и требуемого времени восстановления до исправного состояния. Предлагаемый метод позволяет на основе многоэтапной процедуры уточнения верхнего и нижнего приближений добиться требуемой погрешности оценки коэффициента готовности кластера с учетом влияния ограниченных возможностей его восстановления. Результаты поэтапного уточнения верхней и нижней оценок надежности кластера приведены в таблице.

Нижняя оценка коэффициента готовности кластера может использоваться в качестве показателя надежности кластера, а верхняя для оценки предельной погрешности этой оценки, которая уменьшается на каждой итерации предлагаемой граничной оценки надежности кластера.

Представленные результаты расчета получены для случая, когда условие работоспособности кластера заключаются в требовании работоспособности хотя бы одного его узла. При требовании работоспособности не менее g узлов кластера результаты оценки искомого коэффициента готовности будут ниже представленных в таблице, но поэтапное снижение погрешности расчетов сохранится.

Обсуждение результатов и направления развития

Научная новизна предложенного метода граничной оценки надежности кластера из структурно сложных узлов заключается в декомпозиции модели кластера на основе построения марковской модели узла кластера с учетом поэтапного уточнения влияния на задержки его восстановления других узлов кластера. Значимость представленного метода оценки коэффициента готовности кластера обусловлена тем, что он позволяет избежать лавинообразного роста сложности марковской модели при увеличении числа структурно сложных узлов кластера. Метод предусматривает поэтапное уточнение верхнего и нижнего приближений оценки коэффициента готовности кластера.

Предложенный метод может быть использован при оценке надежности и обосновании выбора организации и дисциплин обслуживания кластеров из структурно сложных узлов с учетом их ограниченного восстановления, сопряженного с образованием очередей на их восстановление.

Ограничение предложенной модели надежности кластера заключается в ее ориентации на кластеры из однородных узлов.

В качестве дальнейшего развития предложенной марковской модели надежности предусматривается ее модернизация для кластеров, содержащих группы из неоднородных структурно сложных узлов.

Заключение

Предложена граничная оценка надежности кластера, каждый из узлов которого имеет множество состояний, характеризующихся различным запасом надежности и требуемого времени восстановления до исправного состояния.

Предлагаемая марковская модель надежности кластера отражает ограниченность восстановления узлов из-за возможного образования очереди на ремонт других узлов кластера.

Представленный метод характеризуется поэтапным последовательным приближением верхней и нижней граничных оценок надежности кластера с уточнением на каждом этапе влияния на замедление восстановления каждого узла кластера очередей на восстановления ранее отказавших элементов остальных его узлов.

Особенность предлагаемого подхода заключается в декомпозиции модели с выделением некоторого отдельного узла кластера и построении его марковской модели с введением состояний ожидания начала восстановлений узла из-за обслуживания очереди на восстановление других ранее отказавших узлов кластера. Определив на марковской модели выделенного узла вероятности всех его состояний, учитывая идентичность всех узлов кластера, на основе формулы перебора гипотез определяются средние задержки до восстановления исправного состояния остальных узлов кластера, имеющих ранее возникшие отказы. Вычисленные средние задержки используются на следующем этапе расчета марковской модели узла с уточнением задержки начала восстановления выделенного узла из-за влияния восстановления остальных узлов кластера.

На основе предлагаемой модели дана оценка коэффициента готовности кластера с учетом накопления отказов узлов, приводящих к снижению их вычислительных возможностей.

Предлагаемая модель позволяет преодолеть проблеме лавинообразного возрастания сложности модели кластера при увеличении числа его узлов и их состояний.

Выполненные по предлагаемой модели расчеты показали поэтапное уточнение верхней и нижней оценок коэффициента готовности кластера из структурно сложных узлов.

