

ISSN 1816-0301 (Print)
ISSN 2617-6963 (Online)

ИНФОРМАТИКА

INFORMATICS

TOM 23
VOL. 23

1 | 2026

ОТ РЕДАКЦИИ

В журнале «Информатика» публикуются оригинальные и обзорные статьи, описывающие результаты фундаментальных и прикладных исследований специалистов академического и вузовского профиля в области информатики и информационных технологий.

Основной целью журнала является публикация наиболее значимых новых результатов в указанной области. Приветствуются статьи, описывающие заключительные результаты научных проектов и диссертационных исследований, открывающие новые направления исследований, которые находятся на стыке информатики и других наук.

Журнал рассчитан на широкий круг специалистов в области информатики и информационных технологий.

Основные разделы журнала:

- биоинформатика;
- математическое моделирование;
- защита информации и надежность систем;
- информационные технологии;
- логическое проектирование;
- обработка сигналов, изображений, речи, текста и распознавание образов;
- автоматизация проектирования;
- интеллектуальные системы.

Префикс DOI: 10.37661

Условия распространения материалов:

контент доступен под лицензией Creative Commons Attribution 4.0 License

Индексирование:

Высшей аттестационной комиссией Республики Беларусь журнал «Информатика» был включен в список научных изданий для опубликования результатов диссертационных исследований.

В декабре 2017 г. включен в базу данных Российского индекса научного цитирования (РИНЦ). С помощью инструментов и сервисов, доступных на платформе eLIBRARY (раздел «Личный кабинет»), можно самостоятельно корректировать список своих публикаций и цитирований в РИНЦ.

В июле 2017 г. включен в базу журналов открытого доступа Directory of Open Access Journals (DOAJ).

С помощью поисковых систем Google Scholar, WorldCat, Соционет можно получить свободный доступ к полному тексту научных публикаций журнала.

Адрес редакции:

ул. Сурганова, 6, к. 305, г. Минск, 220012, Беларусь
Тел. +375 (017) 351 26 22

Editorial address:

Surganova str., 6, of. 305, Minsk, 220012, Belarus
Phone +375 (017) 351 26 22

E-mail: rio@newman.bas-net.by

<https://inf.grid.by/jour>

THE EDITOR'S NOTE

The journal "Informatics" is a scientific publication in computer sciences and information technologies which reviews the results in basic and applied research of scientists from the universities and scientific centers.

The journal focuses on the most significant and modern papers of research projects results and PhD/DSc thesis in computer sciences.

The journal is edited for the specialists in IT and computer sciences research and application.

The main sections of the journal:

- bioinformatics;
- mathematical modeling;
- information protection and system reliability;
- information technology;
- logical design;
- signal, image, speech, text processing and pattern recognition;
- computer-aided design;
- artificial intelligence methods.

DOI Prefix: 10.37661

Distribution:

content is distributed under Creative Commons Attribution 4.0 License

Indexation:

the journal "Informatics" is in the list of scientific publications recommended by the Higher Attestation Commission of the Republic of Belarus for scientists to publish the results of PhD/DSc research.

In December 2017 the journal was included in the database of the Russian Science Citation Index (RISC) and provides free access to reviewed electronic scientific paper, improving scientific information traffic and also raising quotation of works of the authors (please use <https://elibrary.ru> or section for authors https://elibrary.ru_author_tools).

In July 2017 included in the database of open access journals Directory of Open Access Journals (DOAJ).

Using the Google Scholar, WorldCat, Соционет search engine, you can get free access to full text of scientific publications of magazine.

ОБЪЕДИНЕННЫЙ ИНСТИТУТ ПРОБЛЕМ ИНФОРМАТИКИ
НАЦИОНАЛЬНОЙ АКАДЕМИИ НАУК БЕЛАРУСИ

ИНФОРМАТИКА

Informatika

Том 23, № 1, январь-март 2026

Ежеквартальный научный журнал

Издается с января 2004 г.

Учредитель и издатель – государственное научное учреждение «Объединенный институт проблем информатики Национальной академии наук Беларуси» (ОИПИ НАН Беларуси)

Г л а в н ы й р е д а к т о р

Кругликов Сергей Владимирович, д-р воен. наук, канд. техн. наук, доцент,
ОИПИ НАН Беларуси (Минск, Беларусь)

З а м е с т и т е л ь г л а в н о г о р е д а к т о р а

Тузиков Александр Васильевич, д-р физ.-мат. наук, проф., чл.-корр. НАН Беларуси,
ОИПИ НАН Беларуси (Минск, Беларусь)

Р е д а к ц и о н н а я к о л л е г и я

Абламейко Сергей Владимирович, д-р техн. наук, проф., академик НАН Беларуси, БГУ (Минск, Беларусь)

Анищенко Владимир Викторович, канд. техн. наук, доцент, ООО «СофтКлуб» (Минск, Беларусь)

Бобов Михаил Никитич, д-р техн. наук, проф., БГУИР (Минск, Беларусь)

Бондоловский Андрей Михайлович, канд. экон. наук, ОИПИ НАН Беларуси (Минск, Беларусь)

Голуб Юлия Игоревна, канд. техн. наук, ОИПИ НАН Беларуси (Минск, Беларусь)

Долгий Александр Борисович, д-р техн. наук, проф., Высшая инженерная школа Бретани (Нант, Франция)

Дудин Александр Николаевич, д-р физ.-мат. наук, проф., БГУ (Минск, Беларусь)

Карпов Алексей Анатольевич, д-р техн. наук, доцент, СПИИРАН (Санкт-Петербург, Россия)

Килин Сергей Яковлевич, д-р физ.-мат. наук, проф., академик НАН Беларуси, Центр «Квантовая оптика и квантовая информатика» Института физики им. Б. И. Степанова НАН Беларуси (Минск, Беларусь)

Ковалев Михаил Яковлевич, д-р физ.-мат. наук, проф., чл.-корр. НАН Беларуси, ОИПИ НАН Беларуси (Минск, Беларусь)

Краснопрошин Виктор Владимирович, д-р техн. наук, проф., БГУ (Минск, Беларусь)

Крот Александр Михайлович, д-р техн. наук, проф., ОИПИ НАН Беларуси (Минск, Беларусь)

Лиходед Николай Александрович, д-р физ.-мат. наук, проф., БГУ (Минск, Беларусь)

Матус Петр Павлович, д-р физ.-мат. наук, проф., Институт математики НАН Беларуси (Минск, Беларусь)

Скляров Валерий Анатольевич, д-р техн. наук, проф., Университет Авейру (Авейру, Португалия)

Сотсков Юрий Назарович, д-р физ.-мат. наук, проф., ОИПИ НАН Беларуси (Минск, Беларусь)

Харин Юрий Семенович, д-р физ.-мат. наук, проф., академик НАН Беларуси, НИИ ППМИ БГУ (Минск, Беларусь)

Черемисинова Людмила Дмитриевна, д-р техн. наук, проф., БГУИР (Минск, Беларусь)

Чернявский Александр Федорович, д-р техн. наук, проф., академик НАН Беларуси, НИИ ППМ им. А. Н. Севченко БГУ (Минск, Беларусь)

Ярмолик Вячеслав Николаевич, д-р техн. наук, проф., БГУИР (Минск, Беларусь)

Редакционный совет

Ефанов Дмитрий Викторович, Российский университет транспорта (Московский институт инженеров транспорта) (Москва, Россия)

Кумари Мадху, Университетский центр исследований и разработок, Университет Чандигарха (Мохали, Пенджаб, Индия)

Лазарев Александр Алексеевич, Институт проблем управления им. В. А. Трапезникова РАН (Москва, Россия)

Лай Цунг-Чьян, Азиатский университет в Тайчжуне (Китайская Народная Республика, Тайвань)

Марина Нинослав, Университет информационных наук и технологий им. Св. апостола Павла (Охрид, Македония)

Меликян Вазген Шаваршович, Национальный политехнический университет Армении (Ереван, Армения)

Пеш Эрвин, Зигенский университет (Зиген, Германия)

Сингх Таджиндер, Институт инженерии и технологий Сант Лонговал (Лонговал, Пенджаб, Индия)

Ходаченко Максим Леонидович, Институт космических исследований Австрийской академии наук (Грац, Австрия)

Чиулла Карло, Университет Эпока (Тирана, Албания)

Штейнберг Борис Яковлевич, Институт математики, механики и компьютерных наук Южного федерального университета (Ростов-на-Дону, Россия)

ИНФОРМАТИКА

Том 23, № 1, январь-март 2026

Ответственный за выпуск *Мойсейчик Светлана Сергеевна*
Редактор *Гончаренко Галина Борисовна*
Компьютерная верстка *Бутевич Ольга Борисовна*

Сдано в набор 23.02.2025. Подписано в печать 23.03.2025. Формат 60×84 1/8. Бумага офсетная. Ризография. Усл. печ. л. 12,1. Уч.-изд. л. 9,5. Тираж 40 экз. Заказ 1.

Издатель и полиграфическое исполнение:
государственное научное учреждение «Объединенный институт проблем информатики
Национальной академии наук Беларуси»
Свидетельство о государственной регистрации издателя, изготовителя, распространителя печатных изданий
№ 1/274 от 04.04.2014. ЛП № 38200000016516 от 18.12.13. Ул. Сурганова, 6, 220012, Минск, Беларусь.

ISSN 1816-0301 (Print)
ISSN 2617-6963 (Online)

THE UNITED INSTITUTE OF INFORMATICS PROBLEMS
OF THE NATIONAL ACADEMY OF SCIENCES OF BELARUS

INFORMATICS

Vol. 23, no. 1, January-March 2026

Published quarterly

Issued since January 2004

Founder and publisher – State Scientific Institution "The United Institute of Informatics
Problems of the National Academy of Sciences of Belarus" (UIIP NASB)

Editor-in-Chief

Sergey V. Kruglikov, Dr. Sci. (Milit.), Cand. Sci. (Eng.), Assoc. Prof., UIIP NASB (Minsk, Belarus)

Deputy Editor-in-Chief

Alexander V. Tuzikov, Dr. Sci. (Phys.-Math.), Prof., Corr. Member of NASB,
UIIP NASB (Minsk, Belarus)

Editorial Board

Sergey V. Ablameyko, Dr. Sci. (Eng.), Prof., Academician of NASB, BSU (Minsk, Belarus)

Uladimir V. Anishchanka, Cand. Sci. (Eng.), Assoc. Prof., SoftClub Ltd. (Minsk, Belarus)

Mikhail N. Bobov, Dr. Sci. (Eng.), Prof., BSUIR (Minsk, Belarus)

Andrei M. Bandalouski, Cand. Sci. (Econ.), UIIP NASB (Minsk, Belarus)

Yuliya I. Golub, Cand. Sci. (Eng.), UIIP NASB (Minsk, Belarus)

Alexandre B. Dolgui, Dr. Sci. (Eng.), Prof., IMT Atlantique (Nantes, France)

Alexander N. Dudin, Dr. Sci. (Phys.-Math.), Prof., BSU (Minsk, Belarus)

Alexey A. Karpov, Dr. Sci. (Eng.), Assoc. Prof., SPII RAS (Saint Petersburg, Russia)

Sergey Ya. Kilin, Dr. Sci. (Phys.-Math.), Prof., Academician of NASB, Center of Quantum Optics and Quantum
Information of B. I. Stepanov Institute of Physics NASB (Minsk, Belarus)

Mikhail Y. Kovalyov, Dr. Sci. (Phys.-Math.), Prof., Corr. Member of NASB, UIIP NASB (Minsk, Belarus)

Viktor V. Krasnoproshin, Dr. Sci. (Eng.), Prof., BSU (Minsk, Belarus)

Alexander M. Krot, Dr. Sci. (Eng.), Prof., UIIP NASB (Minsk, Belarus)

Nikolai A. Likhoded, Dr. Sci. (Phys.-Math.), Prof., BSU (Minsk, Belarus)

Petr P. Matus, Dr. Sci. (Phys.-Math.), Prof., Institute of Mathematics of NASB (Minsk, Belarus)

Valery A. Sklyarov, Dr. Sci. (Eng.), Prof., University of Aveiro (Aveiro, Portugal)

Yuri N. Sotskov, Dr. Sci. (Phys.-Math.), Prof., UIIP NASB (Minsk, Belarus)

Yuriy S. Kharin, Dr. Sci. (Phys.-Math.), Prof., Academician of NASB, RI APMI BSU (Minsk, Belarus)

Ljudmila D. Cheremisinova, Dr. Sci. (Eng.), Prof., BSUIR (Minsk, Belarus)

Alexander F. Cherniavsky, Dr. Sci. (Eng.), Prof., Academician of NASB, A. N. Sevchenko IAPP BSU (Minsk, Belarus)

Vyacheslav N. Yarmolik, Dr. Sci. (Eng.), Prof., BSUIR (Minsk, Belarus)

Editorial Council

Dmitry V. Efanov, Russian University of Transport (Moscow Institute of Transport Engineers) (Moscow, Russia)

Madhu Kumari, University Center for Research & Development, Chandigarh University (Mohali, Punjab, India)

Alexander A. Lazarev, V. A. Trapeznikov Institute of Control Sciences of the RAS (Moscow, Russia)

Tsung-Chyan Lai, Asia University at Taichung (The People's Republic of China, Taiwan)

Ninoslav Marina, St. Paul the Apostle University of Information Sciences and Technology (Ohrid, Macedonia)

Vazgen Sh. Melikyan, National Polytechnic University of Armenia (Yerevan, Armenia)

Erwin Pesch, University of Siegen (Siegen, Germany)

Tajinder Singh, Sant Longowal Institute of Engineering & Technology (Longowal, Punjab, India)

Maxim L. Khodachenko, Space Research Institute, Austrian Academy of Sciences (Graz, Austria)

Carlo Ciulla, Epoka University (Tirana, Albania)

Boris Steinberg, Institute of Mathematics, Mechanics and Computer Science Southern Federal University (Rostov-on-Don, Russia)

INFORMATICS

Vol. 23, no. 1, January-March 2026

Issue Head *Sviatlana S. Maiseichyk*

Editor *Halina B. Hancharenka*

Computer Imposition *Volha B. Butsevich*

Sent for press 23.02.2025. Output 23.03.2025. Format 60×84 1/8. Offset paper. Riesography.

Printed sheets 12,1. Publisher's signatures 9,5. Circulation 40 copies. Order 1.

Publisher and printing execution:

State Scientific Institution "The United Institute of Informatics Problems of the National Academy of Sciences of Belarus".

Certificate on the state registration of the publisher, manufacturer, distributor of printing editions no. 1/274 dated 04.04.2014. License for the press no. 38200000016516 dated 18.12.13.

6, Surganov Str., 220012, Minsk, Belarus.

ISSN 1816-0301 (Print)
ISSN 2617-6963 (Online)

СОДЕРЖАНИЕ

ЛОГИЧЕСКОЕ ПРОЕКТИРОВАНИЕ

Ярмолик В. Н., Мрозек И., Бранцевич П. Ю. Итерационный метод двухмерного масштабирования управляемых вероятностных тестов 7

ИНТЕЛЛЕКТУАЛЬНЫЕ СИСТЕМЫ

Ляхов Д. А., Бондоловский А. М., Кругликов С. В., Шульган К. К.
BellLitGPT – технологии языковых моделей для белорусского языка 26

ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ

Шульган А. К., Воротницкий Ю. И. Алгоритмы многокритериальной маршрутизации в беспроводных mesh-сетях 39

Лихачёв Д. С., Петровский Н. А., Азаров И. С. Метод повышения пространственного разрешения амбисонического звука на основе разреженного МДКП-представления 58

ОБРАБОТКА СИГНАЛОВ, ИЗОБРАЖЕНИЙ, РЕЧИ, ТЕКСТА И РАСПОЗНАВАНИЕ ОБРАЗОВ

Краснопрошин Д. В., Вашкевич М. И. Распознавание эмоций по речи на основе LSTM-сетей с мультивекторным механизмом внимания 69

Черников К. М., Суров И. А. Тональные пространства векторных языковых моделей 88

ISSN 1816-0301 (Print)
ISSN 2617-6963 (Online)

CONTENTS

LOGICAL DESIGN

Yarmolik V. N., Mrozek I., Brancevich P. Yu. An iterative method for two-dimensional scaling of controlled random tests 7

INTELLIGENT SYSTEMS

Lyakhov D. A., Bandalouski A. M., Kruglikov S. V., Shulgan K. K. BellLitGPT – language model technologies for the Belarusian language 26

INFORMATION TECHNOLOGIES

Shulgan A. K., Vorotnitsky Yu. I. Multicriteria routing algorithms in wireless mesh networks 39

Likhachov D. S., Petrovsky N. A., Azarov E. S. Method for spatial resolution enhancement of ambisonic audio based on sparse MDCT representation 58

SIGNAL, IMAGE, SPEECH, TEXT PROCESSING AND PATTERN RECOGNITION

Krasnoproshin D. V., Vashkevich M. I. Speech emotion recognition based on LSTM networks with multi-vector attention 69

Chernikov K. M., Surov I. A. Tonal spaces of vector language models 88

ЛОГИЧЕСКОЕ ПРОЕКТИРОВАНИЕ

LOGICAL DESIGN

УДК 004.33.054
<https://doi.org/10.37661/1816-0301-2026-23-1-7-25>

Поступила в редакцию | Received 14.12.2025
Подписана в печать | Accepted 08.01.2026
Опубликована | Published 31.03.2026

Итерационный метод двухмерного масштабирования управляемых вероятностных тестов

В. Н. Ярмолик^{1✉}, И. Мрозек², П. Ю. Бранцевич¹
✉E-mail: yarmolik10ru@yahoo.com

¹Белорусский государственный университет информатики и радиоэлектроники,
ул. П. Бровки, 6, Минск, 220013, Беларусь

²Белостокский технический университет,
ул. Вейска, 45А, 15-351, Белосток, Польша

Аннотация

Цели. Целями работы являются исследование ограниченности применения классических подходов генерирования тестовых наборов для управляемых вероятностных тестов, основанных на перечислении кандидатов в тестовые наборы путем их одномерного масштабирования, а также решение задачи построения управляемых вероятностных тестов на основе итерационного метода двухмерного масштабирования исходных шаблонов. Главной целью настоящей статьи является разработка метода построения тестов на базе исходных шаблонов и их расширения до требуемых разрядности и количества тестовых наборов с помощью применения итерационной процедуры.

Методы. Для двухмерного масштабирования исходных шаблонов с заданными характеристиками используются масштабирующие матрицы, которыми, так же как и шаблоны, могут быть управляемые вероятностные тесты. При проведении экспериментальных исследований применялся метод статистических испытаний.

Результаты. Показано, что методы построения управляемых вероятностных тестов, основанные на использовании шаблонов, можно рассматривать как процедуру масштабирования управляемых вероятностных тестов до требуемой их разрядности. Для построения искомым тестов применяются как шаблоны, характеризующиеся минимальной разрядностью тестовых наборов, так и любые управляемые вероятностные тесты. Подобная процедура позволяет увеличивать разрядность тестовых наборов, но сохраняет их количество. Одновременное увеличение разрядности наборов и их количества достигается с помощью предлагаемого подхода, основанного на итерационном двухмерном масштабировании шаблонов с применением масштабирующих матриц. В этом случае результирующие управляемые вероятностные тесты формируются без использования трудоемкой процедуры перечисления кандидатов в тестовые наборы и вычисления для них значений меры (мер) различия. Приведены зависимости основных характеристик результирующего управляемого вероятностного теста от характеристик шаблона и масштабирующей матрицы, которая, так же как и шаблон, может представлять собой управляемый вероятностный тест. Доказано утверждение, которое определяет зависимость характеристик теста, формируемого на k -й итерации, от значений характеристик теста, полученного на $(k-1)$ -й итерации, и масштабирующего теста. Представлены практически полезные следствия и свойства

тестов, построенных на основании предложенной процедуры. Работоспособность и эффективность итерационного метода построения управляемых вероятностных тестов оценены для случая двоичных тестовых наборов. Показано, что управляемые вероятностные тесты, построенные с применением рассмотренной процедуры, имеют заметно большие значения расстояний Хэмминга по сравнению с вероятностными тестами.

Заключение. Рассмотрен итерационный метод формирования управляемых вероятностных тестов путем их двухмерного масштабирования. Основой предложенного метода является использование исходных шаблонов и масштабирующих матриц, которые представляют собой управляемые вероятностные тесты с малым количеством тестовых наборов и небольшой их разрядностью. Показано, что использование различных шаблонов и их двухмерного масштабирования позволяет строить управляемые вероятностные тесты с требуемой разрядностью тестовых наборов и большим их количеством.

Ключевые слова: техническая диагностика, управляемые вероятностные тесты, двоичный тестовый набор, мера различия символьных наборов, расстояние Хэмминга

Для цитирования. Ярмолик, В. Н. Итерационный метод двухмерного масштабирования управляемых вероятностных тестов / В. Н. Ярмолик, И. Мрозек, П. Ю. Бранцевич // Информатика. – 2026. – Т. 23, № 1. – С. 7–25. – <https://doi.org/10.37661/1816-0301-2026-23-1-7-25>.

Конфликт интересов. В. Н. Ярмолик является членом редакционной коллегии журнала «Информатика» с 2010 г., но не имеет никакого отношения к решению опубликовать эту статью. Статья прошла принятую в журнале процедуру рецензирования. Об иных конфликтах интересов авторы не заявляли.

An iterative method for two-dimensional scaling of controlled random tests

Vyacheslav N. Yarmolik^{1✉}, Ireneusz Mrozek², Peter Yu. Brancevich¹

✉E-mail: yarmolik10ru@yahoo.com

¹*Belarusian State University of Informatics and Radioelectronics,
st. P. Brovki, 6, Minsk, 220013, Belarus*

²*Bialystok University of Technology,
Wiejska, 45A, 15-351, Białystok, Poland*

Abstract

Objectives. To study the limitations of classical approaches to generating test patterns for controlled random tests based on enumerating test set candidates through their one-dimensional scaling. To address the problem of constructing controlled random tests using an iterative method for two-dimensional scaling of initial templates. The main goal of the article is to develop a method for constructing tests based on initial templates and expanding them to the required bit size and number of test patterns using an iterative procedure.

Methods. For two-dimensional scaling of initial templates with given characteristics, scaling matrices are used, which, like templates, can also be controlled random tests. Statistical testing method was used during the experimental research.

Results. It is shown that methods for constructing controlled random tests based on the use of templates can be considered as a procedure for scaling controlled random tests to the required bit size. To construct the desired tests, both templates characterized by a minimum test suite capacity and any controllable random tests are used. This procedure allows increasing the test suite capacity while maintaining the number of their patterns. A simultaneous increase in the suite capacity and their number is achieved using the proposed approach, which is based on iterative two-dimensional scaling of templates using scaling matrices. In this case, the resulting controllable random tests are generated without the labor-intensive procedure of listing candidate test suites and calculating the difference measure(s) for them. The dependences of the main characteristics of the resulting controllable random

test on the characteristics of the template and the scaling matrix are presented, which, like a template, can also represent a controllable random test. A statement is proved that determines the dependence of the characteristics of the test generated at the k -th iteration on the values of the characteristics of the test obtained at the $(k-1)$ -th iteration and the scaling test. Useful consequences and properties of tests constructed based on the proposed procedure are presented. The performance and effectiveness of an iterative method for constructing controlled random tests are demonstrated and evaluated for binary test sets. It is shown that controlled random tests constructed using the discussed procedure have significantly larger Hamming distances compared to random tests.

Conclusion. An iterative method for constructing controlled random tests through two-dimensional scaling is considered. The basis of the proposed method is the use of initial templates and scaling matrices, which represent controlled random tests with a small number of test sets and a small bit size. It is shown that the use of various templates and their two-dimensional scaling allows for the construction of controlled random tests with the required bit size and a large number of test patterns.

Keywords: technical diagnostics, controlled random tests, binary test pattern, difference measure of symbol patterns, Hamming distance

For citation. Yarmolik V. N., Mrozek I., Brancevich P. Yu. *An iterative method for two-dimensional scaling of controlled random tests*. Informatika [Informatics], 2026, vol. 23, no. 1, pp. 7–25 (In Russ.). <https://doi.org/10.37661/1816-0301-2026-23-1-7-25>.

Conflict of interests. V. N. Yarmolik has been a member of the editorial board of the journal "Informatics" since 2010 but had no role in the decision to publish this article. The article has undergone the journal's established peer-review process. The authors declare of no conflicts of interest.

Введение

Автоматизация построения тестовых процедур для программного обеспечения, аппаратных средств и запоминающих устройств является важным и активно развивающимся научным и практическим направлением развития современных вычислительных систем [1–3]. Главный компонент подобных процедур – тестовая последовательность, которая должна характеризоваться высокой полнотой покрытия неисправностей вычислительных систем [3, 4]. Еще одним требованием для такой последовательности является небольшое количество в ней тестовых наборов, позволяющее осуществлять процедуру тестирования в реальном масштабе времени.