Литература

1. Половко А.М., Гуров С.В. Основы теории надежности // СПб.: БХВ-Петербург, 2006. 702 с.
2. Shubinsky I.B., Rozenberg I.N., Papic L. Adaptive fault tolerance in real-time information systems // *Reliability Theory and Applications*. 2017. V. 12. N 1. P. 18–25.
3. Черкесов Г.Н. Надежность аппаратно-программных комплексов. СПб.: Питер, 2005. 479 с.
4. Aysan H. Fault-tolerance strategies and probabilistic guarantees for real-time systems. Doctoral dissertation. Mälardalen University. 2012. 109 p.
5. Koren I., Krishna C.M. *Fault-Tolerant Systems*. Morgan Kaufmann, 2007. 378 p.
6. Krasnobaev V., Kuznetsov A., Kiian A., Kuznetsova K. Fault tolerance computer system structures functioning in residue classes // *Proc. of the 11th IEEE International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS)*. 2021. P. 471–474. <https://doi.org/10.1109/idaacs53288.2021.9660919>
7. Кучерявый А.Е. Сети связи с ультра малыми задержками // *Труды Научно-исследовательского института радио*. 2020. № 1. С. 69.
8. Tatarnikova T.M., Sikarev I.A., Bogdanov P.Y., Timochkina T.V. Botnet attack detection approach in IoT networks // *Automatic Control and Computer Sciences*. 2022. V. 56. N 8. P. 838–846. <https://doi.org/10.3103/S0146411622080259>
9. Bogatyrev V.A., Bogatyrev A.V., Bogatyrev S.V. The probability of timeliness of a fully connected exchange in a redundant real-time communication system // *Proc. of the Wave Electronics and its Application in Information and Telecommunication Systems (WECONF)*. 2020. P. 9131517. <https://doi.org/10.1109/WECONF48837.2020.9131517>
10. Burkov A., Rachugin R., Turlikov A. Stabilizing ALOHA using Preamble-based exploration by estimation of the number of active users // *Proc. of the 18th International Symposium Problems of Redundancy in Information and Control Systems, (REDUNDANCY)*. 2023. P. 106–109. <https://doi.org/10.1109/redundancy59964.2023.10330186>
11. Bogatyrev V.A., Bogatyrev A.V., Bogatyrev S.V. Multipath transmission of heterogeneous traffic in acceptable delays with packet replication and destruction of expired replicas in the nodes that make up the path // *Communications in Computer and Information Science*. 2023. V. 1748. P. 104–121. https://doi.org/10.1007/978-3-031-30648-8_9
12. Bogatyrev V.A. Protocols for dynamic distribution of requests through a bus with variable logic ring for reception authority transfer // *Automatic Control and Computer Sciences*. 1999. V. 33. N 1. P. 57–63.
13. Bogatyrev V.A., Bogatyrev S.V., Bogatyrev A.V. Control of multipath transmissions in the nodes of switching segments of reserved paths // *Proc. of the International Conference on Information, Control, and Communication Technologies (ICCT)*. 2022. P. 1–5. <https://doi.org/10.1109/icct56057.2022.9976839>
14. Tatarnikova T.M., Arkhipov E.M. Designing fault-tolerant systems with micro-service architecture // *Proc. of 27th International Conference on Soft Computing and Measurements (SCM)*. 2024. P. 348–351. <https://doi.org/10.1109/scm62608.2024.10554143>
15. Haider S., Nazir B. Fault tolerance in computational grids: perspectives, challenges, and issues // *SpringerPlus*. 2016. V. 5. N 1. P. 1991. <https://doi.org/10.1186/s40064-016-3669-0>
16. Chinnaiiah M.R., Niranjana N. Fault tolerant software systems using software configurations for cloud computing // *Journal of Cloud Computing*. 2018. V. 7. N 1. P. 3. <https://doi.org/10.1186/s13677-018-0104-9>
17. Markoval E., Moltchanov D., Pirmagomedov R., Ivanova D., Koucheryav Y., Samouylov K. Priority-based coexistence of eMBB and URLLC traffic in industrial 5G NR deployments // *Proc. of the 12th International Congress on Ultra Modern Telecommunications and*