Большинство подходов к автоматизированному построению тестов для современных вычислительных систем реализуется с помощью различных форм их вероятностного формирования [5–8]. Простота концепции и реализуемости вероятностного формирования тестовых наборов служит основной причиной его широкого применения для тестирования вычислительных систем и их компонентов [3, 6, 9]. Главным недостатком вероятностного принципа генерирования тестовых наборов является его вычислительная сложность, возникающая из-за необходимости формирования большого количества тестовых наборов. Временные ограничения на реализацию процедур тестирования, и в первую очередь самотестирования, являются критическими для современных встроенных вычислительных систем, что также влияет и на эффективность тестов [2–4, 10]. Схожие проблемы возникают при построении тестовых процедур для современного программного обеспечения и запоминающих устройств, которые характеризуются функциональной сложностью, большими объемами и разнообразием структур организации и построения [11–14]. В связи с этим в настоящее время активно развиваются новые методы построения тестов, в которых случайный фактор подвергается различного рода детерминированным процедурам упорядочивания [10, 15, 16]. Доминирующее по-

ложение среди них занимает управляемое (адаптивное) вероятностное тестирование [16–18]. Этот вид построения тестов основан на вычислении некоторых характеристик для управляемого генерирования их тестовых наборов [19–24]. Тестовый набор выбирается из потенциальных кандидатов в тестовые наборы, сгенерированных случайным образом, по критерию максимальности некоторой либо некоторых характеристик, полученных на основании наборов, уже включенных в формируемый тест. Большинство таких методов построения тестов основано на применении расстояния Хэмминга в качестве характеристики, определяющей выбор очередного тестового набора [10–17]. Очевидно, что чем больше значения критериев выбора, в частности расстояния Хэмминга, тем сложнее процедура выбора тестовых наборов и может быть получено заметно меньшее их количество, которое в конечном счете определяет размерность формируемого управляемого вероятностного теста [15, 16, 22].

Высокая вычислительная сложность процедуры генерирования управляемых вероятностных тестов, требующая выбора тестового набора из большого числа кандидатов в тесты, является главным ее недостатком. Только для определенных значений расстояния Хэмминга и при ряде других ограничений удастся избежать рутинной процедуры поиска тестовых наборов среди ограниченного количества кандидатов в тесты, удовлетворяющих заданным критериям [3, 11, 15].

Главной целью настоящей статьи являются разработка и развитие методов построения управляемых вероятностных тестов с заданным расстоянием Хэмминга как критерием включения кандидата в генерируемый тест, которые характеризуются невысокой вычислительной сложностью. Основными задачами, решаемыми авторами для достижения невысокой вычислительной сложности процедуры синтеза тестов, являются разработка формальных подходов и методов формирования управляемых вероятностных тестов и аналитическое определение их характеристик.

1. Анализ управляемых вероятностных тестов

В области формирования управляемых вероятностных тестов концептуально принимается гипотеза, что для двух тестовых наборов, имеющих максимальное отличие, общее количество обнаруживаемых неисправностей (ошибок) будет максимальным [3, 15–17]. Обобщением данной гипотезы является формирование очередного тестового набора управляемого вероятностного теста, максимально отличающегося от наборов, ранее включенных в тест. В качестве критерия отличия очередного тестового набора T_i от предыдущих наборов T_0, T_1, \dots, T_{i-1} наиболее часто используется расстояние Хэмминга (Hamming distance) $HD(T_i, T_j)$ для $j \in \{0, 1, \dots, i-1\}$. Величина $HD(T_i, T_j)$ определяется числом несовпадающих символов в одноименных разрядах наборов T_i и T_j произвольного алфавита [25, 26]. В качестве критерия выбора тестового набора из множества кандидатов в тестовые наборы используются как пороговое значение расстояния Хэмминга $d = \min HD(T_i, T_j)$, так и его суммарное значение $totalHD(T_i, T_j)$, а также среднее значение $aveHD(T_i, T_j)$ [3, 15–17].

Не нарушая общности изложения материала, будем рассматривать случай управляемого вероятностного теста T_0, T_1, \dots, T_{q-1} , включающего q тестовых наборов $T_i = t_{i,0}, t_{i,1}, \dots, t_{i,r-1}$, где $t_{i,l} \in \{0, 1\}$, для $i \in \{0, 1, \dots, q-1\}$ и $l \in \{0, 1, \dots, r-1\}$. Основываясь на пороговом значении расстояния Хэмминга $d = \min HD(T_i, T_j)$ ($i \neq j \in \{0, 1, \dots, q-1\}$) в качестве критерия выбора тестовых наборов, управляемый вероятностный тест

$CRT(q, d, r)$ описывается тремя параметрами: количеством тестовых наборов q , пороговым значением расстояния Хэмминга d и их разрядностью r . Приведенное определение управляемого вероятностного теста $CRT(q, d, r)$ позволило использовать ранее полученные результаты как в области тестового диагностирования, так и в области помехоустойчивого кодирования для их синтеза [3, 15–17, 25–28]. Примеры множеств двоичных наборов с малой разрядностью r , которые интерпретируются как управляемые вероятностные тесты $CRT(q, d, r) = T_0, T_1, \dots, T_{q-1}$, приведены в табл. 1.

Таблица 1
Примеры управляемых вероятностных тестов

Table 1
Examples of control random tests

$CRT(4, 2, 3)$ ($MMHD(4)$)	$CRT(4, 1, 2)$ ($OCRT_1$)	$CRT(6, 2, 4)$ ($OCRT_2$)	$CRT(5, 2, 4)$ ($PExT_1(2, 4)$)	$CRT(6, 2, 4)$ ($PExT_2(2, 4)$)
$T_0=111$	$T_0=11$	$T_0=1111$	$T_0=1111$	$T_0=0011$
$T_1=100$	$T_1=10$	$T_1=0000$	$T_1=1000$	$T_1=1001$
$T_2=010$	$T_2=01$	$T_2=1100$	$T_2=0100$	$T_2=1100$
$T_3=001$	$T_3=00$	$T_3=0011$	$T_3=0010$	$T_3=0101$
		$T_4=1010$	$T_4=0001$	$T_4=1010$
		$T_5=0101$		$T_5=0110$

Тест $CRT(4, 2, 3)$ представляет собой $MMHD(q)$ (*Maximum Minimum Hamming Distance*)-тест для $q = 4$ с максимальным минимальным расстоянием Хэмминга $\min HD(T_i, T_j) = 2$ [3, 22, 27]. Этот тест характеризуется тем, что для любой пары тестовых наборов сохраняется расстояние Хэмминга, равное 2 [22].

Тесты $CRT(4, 1, 2) = OCRT_1$ и $CRT(6, 2, 4) = OCRT_2$ (*Optimal Controlled Random Tests*) построены с применением двух характеристик различия, а именно расстояния Хэмминга $HD(T_i, T_j)$ и евклидова расстояния $CD(T_i, T_j)$. Они относятся к множеству оптимальных управляемых тестов [3]. Данные тесты характеризуются тем, что для них величины расстояний $HD(T_i, T_j) \geq r/2$ и принимают различные значения. Например, для теста $CRT(6, 2, 4) = OCRT_2$ имеем $HD(T_0, T_1) = 4$, а $HD(T_0, T_2) = 2$, но при этом сохраняется выполнение неравенства $HD(T_i, T_j) \geq 4/2$. Последующие тесты $CRT(5, 2, 4) = PExT_1(2, 4)$ и $CRT(6, 2, 4) = PExT_2(2, 4)$, приведенные в табл. 1, относятся к множеству псевдоисчерпывающих тестов $PExT(k, r)$ (*Pseudo-Exhaustive Tests*), где $k < r$ [3, 11, 28]. Подобные тесты представляют собой множество двоичных наборов, обеспечивающих всевозможные 2^k двоичные комбинации на любых k из r разрядов его тестовых наборов. Оба приведенных псевдоисчерпывающих теста обеспечивают всевозможные двоичные комбинации $\{00, 01, 10, 11\}$ на любых $k = 2$ из $r = 4$ разрядов тестовых наборов.

Рассмотренные в табл. 1 примеры управляемых вероятностных тестов для малых разрядностей r позволяют пояснить понятие шаблона управляемого вероятностного теста $CRT_m(q_t, d_t, r_t)$. Для обозначения шаблона и его характеристик используются индексы, соответствующие его англоязычному смысловому описанию – *template*. Каждый из приведенных примеров можно рассматривать как шаблон $CRT_m(q_t, d_t, r_t)$, под которым будем понимать произвольный управляемый вероятностный тест $CRT(q, d, r)$ с фиксиро-

ваным количеством тестовых наборов q_t , заданным значением порогового расстояния Хэмминга d_t и разрядностью $r_t < m$ тестовых наборов [29, 30]. Значение m определяет разрядность тестовых наборов теста, для формирования которого используется шаблон $CRT_m(q_t, d_t, r_t)$. Применение шаблонов позволяет исключить рутинную процедуру поиска очередного тестового набора из множества потенциальных кандидатов [29, 30]. Управляемый вероятностный тест, используемый как шаблон, применяется для синтеза аналогичных тестов с требуемой разрядностью m на основании правил одномерного масштабирования [28, 29]. Для этой процедуры могут быть использованы как шаблоны, построенные для минимальных разрядностей тестовых наборов, так и любые управляемые вероятностные тесты, для которых $r < m$. Основным недостатком одномерного масштабирования является неизменность количества тестовых наборов, от которой напрямую зависит полнота покрытия неисправностей, обнаруживаемых данным тестом. Как и для вероятностных тестов, полнота покрытия, обеспечиваемая управляемыми вероятностными тестами, увеличивается с возрастанием количества тестовых наборов [3, 8].

С целью одновременного увеличения разрядности тестовых наборов и их количества в работе [29] был рассмотрен подход, основанный на двухмерном масштабировании исходных шаблонов с применением матриц Адамара. Подобные матрицы позволяют выполнять процедуру масштабирования в $n \in \{2, 4, 8, \dots\}$ раз в зависимости от порядка $n = 2^c$ используемой матрицы, где $c \in \mathbb{N}$. Само масштабирование заключается в применении в качестве элемента (+1) оригинальной матрицы Адамара масштабируемого теста $CRT_m(q_t, d_t, r_t)$, а вместо элемента (-1) – его инверсного представления $\overline{CRT_m(q_t, d_t, r_t)}$. Для общего случая двухмерного масштабирования шаблонов либо произвольных управляемых вероятностных тестов с использованием матриц Адамара справедливо следующее утверждение [29].

Утверждение 1. Результатом масштабирования управляемого вероятностного теста $CRT_m(q_t, d_t, r_t)$ при $d_t \leq r_t/2$ с помощью матрицы Адамара n -го порядка является тест $CRT(n \cdot q_t, n \cdot d_t, n \cdot r_t)$, а при $d_t > r_t/2$ – тест $CRT(n \cdot q_t, (n \cdot r_t)/2, n \cdot r_t)$.

Из приведенного утверждения следует, что в отличие от подхода, основанного на одномерном масштабировании, результирующий управляемый вероятностный тест в случае применения матриц Адамара порядка n содержит в n раз больше тестовых наборов. Сама процедура масштабирования в сравнении с классическим подходом построения управляемых вероятностных тестов не требует вычисления меры (мер) различия, что существенно уменьшает вычислительные затраты. Отсутствие необходимости перечисления кандидатов в тестовые наборы сводит задачу синтеза управляемого вероятностного теста к формальной автоматической процедуре [29].

Основываясь на идее двухмерного масштабирования, в работе [30] рассмотрен подход его применения в рамках управляемых вероятностных тестов, когда масштабирующая матрица (*scaling*), так же как и шаблон $CRT_m(q_t, d_t, r_t)$, представляет собой управляемый вероятностный тест $CRT_{sc}(q_s, d_s, r_s)$. Таким образом, в результате двухмерного масштабирования формируется результирующий (*output*) управляемый вероятностный тест $CRT_{ou}(q_o, d_o, r_o)$, характеристики которого определяются следующим утверждением, справедливым для $d_s \geq r_s/2$.

Утверждение 2. Результатом масштабирования управляемого вероятностного теста $CRT_m(q_t, d_t, r_t)$ с помощью масштабирующего теста $CRT_{sc}(q_s, d_s, r_s)$ является тест $CRT(q_s \cdot q_t, d_o, r_s \cdot r_t)$, где $d_o = \min\{(r_s \cdot d_t), (r_t \cdot d_s), (r_s \cdot b_t + r_t \cdot b_s - 2 \cdot (b_s \cdot b_t))\}$.

Для определения d_o использовались как пороговое (минимальное) расстояние $d = \min HD(T_i, T_j)$, так и максимальное расстояние Хэмминга $b = \max HD(T_i, T_j)$. Соответственно, для теста $CRT_{sc}(q_s, d_s, r_s)$ применялось значение $b_s = \max HD(T_i, T_j)$, $i \neq j \in \{0, 1, \dots, q_s - 1\}$, а для $CRT_m(q_t, d_t, r_t)$ – значение $b_t = \max HD(T_i, T_j)$, $i \neq j \in \{0, 1, \dots, q_t - 1\}$. Для теста $CRT(4, 2, 3)$, приведенного в табл. 1, $b = d = 2$, а для теста $CRT(4, 1, 2)$ $b = 2, d = 1$.

Очевидно, что возможны разнообразные сочетания управляемых вероятностных тестов, используемых как для масштабирования, так и для шаблонов (масштабируемых тестов).

Для примера двумерного масштабирования с применением в качестве масштабирующей матрицы теста $CRT_{sc}(q_s, d_s, r_s) = CRT_{sc}(4, 2, 3)$, приведенного в табл. 1, рассмотрим три масштабируемых шаблона, один из которых повторяет тест $CRT(4, 2, 3)$ (см. табл. 1). В качестве второго шаблона используем $CRT_m(4, 1, 3)$, состоящий также из четырех трехразрядных наборов $T_0 = 0\ 0\ 0$, $T_1 = 0\ 0\ 1$, $T_2 = 0\ 1\ 0$ и $T_3 = 1\ 1\ 1$. В качестве третьего шаблона применим нулевую двоичную матрицу размерностью 4×3 , которая состоит только из нулевых элементов и описывается как $CRT_m(4, 0, 3)$. Основные характеристики, такие как d_t , b_t , сумма всех расстояний Хэмминга ($totalHD(T_i, T_j)$) и их средняя величина ($aveHD(T_i, T_j)$), для указанных шаблонов приведены в табл. 2.

Таблица 2

Характеристики управляемых вероятностных тестов $CRT_m(4, 2, 3)$, $CRT_m(4, 1, 3)$ и $CRT_m(4, 0, 3)$

Table 2

Characteristics of control random tests $CRT_m(4, 2, 3)$, $CRT_m(4, 1, 3)$ and $CRT_m(4, 0, 3)$

$CRT_m(q_t, d_t, r_t)$	$CRT_m(4, 2, 3)$	$CRT_m(4, 1, 3)$	$CRT_m(4, 0, 3)$
$d_t = \min HD(T_i, T_j)$	2	1	0
$b_t = \max HD(T_i, T_j)$	2	3	0
$totalHD(T_i, T_j)$	12	11	0
$aveHD(T_i, T_j)$	2	$11/6 = 1,83$	0
$aveHD(T_i, T_j)/r_t$	$2/3 = 0,666$	$1,33/3 = 0,611$	0

Значение $aveHD(T_i, T_j)/r_t$ является приведенным значением среднего расстояния Хэмминга $aveHD(T_i, T_j)$ к одному разряду r_t -разрядного тестового набора.

Отличием рассмотренных шаблонов является их значение $d_t = \min HD(T_i, T_j)$, которое в первом случае равняется двум и удовлетворяет неравенству $d_t > r/2$, а во втором равняется единице, что соответствует соотношению $d_t \leq r/2$. В третьем случае имеем пример вырожденного шаблона, который мог быть получен в результате его вероятностного формирования и для которого $d_t = 0$.

Результаты двумерного масштабирования всех трех шаблонов приведены в табл. 3. Как и в случае масштабирования с помощью матриц Адамара, вместо нулевого значения масштабирующей матрицы $CRT_{sc}(4, 2, 3)$ подставляется соответствующий шаблон, а единичное значение замещается инверсным шаблоном. Для произвольного шаблона указанная процедура замещения показана во втором столбце табл. 3.

Таблица 3

Двухмерное масштабирование двоичных шаблонов $CRT_{im}(4, 2, 3)$, $CRT_{im}(4, 1, 3)$ и $CRT_{im}(4, 0, 3)$

Table 3

Two-dimensional scaling of binary template $CRT_{im}(4, 2, 3)$, $CRT_{im}(4, 1, 3)$ and $CRT_{im}(4, 0, 3)$

$CRT_{sc}(4, 2, 3)$	$CRT_{ou}(q_o, d_o, r_o)$	T_i	$CRT_{ou}(16, 4, 9)$ для / for $CRT_{im}(4, 2, 3)$	$CRT_{ou}(16, 3, 9)$ для / for $CRT_{im}(4, 1, 3)$	$CRT_{ou}(16, 0, 9)$ для / for $CRT_{im}(4, 0, 3)$
0 0 0	$CRT_{im} CRT_{im} CRT_{im}$	T_0	000 000 000	000 000 000	000 000 000
		T_1	011 011 011	001 001 001	000 000 000
		T_2	101 101 101	010 010 010	000 000 000
		T_3	110 110 110	111 111 111	000 000 000
0 1 1	$CRT_{im} \overline{CRT_{im}} \overline{CRT_{im}}$	T_4	000 111 111	000 111 111	000 111 111
		T_5	011 100 100	001 110 110	000 111 111
		T_6	101 010 010	010 101 101	000 111 111
		T_7	110 001 001	111 100 100	000 111 111
1 0 1	$\overline{CRT_{im}} CRT_{im} \overline{CRT_{im}}$	T_8	111 000 111	111 000 111	111 000 111
		T_9	100 011 100	110 001 110	111 000 111
		T_{10}	010 101 010	101 010 101	111 000 111
		T_{11}	001 110 001	100 111 100	111 000 111
1 1 0	$\overline{CRT_{im}} \overline{CRT_{im}} CRT_{im}$	T_{12}	111 111 000	111 111 000	111 111 000
		T_{13}	100 100 011	110 110 001	111 111 000
		T_{14}	010 010 101	101 101 010	111 111 000
		T_{15}	001 001 110	100 100 111	111 111 000

Из табл. 3 видно, что значения q_o и r_o теста $CRT_{ou}(q_o, d_o, r_o)$ увеличиваются соответственно в $q_s = 4$ и $r_s = 3$ раз по сравнению с величинами $q_t = 4$ и $r_t = 3$ для трех масштабируемых шаблонов. Более сложная зависимость наблюдается для порогового значения расстояния Хэмминга d_o , которое зависит как от вида масштабирующей матрицы, так и от масштабируемого шаблона и соответствует утверждению 2 [30].

2. Итерационный метод синтеза управляемых вероятностных тестов

В качестве основы итерационного подхода построения управляемых вероятностных тестов используем метод построения подобных тестов путем двухмерного масштабирования шаблонов [29, 30]. Для описания управляемых вероятностных тестов, основываясь на утверждении 2, в дальнейшем будем применять еще одну их характеристику, а именно максимальное расстояние Хэмминга $b = \max HD(T_i, T_j)$. Тогда масштабирующий управляемый вероятностный тест, шаблон и результат двухмерного масштабирования будут описываться как $CRT_{sc}(q_s, d_s, b_s, r_s)$, $CRT_{im}(q_t, d_t, b_t, r_t)$ и $CRT_{ou}(q_o, d_o, b_o, r_o)$.

Ранее уже отмечалось, что возможны различные варианты сочетания управляемых вероятностных тестов, используемых в качестве масштабирующих матриц и масштабируемых шаблонов. В дальнейшем будем рассматривать схему итерационного масштабирования для фиксированного масштабирующего теста $CRT_{sc}(q_s, d_s, b_s, r_s)$, который применяется для получения на основании теста $CRT_{ou(k-1)}(q_{o(k-1)}, d_{o(k-1)}, b_{o(k-1)}, r_{o(k-1)})$, представляющего собой шаблон для k -й итерации, результирующего теста $CRT_{ou(k)}(q_{o(k)}, d_{o(k)}, b_{o(k)}, r_{o(k)})$. Для первой итерации ($k = 1$) в качестве $CRT_{ou_0}(q_{o(0)}, d_{o(0)}, b_{o(0)}, r_{o(0)})$ используем шаблон, которым может быть управляемый вероятностный тест $CRT_{im}(q_t, d_t,$

$b_t, r_t), q_t > 1$. Один тестовый набор $T_0 (q_t = 1)$, представляющий собой двоичный вектор произвольной разрядности $r_t \geq 1$, также может быть применен в качестве шаблона, который описывается как $CRT_{tm}(1, 0, 0, r_t)$.

Весьма важный вопрос заключается в выборе фиксированного масштабирующего теста $CRT_{sc}(q_s, d_s, b_s, r_s)$. Как отмечалось ранее, критерием включения в тест кандидата в тестовые наборы является его максимальное отличие (удаление) от ранее сгенерированных наборов, определяемое расстоянием Хэмминга. Поэтому в качестве масштабирующего управляемого вероятностного теста $CRT_{sc}(q_s, d_s, r_s)$ необходимо применять тесты, для которых d_s как пороговое значение меры различия принимает максимально возможное значение при небольших величинах r_s и q_s , требующих их масштабирования. Как и для примера, приведенного в табл. 3, используем тест $CRT_{sc}(4, 2, 3)$ в качестве масштабирующего теста $CRT_{sc}(q_s, d_s, b_s, r_s)$. Данный тест согласно теории помехоустойчивого кодирования характеризуется максимальным значением величины $d = 2$ для $q = 4$ и $r = 3$ [3, 26, 28].

Основываясь на масштабирующем тесте $CRT_{sc}(4, 2, 3)$, представим итерационную процедуру синтеза в табл. 4.

Таблица 4
Итерационная процедура построения управляемых вероятностных тестов

Table 4
An iterative procedure for constructing controlled random tests

$CRT_{sc}(4, 2, 3)$	$CRT_{ou(1)}$...	$CRT_{ou(k)}$...
$T_0 = 0 \ 0 \ 0$	CRT_{tm}	CRT_{tm}	CRT_{tm}	...	$CRT_{ou(k-1)}$	$CRT_{ou(k-1)}$	$CRT_{ou(k-1)}$...
$T_1 = 0 \ 1 \ 1$	CRT_{tm}	$\overline{CRT_{tm}}$	$\overline{CRT_{tm}}$...	$CRT_{ou(k-1)}$	$\overline{CRT_{ou(k-1)}}$	$\overline{CRT_{ou(k-1)}}$...
$T_2 = 1 \ 0 \ 1$	$\overline{CRT_{tm}}$	CRT_{tm}	$\overline{CRT_{tm}}$...	$\overline{CRT_{ou(k-1)}}$	$CRT_{ou(k-1)}$	$\overline{CRT_{ou(k-1)}}$...
$T_3 = 1 \ 1 \ 0$	$\overline{CRT_{tm}}$	$\overline{CRT_{tm}}$	CRT_{tm}	...	$\overline{CRT_{ou(k-1)}}$	$\overline{CRT_{ou(k-1)}}$	$CRT_{ou(k-1)}$...

Результаты применения $CRT_{sc}(4, 2, 3)$ в качестве масштабирующего теста для первой итерации и трех различных шаблонов приведены в табл. 3, а характеристики результирующих тестов $CRT_{ou_1}(q_{o(1)}, d_{o(1)}, b_{o(1)}, r_{o(1)}) = CRT_{ou}(q_o, d_o, b_o, r_o)$ – в табл. 5.

Таблица 5
Характеристики тестов, построенных для шаблонов $CRT_{tm}(4, 2, 2, 3)$, $CRT_{tm}(4, 1, 2, 3)$ и $CRT_{tm}(4, 0, 0, 3)$

Table 5
Characteristics of tests built for templates $CRT_{tm}(4, 2, 0, 3)$, $CRT_{tm}(4, 1, 3, 3)$ and $CRT_{tm}(4, 0, 0, 3)$

$CRT_{ou}(q_o, d_o, b_o, r_o)$	$CRT_{ou}(16, 4, 6, 9)$ для / for $CRT_{tm}(4, 2, 2, 3)$	$CRT_{ou}(16, 3, 9, 9)$ для / for $CRT_{tm}(4, 1, 3, 3)$	$CRT_{ou}(16, 0, 6, 9)$ для / for $CRT_{tm}(4, 0, 0, 3)$
$d_o = \min HD(T_i, T_j)$	4	3	0
$b_o = \max HD(T_i, T_j)$	6	9	6
$total HD(T_i, T_j)$	576	576	576
$ave HD(T_i, T_j)$	$576/120 = 4,8$	$576/120 = 4,8$	$576/120 = 4,8$
$ave HD(T_i, T_j)/r_o$	$4,8/9 = 0,533$	$4,8/9 = 0,533$	$4,8/9 = 0,533$

Анализ приведенных результатов показывает нивелирование отличий характеристик использованных шаблонов (см. табл. 2). Действительно, как видно из табл. 5, в результирующих тестах $CRT_{ou}(q_o, d_o, b_o, r_o)$ отличаются только значения $d_o = \min HD(T_i, T_j)$ и $b_o = \max HD(T_i, T_j)$, а остальные характеристики принимают одинаковые величины даже для абсолютно неприемлемого шаблона $CRT_{tm}(4, 0, 0, 3)$, состоящего только из нулевых элементов. Для всевозможных шаблонов $CRT_{tm}(q_t, d_t, b_t, r_t)$, для которых $d_t > 0$, был проведен исчерпывающий анализ характеристик результирующих тестов $CRT_{ou}(q_o, d_o, b_o, r_o)$, полученных для $CRT_{sc}(q_s, d_s, b_s, r_s) = CRT_{sc}(4, 2, 2, 3)$, т. е. использовались только такие шаблоны, в которых отсутствуют повторяющиеся тестовые наборы. Результаты показали, что для тестов $CRT_{ou}(q_o, d_o, b_o, r_o)$ величины их характеристик d_o и b_o принадлежат диапазонам $3 \leq d_o \leq 4$ и $6 \leq b_o \leq 9$, а $aveHD(T_i, T_j)$ во всех случаях принимает постоянное значение, равное 4, 8.