References

1. Polovko A.M., Gurov S.V. *Fundamentals of Reliability Theory*. St. Petersburg, BHV-Petersburg Publ., 2006, 702 p. (in Russian)
2. Shubinsky I.B., Rozenberg I.N., Papic L. Adaptive fault tolerance in real-time information systems. *Reliability Theory and Applications*, 2017, vol. 12, no. 1, pp. 18–25.
3. Cherkessov G.N. *Reliability of Hardware and Software Complexes*. St. Petersburg, Piter Publ., 2005, 479 p. (in Russian)
4. Aysan H. *Fault-tolerance strategies and probabilistic guarantees for real-time systems*. Doctoral dissertation. Mälardalen University. 2012. 109 p.
5. Koren I., Krishna C.M. *Fault-Tolerant Systems*. Morgan Kaufmann, 2007, 378 p.
6. Krasnobaev V., Kuznetsov A., Kiian A., Kuznetsova K. Fault tolerance computer system structures functioning in residue classes. *Proc. of the 11th IEEE International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS)*, 2021, pp. 471–474. <https://doi.org/10.1109/idaacs53288.2021.9660919>
7. Kucheryav Y.A.E. *Ultra Low Latency communication networks*. Trudy NIIR, 2020, no. 1, pp. 69 (in Russian)
8. Tatarnikova T.M., Sikarev I.A., Bogdanov P.Y., Timochkina T.V. Botnet attack detection approach in IoT networks. *Automatic Control and Computer Sciences*, 2022, vol. 56, no. 8, pp. 838–846. <https://doi.org/10.3103/S0146411622080259>
9. Bogatyrev V.A., Bogatyrev A.V., Bogatyrev S.V. The probability of timeliness of a fully connected exchange in a redundant real-time communication system. *Proc. of the Wave Electronics and its Application in Information and Telecommunication Systems (WECONF)*, 2020, pp. 9131517. <https://doi.org/10.1109/WECONF48837.2020.9131517>
10. Burkov A., Rachugin R., Turlikov A. Stabilizing ALOHA using Preamble-based exploration by estimation of the number of active users. *Proc. of the 18th International Symposium Problems of Redundancy in Information and Control Systems, (REDUNDANCY)*, 2023, pp. 106–109. <https://doi.org/10.1109/redundancy59964.2023.10330186>
11. Bogatyrev V.A., Bogatyrev A.V., Bogatyrev S.V. Multipath transmission of heterogeneous traffic in acceptable delays with packet replication and destruction of expired replicas in the nodes that make up the path. *Communications in Computer and Information Science*, 2023, vol. 1748, pp. 104–121. https://doi.org/10.1007/978-3-031-30648-8_9
12. Bogatyrev V.A. Protocols for dynamic distribution of requests through a bus with variable logic ring for reception authority transfer. *Automatic Control and Computer Sciences*, 1999, vol. 33, no. 1, pp. 57–63.
13. Bogatyrev V.A., Bogatyrev S.V., Bogatyrev A.V. Control of multipath transmissions in the nodes of switching segments of reserved paths. *Proc. of the International Conference on Information, Control, and Communication Technologies (ICCT)*, 2022, pp. 1–5. <https://doi.org/10.1109/icct56057.2022.9976839>
14. Tatarnikova T.M., Arkhipov E.M. Designing fault-tolerant systems with micro-service architecture. *Proc. of the 27th International Conference on Soft Computing and Measurements (SCM)*, 2024, pp. 348–351. <https://doi.org/10.1109/scm62608.2024.10554143>
15. Haider S., Nazir B. Fault tolerance in computational grids: perspectives, challenges, and issues. *SpringerPlus*, 2016, vol. 5, no. 1, pp. 1991. <https://doi.org/10.1186/s40064-016-3669-0>
16. Chinnaiiah M.R., Niranjana N. Fault tolerant software systems using software configurations for cloud computing. *Journal of Cloud Computing*, 2018, vol. 7, no. 1, pp. 3. <https://doi.org/10.1186/s13677-018-0104-9>
17. Markoval E., Moltchanov D., Pirmagomedov R., Ivanova D., Koucheryav Y., Samouylov K. Priority-based coexistence of eMBB and URLLC traffic in industrial 5G NR deployments. *Proc. of the 12th International Congress on Ultra Modern Telecommunications and*

- Control Systems and Workshops (ICUMT). 2020. P. 1–6. <https://doi.org/10.1109/ICUMT51630.2020.9222433>
18. Ji H., Park S., Yeo J., Kim Y., Lee J., Shim B. Ultra-Reliable and Low-Latency Communications in 5G Downlink: physical layer aspects // *IEEE Wireless Communications*. 2018. V. 25. N 3. P. 124–130. <https://doi.org/10.1109/mwc.2018.1700294>
 19. Gurjanov A.V., Korobeynikov A.G., Zharinov I.O., Zharinov O.O. Edge, fog and cloud computing in the cyber-physical systems networks // *Ceur Workshop Proceedings*. 2021. P. 103–108.
 20. Srivastava A., Kumar N. Queueing model based dynamic scalability for containerized cloud // *International Journal of Advanced Computer Science and Applications*. 2023. V. 14. N 1. P. 465–472. <https://doi.org/10.14569/IJACSA.2023.0140150>
 21. Астахова Т.Н., Верзун Н.А., Касаткин В.В., Колбанев М.О., Шамин А.А. Исследование моделей связности сенсорных сетей // *Информационно-управляющие системы*. 2019. № 5 (102). С. 38–50. <https://doi.org/10.31799/1684-8853-2019-5-38-50>
 22. Gurov S.V., Utkin L.V. Reliability of repairable reserved systems with failure aftereffect // *Automation and Remote Control*. 2017. V. 78. N 1. P. 113–124. <https://doi.org/10.1134/S000511791701009X>
 23. Bogatyrev V., Vinokurova M. Control and safety of operation of duplicated computer systems // *Communications in Computer and Information Science*. 2017. V. 700. P. 331–342. https://doi.org/10.1007/978-3-319-66836-9_28
 24. Леонтьев А.С., Тимошкин М.С. Математические модели оценки показателей надежности для исследования вероятностно-временных характеристик многомашинных комплексов с учетом отказов // *Международный научно-исследовательский журнал*. 2023. № 1(127). С. 18. <https://doi.org/10.23670/IRJ.2023.127.27>
 25. Клейнрок Л. Теория массового обслуживания. М.: Машиностроение, 1979. 432 с.
 26. Богатырев В.А., Богатырев С.В., Богатырев А.В. Оценка готовности компьютерной системы к своевременному обслуживанию запросов при его совмещении с информационным восстановлением памяти после отказов // *Научно-технический вестник информационных технологий, механики и оптики*. 2023. Т. 23. № 3. С. 608–617. <https://doi.org/10.17586/2226-1494-2023-23-3-608-617>
 27. Qi Y., Meng H., Hou D., Chen Y. A study on software rejuvenation model of application server cluster in two-dimension state space using Markov process // *Information Technology Journal*. 2008. V. 7. N 1. P. 98–104. <https://doi.org/10.3923/itj.2008.98.104>
 28. Rahman P.A. Advanced reliability model of the fault-tolerant disk arrays with data striping and single disk redundancy // *Proc. of the International Scientific and Practical Conference*. 2017. P. 20–25.
 29. Uspenskaya N.N. Estimation of availability factor for the data storage systems based on redundant disk arrays with the backup // *Proc. of the International Scientific and Practical Conference*. 2016. P. 20–23.
 30. Рахман П.А., Шарипов М.И. Модель надежности двухузлов кластера приложений высокой готовности в системах управления предприятием // *Экономика и менеджмент систем управления*. 2015. № 3 (17). С. 85–102.
 31. Хомоненко А.Д., Благовещенская Е.А., Проурзин О.В., Андрук А.А. Прогноз надежности кластерной вычислительной системы с помощью полумарковской модели альтернирующих процессов и мониторинга // *Наукоемкие технологии в космических исследованиях Земли*. 2018. Т. 10. № 4. С. 72–82. <https://doi.org/10.24411/2409-5419-2018-10099>
 32. Terskov V., Sakash I. The reliability evaluation of local computer networks using markov model of multiple heterogeneous groups of switches // *E3s Web of Conferences*. 2024. V. 592. P. 03036.