Из табл. 3 и 5 видно, что параметры $q_{o(k)}$ и $r_{o(k)}$ результирующего теста $CRT_{ou_k}(q_{o(k)}, d_{o(k)}, b_{o(k)}, r_{o(k)})$, сформированного на k -й итерации процедуры синтеза тестов, очевидным образом зависят от q_s и r_s масштабирующего теста $CRT_{sc}(q_s, d_s, b_s, r_s)$. Соответственно, $q_{o(k)} = q_s \cdot q_{o(k-1)}$, а $r_{o(k)} = r_s \cdot r_{o(k-1)}$. Более сложная зависимость наблюдается для минимального $d_{o(k)}$ и максимального $b_{o(k)}$ значений расстояния Хэмминга. Для оценки указанных характеристик примем гипотезу неизменности масштабирующего управляемого вероятностного теста $CRT_{sc}(q_s, d_s, b_s, r_s)$ для всех итераций процедуры построения результирующего теста $CRT_{ou_k}(q_{o(k)}, d_{o(k)}, b_{o(k)}, r_{o(k)})$. Этот тест может быть и отличным от $CRT_{sc}(4, 2, 2, 3)$, поэтому определим зависимость характеристик $d_{o(k)}$ и $b_{o(k)}$ от характеристик как произвольного масштабирующего теста, так и произвольного шаблона. Отметим, что в качестве шаблона на k -й итерации применяется тест $CRT_{ou(k-1)}(q_{o(k-1)}, d_{o(k-1)}, b_{o(k-1)}, r_{o(k-1)})$, сформированный на предыдущей итерации.

Для общего случая итерационной процедуры двухмерного масштабирования с целью получения управляемого вероятностного теста справедливо утверждение 3, которое во многом основывается на следующих свойствах расстояния Хэмминга:

$$\begin{aligned} HD(T_i, T_i) = 0; \quad HD(T_i, \bar{T}_i) = r; \quad HD(T_i, T_j) = HD(\bar{T}_i, \bar{T}_j); \\ HD(T_i, \bar{T}_j) = HD(\bar{T}_i, T_j); \quad HD(T_i, T_j) + HD(T_i, \bar{T}_j) = r. \end{aligned} \quad (1)$$

Приведенное ниже утверждение определяет зависимость характеристик результата $CRT_{ou_k}(q_{o(k)}, d_{o(k)}, b_{o(k)}, r_{o(k)})$ k -й итерации двухмерного масштабирования шаблона $CRT_{ou(k-1)}(q_{o(k-1)}, d_{o(k-1)}, b_{o(k-1)}, r_{o(k-1)})$ с помощью масштабирующего теста $CRT_{sc}(q_s, d_s, b_s, r_s)$.

Утверждение 3. *Результатом k -й итерации двухмерного масштабирования шаблона $CRT_{ou(k-1)}(q_{o(k-1)}, d_{o(k-1)}, b_{o(k-1)}, r_{o(k-1)})$ с помощью масштабирующего теста $CRT_{sc}(q_s, d_s, b_s, r_s)$ является тест $CRT_{ou_k}(q_s \cdot q_{o(k-1)}, d_{o(k)}, b_{o(k)}, r_s \cdot r_{o(k-1)})$, где для $d_s \geq r_s/2$ значение $d_{o(k)} = \min\{(r_s \cdot d_{o(k-1)}), (r_{o(k-1)} \cdot d_s), (r_s \cdot b_{o(k-1)} + r_{o(k-1)} \cdot b_s - 2 \cdot (b_s \cdot b_{o(k-1)}))\}$, значение $b_{o(k)} = \max\{(r_s \cdot b_{o(k-1)}), (r_{o(k-1)} \cdot b_s), (r_s \cdot d_{o(k-1)} + r_{o(k-1)} \cdot d_s - 2 \cdot (d_s \cdot d_{o(k-1)}))\}$, а для $d_s < r_s/2$ значение $d_{o(k)} = \min\{(r_s \cdot d_{o(k-1)}), (r_{o(k-1)} \cdot d_s), (r_s \cdot d_{o(k-1)} + r_{o(k-1)} \cdot d_s - 2 \cdot (d_s \cdot d_{o(k-1)}))\}$, $b_{o(k)} = \max\{(r_s \cdot b_{o(k-1)}), (r_{o(k-1)} \cdot b_s), (r_s \cdot b_{o(k-1)} + r_{o(k-1)} \cdot b_s - 2 \cdot (b_s \cdot b_{o(k-1)}))\}$.*

Доказательство. Результирующий тест CRT_{ou_k} содержит $q_{o(k)} = q_s \cdot q_{o(k-1)}$ наборов, каждый из которых состоит из $r_{o(k)} = r_s \cdot r_{o(k-1)}$ двоичных разрядов. Это следует из того факта, что при масштабировании разряды масштабирующего теста $CRT_{sc}(q_s, d_s, b_s, r_s)$

заменяются шаблоном либо его инверсией. Инверсное представление шаблона и сам шаблон состоят из $q_{o(k-1)}$ наборов, каждый из которых представлен $r_{o(k-1)}$ разрядами. В матричном представлении результатом масштабирования является матрица размерностью $(q_s \cdot q_{o(k-1)}) \times (r_s \cdot r_{o(k-1)})$, в которой строки группируются в блоки, состоящие из $q_{o(k-1)}$ строк. Каждый из q_s блоков определяется строкой масштабирующей матрицы теста $CRT_{sc}(q_s, d_s, r_s)$, используемой для масштабирования так, как это показано в табл. 3 и 4.

Рассмотрим значение величины расстояния Хэмминга $h = HD(T_i, T_j)$ между двумя произвольными строками теста CRT_{ou_k} , полученного в результате k -й итерации масштабирования. Строки текста представляют собой два тестовых набора $T_k \neq T_j \in \{T_0, T_1, \dots, T_{q_{o(k-1)}-1}, T_{q_{o(k-1)}}, T_{q_{o(k-1)}+1}, \dots, T_{2q_{o(k-1)}-1}, \dots\}$, каждый из которых состоит из $r_s \cdot r_{o(k-1)}$ двоичных значений. Например, для случая, приведенного в табл. 3, $T_k \neq T_j \in \{T_0, T_1, \dots, T_{15}\}$, а два произвольных тестовых набора T_i и T_j могут принадлежать как одному из $q_s = 4$ блоков, так и разным блокам.

Структура каждого из q_s блоков повторяет результат применения процедуры одномерного масштабирования, когда для увеличения разрядности используется повторение шаблона либо его инверсного представления [29]. В этом случае сохраняется отношение значения расстояния Хэмминга к разрядности наборов. Соответственно, если тестовые наборы принадлежат одному из q_s блоков строк, то для них сохраняются значения $h = HD(T_i, T_j)$. Например, для всех трех примеров масштабирования, приведенных в табл. 3, $HD(T_1, T_3) = HD(T_5, T_7) = HD(T_9, T_{11}) = HD(T_{13}, T_{15})$ и принимают значения 6, 3 и 0 соответственно. Постоянство величины h в данном случае объясняется свойством расстояния Хэмминга, задаваемым равенством $HD(T_i, T_j) = HD(\bar{T}_i, \bar{T}_j)$. Среди множества величин h для тестовых наборов T_i и T_j , принадлежащих одному блоку, можно выделить минимальное значение, равное $r_s \cdot d_{o(k-1)}$, и максимальное $r_s \cdot b_{o(k-1)}$, определяемые минимальным $d_{o(k-1)}$ и максимальным $b_{o(k-1)}$ значениями h шаблона $CRT_{ou(k-1)}$.

В случае когда тестовые наборы T_i и T_j принадлежат разным блокам, идентифицированным различными строками масштабирующего теста CRT_{sc} , возможны два варианта, а именно – эти наборы порождаются совпадающими или разными наборами шаблона $CRT_{ou(k-1)}$. Например, наборы T_0, T_4, T_8 и T_{12} , приведенные в табл. 3, сформированы на базе одного совпадающего набора 0 0 0 для всех трех примеров масштабирования. Для подобных случаев $h = HD(T_i, T_j)$ вычисляется как $0 \cdot (r_s - h_s) + r_{o(k-1)} \cdot h_s = r_{o(k-1)} \cdot h_s$, где h_s представляет собой расстояние Хэмминга для двух различных строк масштабирующего теста CRT_{sc} . Первая из этих строк идентифицирует блок, к которому принадлежит T_i , а вторая определяет блок, включающий набор T_j , который порождается набором T_i . В приведенном выражении $h = 0 \cdot (r_s - h_s) + r_{o(k-1)} \cdot h_s$ первое слагаемое определяется повторяющимися наборами, например 0 0 0, имеющими расстояние Хэмминга, равное 0. Второе слагаемое соответствует случаю, когда расстояние Хэмминга равняется $r_{o(k-1)}$ как результат вычисления этого значения для инверсных наборов, например 0 0 0 и 1 1 1 (1).

Во втором случае тестовые наборы T_i и T_j порождаются разными наборами шаблона $CRT_{ou(k-1)}$, например T_0, T_5 . Для таких наборов применима зависимость расстояния Хэмминга $HD(T_i, T_j) + HD(T_i, \bar{T}_j) = r$, в которой фигурируют значения $h_i = HD(T_i, T_j)$ для наборов шаблона $CRT_{ou(k-1)}$ (1). Тогда, подобно первому случаю, расстояние Хэмминга будет вычисляться как $h = h_i \cdot (r_s - h_s) + h_s \cdot (r_{o(k-1)} - h_i) = h_i \cdot r_s + r_{o(k-1)} \cdot h_s - 2 \cdot h_i \cdot h_s$. В свою

очередь, минимальное значение расстояния Хэмминга h между наборами будет определяться величинами максимальных расстояний b_t и b_s для h_t и h_s либо минимальных их значений d_t и d_s . Это следует из того факта, что после преобразования выражения $h = h_t(r_s - h_s) + h_s(r_{o(k-1)} - h_t)$ к виду $h = h_t(r_s - 2 \cdot h_s) + h_s \cdot r_{o(k-1)}$ минимальное и максимальное значения величины h зависят от соотношения величин r_s и h_s . Данные величины определяют знак первого слагаемого в соответствии с выражением $(r_s - 2 \cdot h_s)$. Отметим, что все величины, участвующие в определении значения h , являются положительными. Таким образом, для определения предельных значений $\min h$ и $\max h$ важными являются соотношения $h_s = r_s/2$, $h_s > r_s/2$ и $h_s < r_s/2$. При выполнении равенства $h_s = r_s/2$ имеем $\min h = \max h = h_s \cdot r_{o(k-1)}$. В случае выполнения неравенства $h_s > r_s/2$ имеем знак минус в первом слагаемом для выражения h . Соответственно, $\max h$ достигается для $d_{o(k-1)}$ в заданном шаблоне $CRT_{ou(k-1)}$, а $\min h$ достигается для $b_{o(k-1)}$. Наконец, для соотношения $h_s < r_s/2$ зависимость обратная, т. е. $\max h$ получается для $b_{o(k-1)}$, а $\min h$ для $d_{o(k-1)}$.

Таким образом, для $h_s \geq r_s/2$ значение $d_{o(k)} = \min\{(r_s \cdot d_{o(k-1)}), (r_{o(k-1)} \cdot d_s), (r_s \cdot b_{o(k-1)} + r_{o(k-1)} \cdot b_s - 2 \cdot (b_s \cdot b_{o(k-1)}))\}$, $b_{o(k)} = \max\{(r_s \cdot b_{o(k-1)}), (r_{o(k-1)} \cdot b_s), (r_s \cdot d_{o(k-1)} + r_{o(k-1)} \cdot d_s - 2 \cdot (d_s \cdot d_{o(k-1)}))\}$, а для $h_s < r_s/2$ — $d_{o(k)} = \min\{(r_s \cdot d_{o(k-1)}), (r_{o(k-1)} \cdot d_s), (r_s \cdot d_{o(k-1)} + r_{o(k-1)} \cdot d_s - 2 \cdot (d_s \cdot d_{o(k-1)}))\}$, $b_{o(k)} = \max\{(r_s \cdot b_{o(k-1)}), (r_{o(k-1)} \cdot b_s), (r_s \cdot b_{o(k-1)} + r_{o(k-1)} \cdot b_s - 2 \cdot (b_s \cdot b_{o(k-1)}))\}$. Что и требовалось доказать.

Утверждение 3 позволяет вычислять характеристики формируемых тестов. Например, для $CRT_{sc}(q_s, d_s, b_s, r_s) = CRT(4, 2, 2, 3)$ и $CRT_{ou(k-1)}(q_{o(k-1)}, d_{o(k-1)}, b_{o(k-1)}, r_{o(k-1)}) = CRT(4, 1, 3, 3)$ имеем $q_{o(k)} = q_s \cdot q_{o(k-1)} = 4 \cdot 4 = 16$; $d_{o(k)} = \min\{(r_s \cdot d_{o(k-1)}), (r_{o(k-1)} \cdot d_s), (r_s \cdot b_{o(k-1)} + r_{o(k-1)} \cdot b_s - 2 \cdot (b_s \cdot b_{o(k-1)}))\} = \min\{(3 \cdot 1), (3 \cdot 2), (3 \cdot 3 + 3 \cdot 2 - 2 \cdot (2 \cdot 3))\} = 3$; $b_{o(k)} = \max\{(r_s \cdot b_{o(k-1)}), (r_{o(k-1)} \cdot b_s), (r_s \cdot d_{o(k-1)} + r_{o(k-1)} \cdot d_s - 2 \cdot (d_s \cdot d_{o(k-1)}))\} = \max\{(3 \cdot 3), (3 \cdot 2), (3 \cdot 1 + 3 \cdot 2 - 2 \cdot (2 \cdot 1))\} = 9$; $r_{o(k)} = r_s \cdot r_{o(k-1)} = 3 \cdot 3 = 9$. В результате формируется тест $CRT_{ou}(q_o, d_o, b_o, r_o) = CRT(16, 3, 9, 9)$, представленный в табл. 3 и 5. В этих же таблицах приведен тест $CRT_{ou}(q_o, d_o, b_o, r_o) = CRT_{ou}(16, 4, 6, 9)$, полученный как результат масштабирования теста $CRT(4, 2, 2, 3)$ на основании такого же масштабирующего теста.

Цель итерационной процедуры двухмерного масштабирования заключается в формировании управляемого вероятностного теста с требуемыми характеристиками, которые определяются в соответствии с утверждением 3. В первую очередь важным является обеспечение требуемой разрядности m результирующего теста, которая определяет количество k итераций предложенной процедуры. Соотношения для $r_{o(k)} = r_s \cdot r_{o(k-1)}$, приведенные ниже, для фиксированного масштабирующего теста $CRT_{sc}(q_s, d_s, b_s, r_s)$, который используется для получения $CRT_{ou(k)}(q_{o(k)}, d_{o(k)}, b_{o(k)}, r_{o(k)})$ на основании $CRT_{tm}(q_t, d_t, b_t, r_t)$ или $CRT_{tm}(1, 0, 0, r_t)$, позволяют определить величину k . В обоих случаях $r_{o(k)} = r_s^k \cdot r_t$, тогда k может быть определено следующим образом:

$$k \geq \left\lceil \log_{r_s} (m / r_t) \right\rceil. \quad (2)$$

Полученная величина k , обеспечивающая требуемую разрядность m , позволяет вычислять количество тестовых наборов в результирующем тесте $CRT_{ou(k)}$. Для $CRT_{tm}(q_t, d_t, b_t, r_t)$ имеем $q_{o(k)} = q_s^k \cdot q_t$, а для $CRT_{tm}(1, 0, 0, r_t)$, соответственно, $q_{o(k)} = q_s^k$.

Практически важным является случай, когда разрядность сформированного на основании итерационной процедуры теста кратна 2^k , где $k \in \{1, 2, 3, \dots\}$, а пороговое зна-

чение $d_{o(k)}$ расстояния Хэмминга принимает приемлемые величины. Для построения подобных тестов может быть применено одно из следствий, вытекающих из утверждения 3.

Следствие. Результатом k -й итерации двумерного масштабирования с помощью масштабирующего теста $CRT_{sc}(q_s, d_s, b_s, r_s) = CRT_{sc}(2, 1, 1, 2)$ для $CRT_{ou_0}(r, r/2, r/2, r)$ и четного r является тест $CRT_{ou_k}(2^k \cdot r, 2^{k-1} \cdot r, 2^{k-1} \cdot r, 2^k \cdot r)$.

Доказательство. Согласно утверждению 3 для $k = 1$ имеем $CRT_{ou_1}(2 \cdot r, d_{o(1)}, b_{o(1)}, 2 \cdot r)$, где значение $d_{o(1)} = \min\{(2 \cdot r/2), (r-1), (2 \cdot r/2 + r-1 - 2 \cdot (1 \cdot r/2))\} = r$. Аналогично и $b_{o(1)} = r$. Соответственно, $CRT_{ou_1}(2 \cdot r, d_{o(1)}, b_{o(1)}, 2 \cdot r) = CRT_{ou_1}(2 \cdot r, r, r, 2 \cdot r)$. Далее на основании $CRT_{ou_1}(2 \cdot r, r, r, 2 \cdot r)$ ($k = 2$) формируется $CRT_{ou_2}(2^2 \cdot r, 2^1 \cdot r, 2^1 \cdot r, 2^2 \cdot r)$. Для произвольного k имеем $CRT_{ou_k}(2^k \cdot r, 2^{k-1} \cdot r, 2^{k-1} \cdot r, 2^k \cdot r)$, что и требовалось доказать.

Для случая $CRT_{ou_0}(2, 1, 1, 2) = \{T_0 = 1\ 1, T_1 = 0\ 1\}$ в табл. 6 приведен пример трех итераций для $CRT_{sc}(2, 1, 1, 2) = \{T_0 = 0\ 0, T_1 = 0\ 1\}$.

Таблица 6

Итерационная процедура построения управляемых вероятностных тестов для $CRT_{sc}(2, 1, 1, 2)$

Table 6

An iterative procedure for constructing controlled random tests for $CRT_{sc}(2, 1, 1, 2)$

CRT_{sc}	$CRT_{ou(1)}$	CRT_{sc}	$CRT_{ou(2)}$	CRT_{sc}	$CRT_{ou(3)}$
$T_0 = 0\ 0$	11 11 01 01	$T_0 = 0\ 0$	11 11 11 11 01 01 01 01 11 00 11 00 01 10 01 10	$T_0 = 0\ 0$	11 11 11 11 11 11 11 11
					01 01 01 01 01 01 01 01
					11 00 11 00 11 00 11 00
					01 10 01 10 01 10 01 10
					11 11 00 00 11 11 00 00
					01 01 10 10 01 01 10 10
$T_1 = 0\ 1$	11 00 01 10	$T_1 = 0\ 1$	11 11 00 00 01 01 10 10 11 00 00 11 01 10 10 01	$T_1 = 0\ 1$	11 11 11 11 00 00 00 00
					01 01 01 01 10 10 10 10
					11 00 11 00 00 11 00 11
					01 10 01 10 10 01 10 01
					11 11 00 00 00 00 11 11
					01 01 10 10 10 10 01 01
					11 00 00 11 00 11 11 00
					01 10 10 01 10 01 01 10

Из табл. 6 видно, что пороговое расстояние Хэмминга для всех трех тестов, построенных в результате итерационной процедуры, соответствует следствию. Действительно, $d_{o(1)} = 2$, $d_{o(2)} = 4$ и $d_{o(3)} = 8$.

3. Результаты экспериментальных исследований

Для подтверждения полученных авторами результатов был проведен ряд вычислительных и практических экспериментов. Исследовалась эффективность управляемых вероятностных тестов, полученных в результате масштабирования исходных шаблонов. Эксперимент состоял в определении эффективности предлагаемой процедуры построения управляемых вероятностных тестов $CRT_{ou}(q_o, d_o, b_o, r_o)$ на основании $CRT_{sc}(q_s, d_s,$

$b_s, r_s) = CRT_{sc}(4, 2, 2, 3)$ и $CRT_{tm}(1, 0, 0, r_t) = CRT_{tm}(1, 0, 0, 341)$. Выбор указанного CRT_{sc} обоснован его характеристиками, а размерность $r_t = 341$ – значением $r_o = r_s \cdot r_t = 3 \cdot 341 = 1023$, определяемым емкостью запоминающего устройства для последующих экспериментов [32, 33]. Выполнялась одна итерация предложенной процедуры, в результате которой формировался результирующий тест $CRT_{ou}(4, 682, 682, 1023)$ как результат масштабирования шаблона $CRT_{tm}(1, 0, 0, 341)$.

Эффективность тестов оценивалась по разнообразию (отличию) тестовых наборов, количественно определяемому с помощью расстояний Хэмминга между наборами в каждом тесте $CRT_{ou}(4, 682, 682, 1023)$. Для сравнения таких оценок использовались аналогичные оценки, полученные для вероятностных тестов, также включающих четыре случайных тестовых набора, которые служили статистической базой [31]. Большие значения расстояния Хэмминга указывают на более сильное взаимное различие тестовых наборов, что необходимо с точки зрения покрытия пространства тестовых наборов [32, 33]. Каждый тест состоял из четырех двоичных наборов длиной 1023 бита, для каждого из таких тестов определялись все шесть значений расстояния Хэмминга. Основными статистическими мерами, используемыми при анализе, были среднее расстояние Хэмминга $aveHD(T_i, T_j)$, минимальное расстояние Хэмминга $d = \min HD(T_i, T_j)$ и максимальное расстояние Хэмминга $b = \max HD(T_i, T_j)$, получаемые для каждого теста.

Чтобы экспериментально оценить статистическую стабильность указанных параметров и определить, сколько вероятностных тестов необходимо для получения репрезентативных средних значений, было проанализировано стандартное отклонение расстояний Хэмминга. При $r_o = 1023$ среднее расстояние Хэмминга $aveHD(T_i, T_j)$ между двумя случайными тестовыми наборами, являющееся также случайной величиной, можно описать нормальным законом распределения с математическим ожиданием $\mu = r_o/2$ и дисперсией $\sigma^2 = r_o/4$. С помощью этой модели дисперсии и нормального приближения количество независимых случайных тестов N , необходимое для достижения 95%-го доверительного интервала с максимальной погрешностью ± 5 бит, было оценено как $N \approx 7$. Аналогичный анализ был применен к случайным величинам d и b . На основании данного анализа количество случайных тестов, использованных в эксперименте, было установлено равным 15. Это обеспечивает статистическую точность лучше ± 5 бит для всех анализируемых параметров, а именно $aveHD(T_i, T_j)$, $d = \min HD(T_i, T_j)$ и $b = \max HD(T_i, T_j)$. Каждый из 15 тестов состоял из четырех независимых 1023-битных случайных наборов, и для каждого теста были определены все парные расстояния Хэмминга $HD(T_i, T_j)$. Полученные значения $aveHD(T_i, T_j)$, $d = \min HD(T_i, T_j)$ и $b = \max HD(T_i, T_j)$ затем усреднялись по тестам для оценки характеристик расстояний вероятностных тестов. Подробные результаты этого эксперимента представлены в табл. 7.

Значения $aveHD(T_i, T_j)$ округлены до целых величин. Как видно из табл. 7, экспериментальные значения d , $aveHD(T_i, T_j)$ и b для 15 вероятностных тестов лежат в диапазонах $470 \leq d \leq 511$, $498 \leq aveHD(T_i, T_j) \leq 526$ и $519 \leq b \leq 558$. Средние значения для вероятностных тестов составляют около 512,76 для $aveHD(T_i, T_j)$, а также 492,80 и 531,80 для d и b , тогда как управляемые вероятностные тесты $CRT_{ou}(4, 682, 682, 1023)$ имеют постоянное значение 682 во всех трех случаях.

Таблица 7

Экспериментальные статистические результаты для 15 вероятностных тестов

Table 7

Experimental statistical results for 15 random tests

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
$d = \min HD(T_i, T_j)$	504	502	470	475	489	476	500	511	490	484	497	501	501	490	502
$ave HD(T_i, T_j)$	508	516	498	506	514	501	517	519	507	511	509	512	526	518	522
$b = \max HD(T_i, T_j)$	519	541	520	539	535	529	527	537	522	525	527	520	558	533	545

Подобные управляемые вероятностные тесты $CRT_{ou}(q_o, d_o, b_o, r_o)$, построенные в результате масштабирования шаблона $CRT_m(1, 0, 0, 341)$, описываются как $CRT_{ou}(q_s, d_s \cdot r_t, b_s \cdot r_t, r_s \cdot r_t) = CRT_{ou}(4, 682, 682, 1023)$, что следует из утверждения 3.

Обобщенные результаты, представленные в табл. 7, наглядно показывают разницу в распределении расстояний Хэмминга между вероятностными и управляемыми вероятностными тестами. В то время как вероятностные тесты демонстрируют заметную вариабельность между минимальным, максимальным и средним расстояниями, управляемые вероятностные тесты CRT_{ou} сохраняют стабильные более высокие значения этих характеристик. Для любого теста CRT_{ou} , полученного, например, на основании шаблона $CRT_m(1, 0, 0, r_t)$, который использует $r_t = 341$ -разрядный случайный набор, все указанные характеристики принимают значение 682. Более того, указанные значения расстояний Хэмминга значительно выше по сравнению с аналогичными значениями вероятностных тестов. Это подтверждает повышение эффективности предлагаемого метода для формирования более разнообразных и хорошо разделенных тестовых наборов, что крайне желательно с точки зрения покрытия тестового пространства [32, 33].