 - Control Systems and Workshops (ICUMT), 2020, pp. 1–6. <https://doi.org/10.1109/ICUMT51630.2020.9222433>
 18. Ji H., Park S., Yeo J., Kim Y., Lee J., Shim B. Ultra-Reliable and Low-Latency Communications in 5G Downlink: physical layer aspects. *IEEE Wireless Communications*, 2018, vol. 25, no. 3, pp. 124–130. <https://doi.org/10.1109/mwc.2018.1700294>
 19. Gurjanov A.V., Korobeynikov A.G., Zharinov I.O., Zharinov O.O. Edge, fog and cloud computing in the cyber-physical systems networks. *Ceur Workshop Proceedings*, 2021, pp. 103–108.
 20. Srivastava A., Kumar N. Queueing model based dynamic scalability for containerized cloud. *International Journal of Advanced Computer Science and Applications*, 2023, vol. 14, no. 1, pp. 465–472. <https://doi.org/10.14569/IJACSA.2023.0140150>
 21. Astakhova T.N., Verzun N.A., Kasatkin V.V., Kolbanev M.O., Shamin A.A. Sensor network connectivity models. *Informatsionno Upravliaiushchie Sistemy*, 2019. N 5 (102). P. 38–50. (in Russian). <https://doi.org/10.31799/1684-8853-2019-5-38-50>
 22. Gurov S.V., Utkin L.V. Reliability of repairable reserved systems with failure aftereffect. *Automation and Remote Control*, 2017, vol. 78, no. 1, pp. 113–124. <https://doi.org/10.1134/S000511791701009X>
 23. Bogatyrev V., Vinokurova M. Control and safety of operation of duplicated computer systems. *Communications in Computer and Information Science*, 2017, vol. 700, pp. 331–342. https://doi.org/10.1007/978-3-319-66836-9_28
 24. Leontev A.S., Timoshkin M.S. Mathematical models for evaluating reliability indicators to study the probabilistic and temporal characteristics of multi-machine complexes with regard to failures. *International Research Journal*, 2023, no. 1 (127). pp. 18. (in Russian). <https://doi.org/10.23670/IRJ.2023.127.27>
 25. Kleinrock L. *Queueing Systems: Theory*. Wiley, 1974, 417 p.
 26. Bogatyrev V.A., Bogatyrev S.V., Bogatyrev A.V. Assessment of the readiness of a computer system for timely servicing of requests when combined with information recovery of memory after failures. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2023, vol. 23, no. 3, pp. 608–617. (in Russian). <https://doi.org/10.17586/2226-1494-2023-23-3-608-617>
 27. Qi Y., Meng H., Hou D., Chen Y. A study on software rejuvenation model of application server cluster in two-dimension state space using Markov process. *Information Technology Journal*, 2008, vol. 7, no. 1, pp. 98–104. <https://doi.org/10.3923/itj.2008.98.104>
 28. Rahman P.A. Advanced reliability model of the fault-tolerant disk arrays with data striping and single disk redundancy. *Proc. of the International Scientific and Practical Conference*, 2017, pp. 20–25.
 29. Uspenskaya N.N. Estimation of availability factor for the data storage systems based on redundant disk arrays with the backup. *Proc. of the International Scientific and Practical Conference*, 2016, pp. 20–23.
 30. Rakhman P.A., Sharipov M.I. Reliability model of a two-node cluster of high-availability applications in enterprise management systems. *Economics and Management of Management Systems*. 2015. no. 3 (17). pp. 85–102. (in Russian)
 31. Khomonenko A.D., Blagoveshchenskaya E.A., Prouzrin O.V., Andruk A.A. Forecasting the reliability of a cluster computing system using a semi-Markov model of alternating processes and monitoring. *High Technologies in Earth Space Research. H&ES Research*, 2018, vol. 10, no. 4, pp. 72–82. (in Russian). <https://doi.org/10.24411/2409-5419-2018-10099>
 32. Terskov V. Sakash I. The reliability evaluation of local computer networks using markov model of multiple heterogeneous groups of switches. *E3s Web of Conferences*, 2024. vol. 592, pp. 03036.