Заключение

Рассмотрен подход к генерированию тестовых наборов при формировании управляемых вероятностных тестов с помощью одномерного и двухмерного масштабирования исходных шаблонов. Предложен итерационный подход для синтеза управляемых вероятностных тестов. Основой предложенного авторами подхода является двухмерное масштабирование исходных шаблонов с применением масштабирующих матриц, которыми, так же как и шаблоны, могут быть управляемые вероятностные тесты с небольшим числом наборов и малой их разрядностью. Показано, что использование различных шаблонов и масштабирующих тестов для двухмерного масштабирования позволяет строить управляемые вероятностные тесты с требуемой разрядностью тестовых наборов и бóльшим их количеством. Управляемые вероятностные тесты формируются без необходимости перечисления кандидатов в тестовые наборы, что сводит задачу синтеза управляемого вероятностного теста к формальной процедуре, которая исключает необходимость вычисления меры (мер) различия и не требует вычислительных затрат. В работе показано, что количество наборов результирующего теста и их свойства зависят от основных характеристик шаблона и масштабирующего теста. Приведены основные соотношения для определения этих характеристик, и экспериментально показаны преимущества синтезируемых тестов по отношению к вероятностным тестам.

В рамках предложенного авторами подхода важным является дальнейшее исследование в части нахождения оптимального сочетания шаблона с масштабирующим тестом, что во многом зависит от целевых параметров синтезируемого теста, которыми могут быть не только разрядность тестовых наборов и их количество. Интересным представляется дальнейшее исследование практической применимости тестов, синтезированных по предложенному авторами методу, для тестирования современного программного обеспечения и запоминающих устройств.

Вклад авторов. В. Н. Ярмолик предложил итерационную процедуру построения управляемых вероятностных тестов, основанную на применении двухмерного масштабирования исходных шаблонов. И. Мрозек провел большой объем экспериментальных исследований, принял участие в обобщении, анализе и оформлении полученных результатов. П. Ю. Бранцевич принял участие в обобщении, анализе, редактировании и оформлении полученных результатов.

Список использованных источников

1. Ledin, J. *Modern Computer Architecture and Organization* / J. Ledin. – Birmingham : Packt Publishing Ltd., 2020. – 536 p.
2. Karmore, S. P. Testing of embedded system, an issues and challenges / S. P. Karmore, A. R. Mahajan // *International Journal of Enhanced Research in Science, Technology & Engineering*. – 2015. – Vol. 4, no. 8. – P. 181–186.
3. Ярмолик, В. Н. *Контроль и диагностика вычислительных систем* / В. Н. Ярмолик. – Мн. : Бестпринт, 2019. – 387 с.
4. Krupp, A. A systematic approach to the test of combined HW/SW systems / A. Krupp, W. Muller // *Proc. of IEEE Conf. on the Testing and Automation of Embedded Systems (DATE 2010)*, Dresden, Germany, 08–12 Mar. 2010. – Dresden, 2010. – P. 323–326.
5. Teller-Giron, R. Random fault detection in logical networks / R. Teller-Giron, R. David // *Proc. of Intern. Symp. on Discrete Systems*, Riga, USSR, 30 Sept. – 4 Oct. 1974. – Riga, 1974. – P. 232–241.
6. Agrawal, V. D. When to use random testing / V. D. Agrawal // *IEEE Transactions on Computers*. – 1978. – Vol. C-27, no 11. – P. 1054–1055.
7. Bernet, G. A theory of probabilistic functional testing / G. Bernet, L. Bouaziz, P. LeGall // *Proc. of the 1997 Intern. Conf. on Software Engineering*, Boston, Massachusetts, USA, 17–23 May 1997. – Boston, 1997. – P. 216–226.
8. Arcuri, A. Random testing: Theoretical results and practical implications / A. Arcuri, M. Z. Iqbal, L. Briand // *IEEE Transactions on Software Engineering*. – 2011. – Vol. 38, no. 2. – P. 258–277.
9. Bushnell, M. *Essentials of Electronic Testing for Digital, Memory and Mixed-Signal VLSI Circuits (Frontiers in Electronic Testing)* / M. Bushnell, V. Agrawal. – Dordrecht, Netherlands : Springer, 2004. – 690 p.
10. Yarmolik, V. N. *Self-Testing VLSI Design* / V. N. Yarmolik, I. V. Kachan. – Amsterdam : Elsevier Science Publishers, 1993. – 345 p.
11. An orchestrated survey on automated software test case generation / S. Anand, E. K. Burke, T. Y. Chen [et al.] // *Journal of Systems and Software*. – 2014. – Vol. C-39, no. 4. – P. 582–586.
12. Myers, G. J. *The Art of Software Testing* / G. J. Myers, C. Sandler, T. Badgett. – 3rd ed. – Canada : John Wiley & Sons, Inc., 2012. – 240 p.
13. Testing embedded software: A survey of the literature / V. Garousi, M. Felderer, C. M. Karapıçak, U. Yılmaz // *Information and Software Technology*. – 2018. – Vol. 104. – P. 14–45.

14. Goor, A. J. *Testing Semiconductor Memories, Theory and Practice* / A. J. Goor. – Chichester, UK : John Wiley & Sons, 1991. – 536 p.
15. A survey on adaptive random testing / R. Huang, W. Sun, Y. Xu [et al.] // *IEEE Transactions on Software Engineering*. – 2021. – Vol. 47, no. 10. – P. 2052–2083.
16. Adaptive random testing: The art of test case diversity / T. Y. Chen, F. C. Kuo, R. G. Merkel, T. H. Tse // *Journal of Systems and Software*. – 2010. – Vol. 83. – P. 60–66.
17. Alamgir, A. Adaptive random testing with total Cartesian distance for black box circuit under test / A. Alamgir // *Indonesian Journal of Electrical Engineering and Computer Science*. – 2020. – Vol. 20, no. 2. – P. 720–726.
18. Antirandom testing: A distance-based approach / S. N. Wu, S. Jandhyala, Y. K. Malaiya, A. P. Jayasumana // *Hindawi Publishing Corporation VLSI Design*. – 2008. – Vol. 2008, art. ID 165709. – 9 p. – <https://doi.org/10.1155/2008/165709>.
19. Xu, S. Maximum distance testing / S. Xu, J. Chen // *Proc. of the 11th IEEE Asian Test Symp. (ATS'02), Guam, USA, 18–20 Nov. 2002*. – Guam, 2002. – P. 15–20.
20. Xu, S. Orderly random testing for hardware and software / S. Xu // *Proc. of the 2008 14th IEEE Pacific Rim Intern. Symp. on Dependable Computing, Washington, DC, USA, 15–17 Dec. 2008*. – Washington, 2008. – P. 160–167.
21. Mrozek, I. Multiple controlled random testing / I. Mrozek, V. N. Yarmolik // *Fundamenta Informaticae*. – 2016. – Vol. 144, no. 1. – P. 23–43.
22. Yarmolik, S. V. The synthesis of probability tests with a small number of kits / S. V. Yarmolik, V. N. Yarmolik // *Automatic Control and Computer Sciences*. – 2011. – Vol. 45, no. 3. – P. 133–141.
23. Ярмолик, С. В. Управляемое случайное тестирование / С. В. Ярмолик, В. Н. Ярмолик // *Информатика*. – 2011. – № 1(29). – С. 79–88.
24. Yarmolik, S. V. Controlled Random Tests / S. V. Yarmolik, V. N. Yarmolik // *Automation and Remote Control*. – 2012. – Vol. 73, no. 10. – P. 1704–1714.
25. Hamming, R. W. Error Detecting and Error Correcting Codes / R. W. Hamming // *The Bell System Technical Journal*. – 1950. – Vol. 29, no. 2. – P. 147–160.
26. Peterson, W. W. *Error-Correction Codes* / W. W. Peterson, E. J. Weldon. – Cambridge, Massachusetts, London, England : The MIT Press, 1972. – 560 p.
27. Plotkin, M. Binary codes with specified minimum distance / M. Plotkin // *IRE Transactions on Information Theory*. – 1960. – Vol. 6, no. 4. – P. 445–450.
28. MacWilliams, F. J. *The Theory of Error-Correcting Codes* / F. J. MacWilliams, N. J. A. Sloane. – Amsterdam, The Netherlands : Elsevier-North-Holland Publishing Co., 1977. – 762 p.
29. Ярмолик, В. Н. Масштабирование управляемых вероятностных тестов с применением матриц Адамара / В. Н. Ярмолик, Н. А. Шевченко, В. В. Петровская // *Информатика*. – 2025. – Т. 22, № 2. – С. 63–80.
30. Метод построения управляемых вероятностных тестов / В. Н. Ярмолик, И. Мрозек, П. Ю. Бранцевич [и др.] // *Доклады БГУИР*. – 2025. – Т. 23, № 6. – С. 87–95.
31. Hahn, G. J. *Statistical Models in Engineering* / G. J. Hahn, S. S. Shapiro. – N. Y., USA : John Wiley & Sons, 1994. – 376 p.
32. Ярмолик, В. Н. Многократные неразрушающие маршевые тесты с изменяемыми адресными последовательностями / В. Н. Ярмолик, С. В. Ярмолик // *Автоматика и телемеханика*. – 2007. – Вып. 4. – С. 126–137.
33. Mrozek, I. Problemy funkcjonalnego testowania pamięci RAM / I. Mrozek, V. Yarmolik. – Białystok, Polska : Politechnika Piałostocka, 2009. – 264 p.

References

1. Ledin J. *Modern Computer Architecture and Organization*. Birmingham, Packt Publishing Ltd., 2020, 536 p.
2. Karmore S. P., Mahajan A. R. Testing of embedded system, an issues and challenges. *International Journal of Enhanced Research in Science, Technology & Engineering*, 2015, vol. 4, no. 8, pp. 181–186.
3. Yarmolik V. N. Control' i diagnostika vuchislitel'nuch system. *Computer Systems Testing and Diagnoses*. Minsk, Bestprint, 2019, 387 p. (In Russ.).
4. Krupp A., Muller W. A systematic approach to the test of combined HW/SW systems. *Proceedings of the IEEE Conference on the Testing and Automation of Embedded Systems (DATE 2010), Dresden, Germany, 08–12 March 2010*. Dresden, 2010, pp. 323–326.
5. Teller-Giron R., David R. Random fault detection in logical networks. *Proceedings of the International Symposium on Discrete Systems, Riga, USSR, 30 September – 4 October 1974*. Riga, 1974, pp. 232–241.
6. Agrawal V. D. When to use random testing. *IEEE Transactions on Computers*, 1978, vol. C-27, no. 11, pp. 1054–1055.
7. Bernet G., Bouaziz L., LeGall P. A theory of probabilistic functional testing. *Proceedings of the 1997 International Conference on Software Engineering, Boston, Massachusetts, USA, 17–23 May 1997*. Boston, 1997, pp. 216–226.
8. Arcuri A., Iqbal Z., Briand L. Random testing: Theoretical results and practical implications. *IEEE Transactions on Software Engineering*, 2011, vol. 38, no. 2, pp. 258–277.
9. Bushnell M., Agrawal V. *Essentials of Electronic Testing for Digital, Memory and Mixed-Signal VLSI Circuits (Frontiers in Electronic Testing)*. Dordrecht, Netherlands, Springer, 2004, 690 p.
10. Yarmolik V. N., Kachan I. V. *Self-Testing VLSI Design*. Amsterdam, Elsevier Science Publishers, 1993, 345 p.
11. Anand S., Burke E. K., Chen T. Y., Clark J., Cohen M. B., ..., Zhu H. An orchestrated survey on automate software test case generation. *Journal of Systems and Software*, 2014, vol. C-39, no. 4, pp. 582–586.
12. Myers G. J., Sandler C., Badgett T. *The Art of Software Testing 3rd Edition*. Canada, John Wiley & Sons Inc., 2012, 240 p.
13. Garousi V., Felderer M., Karapıçak C. M., Yılmaz U. Testing embedded software: A survey of the literature. *Information and Software Technology*, 2018, vol. 104, pp. 14–45.
14. Goor A. J. *Testing Semiconductor Memories, Theory and Practice*. Chichester, UK, John Wiley & Sons Inc., 1991, 536 p.
15. Huang R., Sun W., Xu Y., Chen H., Towey D., Xia X. A survey on adaptive random testing. *IEEE Transactions on Software Engineering*, 2021, vol. 47, no. 10, pp. 2052–2083.
16. Chen T. Y., Kuo F. C., Merkel R. G., Tse T. H. Adaptive random testing: The art of test case diversity. *Journal of Systems and Software*, 2010, vol. 83, pp. 60–66.
17. Alamgir A. Adaptive random testing with total Cartesian distance for black box circuit under test. *Indonesian Journal of Electrical Engineering and Computer Science*, 2020, vol. 20, no. 2, pp. 720–726.
18. Wu S. H., Jandhyala S., Malaiya Y. K., Jayasumana A. P. Antirandom testing: A distance-based approach. *Hindawi Publishing Corporation VLSI Design*, 2008, vol. 2008, art. ID 165709, 9 p. <https://doi.org/10.1155/2008/165709>.
19. Xu S., Chen J. Maximum distance testing. *Proceedings of the 11th Asian Test Symposium (ATS'02), Guam, USA, 18–20 November 2002*. Guam, 2002, pp. 15–20.