Авторы


Богатырев Владимир Анатольевич — доктор технических наук, профессор, Санкт-Петербургский государственный университет аэрокосмического приборостроения, Санкт-Петербург, 190000, Российская Федерация; профессор, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, [sc 7006571069](https://orcid.org/0000-0003-0213-0223), <https://orcid.org/0000-0003-0213-0223>, vladimir.bogatyrev@gmail.com


Богатырев Станислав Владимирович — инженер-консультант, ООО «Центр разработки облачных хранилищ Ядро», Санкт-Петербург, 195027, Российская Федерация; аспирант, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, [sc 57183002200](https://orcid.org/0000-0003-0836-8515), <https://orcid.org/0000-0003-0836-8515>, realloc@gmail.com

Authors

Vladimir A. Bogatyrev — D.Sc., Professor, Saint Petersburg State University of Aerospace Instrumentation (SUAI), Saint Petersburg, 190000, Russian Federation; Professor, ITMO University, Saint Petersburg, 197101, Russian Federation, [sc 7006571069](https://orcid.org/0000-0003-0213-0223), <https://orcid.org/0000-0003-0213-0223>, vladimir.bogatyrev@gmail.com

Stanislav V. Bogatyrev — Consulting Engineer, Yadro Cloud Storage Development Center, Saint Petersburg, 195027, Russian Federation; PhD Student, ITMO University, Saint Petersburg, 197101, Russian Federation, [sc 57183002200](https://orcid.org/0000-0003-0836-8515), <https://orcid.org/0000-0003-0836-8515>, realloc@gmail.com

Богатырев Анатолий Владимирович — кандидат технических наук, инженер-консультант, ООО «Центр разработки облачных хранилищ Ядро», Санкт-Петербург, 195027, Российская Федерация,  56549712700, <https://orcid.org/0000-0001-5447-7275>, gangleon@gmail.com

Anatoly V. Bogatyrev — PhD, Consulting Engineer, Yadro Cloud Storage Development Center, Saint Petersburg, 195027, Russian Federation,  56549712700, <https://orcid.org/0000-0001-5447-7275>, gangleon@gmail.com

*Статья поступила в редакцию 18.01.2025
Одобрена после рецензирования 03.05.2025
Принята к печати 28.05.2025*

*Received 18.01.2025
Approved after reviewing 03.05.2025
Accepted 28.05.2025*



Работа доступна по лицензии
Creative Commons
«Attribution-NonCommercial»

Уважаемые подписчики научно-технической литературы!

Журнал выходит 6 раз в год.

На журнал «Научно-технический вестник информационных технологий, механики и оптики» можно оформить подписку в почтовых отделениях по объединенному каталогу «Пресса России», подписные индексы 47197 (полугодовая подписка),

а также online

по объединенному каталогу «Пресса России» и по каталогу «Пресса по подписке», подписные индексы Э47197 (полугодовая подписка) и по электронному каталогу Почты России, подписной индекс ПС543.

Корпоративная подписка и подписка физических лиц возможна по каталогу компаний «Урал-Пресс», подписные индексы 47197 (полугодовая подписка) и 70522 (годовая подписка).

Сведения о подписке можно уточнить в редакции журнала по адресу:

Санкт-Петербург, ул. Ломоносова, д.9, литера А, комн. 2136.

Тел.: +7(812) 480 02 75