20. Xu S. Orderly random testing for both hardware and software. *Proceedings of the 2008 14th IEEE Pacific Rim International Symposium on Dependable Computing, Washington, DC, USA, 15–17 December 2008*. Washington, 2008, pp. 160–167.
21. Mrozek I., Yarmolik V. N. Multiple Controlled Random Testing. *Fundamenta Informaticae*, 2016, vol. 144, no. 1, pp. 23–43.
22. Yarmolik S. V., Yarmolik V. N. The synthesis of probability tests with a small number of kits. *Automatic Control and Computer Sciences*, 2011, vol. 45, no. 3, pp. 133–141.
23. Yarmolik S. V., Yarmolik V. N. *Controlled random testing*. *Informatika [Informatics]*, 2011, no. 1(29), pp. 79–88 (In Russ.).
24. Yarmolik S. V., Yarmolik V. N. Controlled random tests. *Automation and Remote Control*, 2012, vol. 73, no. 10, pp. 1704–1714.
25. Hamming R. W. Error detecting and error correcting codes. *The Bell System Technical Journal*, 1950, vol. 29, no. 2, pp. 147–160.
26. Peterson W. W., Weldon E. J. *Error-Correction Codes*. Cambridge, Massachusetts, London, England, The MIT Press, 1972, 560 p.
27. Plotkin M. Binary codes with specified minimum distance. *IRE Transactions on Information Theory*, 1960, vol. 6, no. 4, pp. 445–450.
28. MacWilliams F. J., Sloane N. J. A. *The Theory of Error-Correcting Codes*. Amsterdam, The Netherland, Elsevier-North-Holland Publishing Co., 1977, 762 p.
29. Yarmolik V. N., Shauchenka M. A., Petrovskaya V. V. *Scaling controlled random tests based on Hadamard matrices*. *Informatika [Informatics]*, 2025, vol. 22, no. 2, pp. 63–80 (In Russ.).
30. Yarmolik V. N., Mrozek I., Brancevich P. Yu., Demenkovets D. V., Levantsevich V. A. *Method of controlled random tests generation*. *Doklady BGUIR [BSUIR Proceedings]*, 2025, vol. 23, no. 6, pp. 87–95 (In Russ.).
31. Hahn G. J., Shapiro S. S. *Statistical Models in Engineering*. New York, USA, John Wiley & Sons, 1994, 376 p.
32. Yarmolik V. N., Yarmolik S. V. Multiple non-destructive marching tests with variable address sequences. *Automation and Remote Control*, 2007, vol. 4, pp. 126–137.
33. Mrozek I., Yarmolik V. N. *Problemy funkcjonalnego testowania pamięci RAM*. Białystok, Polska, Politechnika Piałostocka, 2009, 264 p.

Информация об авторах

Ярмолик Вячеслав Николаевич, доктор технических наук, профессор, Белорусский государственный университет информатики и радиоэлектроники.
E-mail: yarmolik10ru@yahoo.com

Мрозек Иренеуш, доктор, профессор, Белостокский технический университет.
E-mail: i.mrozek@pb.edu.pl

Бранцевич Петр Юльевич, доктор технических наук, профессор, Белорусский государственный университет информатики и радиоэлектроники.
E-mail: branc@bsuir.edu.by

Information about the authors

Vyacheslav N. Yarmolik, Dr. Sci. (Eng.), Prof., Belarusian State University of Informatics and Radioelectronics.
E-mail: yarmolik10ru@yahoo.com

Ireneusz Mrozek, Dr., Prof., Bialystok University of Technology.
E-mail: i.mrozek@pb.edu.pl

Peter Yu. Brancevich, Dr. Sci. (Eng.), Prof., Belarusian State University of Informatics and Radioelectronics.
E-mail: branc@bsuir.edu.by

ИНТЕЛЛЕКТУАЛЬНЫЕ СИСТЕМЫ

INTELLIGENT SYSTEMS

УДК 004.8
<https://doi.org/10.37661/1816-0301-2026-23-1-26-38>

Поступила в редакцию | Received 11.02.2026
Подписана в печать | Accepted 03.03.2026
Опубликована | Published 31.03.2026

BellLitGPT – технологии языковых моделей для белорусского языка

Д. А. Ляхов, А. М. Бондоловский[✉], С. В. Кругликов, К. К. Шульган
[✉]E-mail: a.bandalouski@newman.bas-net.by

*Объединенный институт проблем информатики
Национальной академии наук Беларуси,
Сурганова, 6, Минск, 220012, Беларусь*

Аннотация

Цели. Работа выполнена в области исследования специализированных генеративных нейронных сетей для белорусского языка. Поставлена цель сделать первый шаг для построения национальной генеративной языковой модели.

Методы. Описывается процесс разработки модели BellLitGPT (700 млн параметров), который основан на стратегии трансферного обучения русскоязычной модели ruGPT-3 и состоит из трех этапов: подготовки корпуса, адаптации токенизатора и обучения модели. Обучающий корпус составлен из золотого фонда классической белорусской прозы и подготовленных статей из Википедии. Подробно описываются методика адаптации токенизатора для расширения словарного запаса специфическими белорусскими лексемами, процесс обучения и тестирования модели.

Результаты. Результаты исследования подтверждают способность модели BellLitGPT генерировать связанные, грамматически и стилистически корректные тексты. Особое внимание уделено созданию гибридного нейросимвольного подхода для генерации четверостиший с соблюдением ритма и рифмы.

Заключение. Эксперимент по масштабированию архитектуры показал сложности в обучении крупной модели (13 млрд параметров) в условиях дефицита данных.

Ключевые слова: большие языковые модели, трансферное обучение, нейросимвольный подход, генерация стихов, модель BellLitGPT, белорусский язык

Благодарности. Авторы выражают благодарность следующим сотрудникам ОИПИ НАН Беларуси: старшему научному сотруднику лаборатории распознавания и синтеза речи Н. В. Супрунчуку за проверку текстового корпуса, младшему научному сотруднику названной лаборатории А. Е. Дрогун за помощь в сборе и подготовке данных, заместителю генерального директора по научной работе С. Н. Касанину за успешные переговоры по получению в безвозмездное пользование вычислительного устройства AMD Ryzen AI Max+ PRO 395 на время проведения эксперимента и заведующему лабораторией высокопроизводительных систем О. П. Чижу за содействие в настройке устройства.

Для цитирования. BellLitGPT – технологии языковых моделей для белорусского языка / Д. А. Ляхов, А. М. Бондоловский, С. В. Кругликов, К. К. Шульган // Информатика. – 2026. – Т. 23, № 1. – С. 26–38. – <https://doi.org/10.37661/1816-0301-2026-23-1-26-38>.

Конфликт интересов. А. М. Бондоловский является членом редакционной коллегии журнала «Информатика» с 2026 г., но не имеет никакого отношения к решению опубликовать эту статью. С. В. Кругликов является членом редакционной коллегии журнала «Информатика» с 2016 г., но не имеет никакого отношения к решению опубликовать эту статью. Статья прошла принятую в журнале процедуру рецензирования. Об иных конфликтах интересов авторы не заявляли.



BelLitGPT – language model technologies for the Belarusian language

Dmitry A. Lyakhov, Andrei M. Bandalouski[✉], Sergey V. Kruglikov, Konstantin K. Shulgan

[✉]E-mail: a.bandalouski@newman.bas-net.by

*The United Institute of Informatics Problems
of the National Academy of Sciences of Belarus,
st. Surganova, 6, Minsk, 220012, Belarus*

Abstract

Objectives. The research is conducted in the field of specialized generative neural networks for the Belarusian language. The authors aim to take the first step towards building a national generative language model.

Methods. The paper describes the development process of the BelLitGPT model (700 million parameters). It is based on a transfer learning strategy using the Russian-language model ruGPT-3 and consists of three stages: corpus preparation, tokenizer adaptation methodology and model training. The training corpus is compiled from the golden fund of classic Belarusian prose and prepared Wikipedia articles. The paper details the tokenizer adaptation method for expanding the vocabulary with specific Belarusian lexemes, as well as the model training and testing process.

Results. The research results confirm that BelLitGPT can generate coherent, grammatically and stylistically correct texts. Special attention is given to the creation of a hybrid neuro-symbolic approach for generating quatrains that adhere to rhythm and rhyme.

Conclusion. The experiment on scaling the architecture revealed difficulties in training a large model (13 billion parameters) under conditions of data scarcity.

Keywords: large language models (LLM), transfer learning, neuro-symbolic approach, poetry generation, model BelLitGPT, Belarusian language

Acknowledgments. The authors express their gratitude to the following coworkers of the UIIP NAS of Belarus: Mikita V. Suprunchuk, Senior Researcher at the Laboratory of Speech Synthesis and Recognition, for verifying the text corpus; Anastasia E. Drogun, Junior Researcher at the Laboratory of Speech Synthesis and Recognition, for assistance with data collection and preparation; Sergey N. Kasanin, Deputy General Director for Research, for successful negotiations regarding the free use of an AMD Ryzen AI Max+ PRO 395 computing device for the duration of the experiment; and Oleg P. Chizh, Head of the High-Performance Systems Laboratory, for assistance with device setup.

For citation. Lyakhov D. A., Bandalouski A. M., Kruglikov S. V., Shulgan K. K. *BelLitGPT – language model technologies for the Belarusian language*. *Informatika [Informatics]*, 2026, vol. 23, no. 1, pp. 26–38 (In Russ.). <https://doi.org/10.37661/1816-0301-2026-23-1-26-38>.

Conflict of interests. A. M. Bandalouski has been a member of the editorial board of the journal "Informatics" since 2026 but had no role in the decision to publish this article. S. V. Kruglikov has been a member of the editorial board of the journal "Informatics" since 2016 but had no role in the decision to publish this article. The article has undergone the journal's established peer-review process. The authors declare of no other conflicts of interest.

Введение

Последнее десятилетие ознаменовалось стремительным развитием области обработки естественного языка (англ. natural language processing, NLP), которое стало возможным благодаря появлению и широкому распространению больших языковых моделей (БояМ, англ. large language models, LLM) [1–3]. Эти модели обладают хорошими способностями в создании текстов, выполнении переводов и решении интеллектуальных задач. Тем не менее, несмотря на значительные международные достижения в данной

области, большинство современных решений остаются ориентированными преимущественно на английский язык и ограниченное число наиболее распространенных мировых языков. Для языков с небольшим объемом доступных цифровых материалов, включая белорусский, существующие технологии часто показывают недостаточно высокое качество генерации, склонны порождать артефакты (так называемые галлюцинации) и недостаточно хорошо понимают специфику культурного контекста. В связи с этим создание национальной большой языковой модели не только является технической задачей, но и становится важнейшей предпосылкой для поддержания культурной самобытности и укрепления цифрового суверенитета в изменяющемся мире. Научная область исследований специализированных генеративных нейронных сетей для белорусского языка новая и малоизученная. Можно сказать, что на сегодняшний день практически отсутствуют открыто обученные модели, которые показывают высокое владение белорусским языком.

Данная работа призвана преодолеть указанный недостаток и положить начало систематическим исследованиям в области создания национальной БояМ. Исследование рассматривается как первый шаг на пути к построению развитой цифровой экосистемы сервисов генеративного белорусского искусственного интеллекта (ИИ). Под «экосистемой белорусского ИИ» понимается система решений, которые выстроены на одной платформе вокруг мощной базовой языковой модели, обладающей знанием белорусского языка и национального контекста. К окружающим «ядро» (языковую модель) решениям могут относиться такие сервисы и надстройки, как поиск в интернете, управление подкасками при запросах, обеспечение безопасности содержания, диалоговые интерфейсы для конечных пользователей (чат-боты), API для интеграции в сторонние приложения, специализированные ИИ-ассистенты (для образования, анализа документов, юридической или медицинской сферы) и др.

В качестве отправной точки для экспериментов была выбрана стратегия трансферного обучения [4]. В статье описываются шаги ее осуществления, в частности процесс адаптации и дообучения модели на специально собранном корпусе белорусских текстов. Авторы предположили, что нативные русскоязычные модели наиболее легко и успешно переобучаются белорусскому языку, и выбрали предварительно обученную малую языковую модель ruGPT-3¹ с 700 млн параметров (далее – ruGPT-3). Благодаря заложенным знаниям в языковой модели с родственным языком выбранный подход позволяет эффективно начать разработку суверенной БояМ даже при ограниченных вычислительных ресурсах и данных.

1. Подготовка данных

Цель подготовительных работ – создание компактного набора данных, который подходит для переобучения современных языковых моделей, приведен к единому орфографическому стандарту и содержит в себе признанные социокультурные ценности белорусского народа. Далее опишем процесс создания текстового корпуса для переобучения выбранной модели.

¹SberDevices, Sber AI, and SberCloud. rugpt-3: Open source russian gpt-3 models. GitHub repository, 2020.

Корпус включил в себя произведения классической белорусской литературы и статьи из белорусской Википедии. Основной набор данных корпуса сформирован из золотого фонда классической белорусской прозы. В него включены труды 20 века следующих авторов:

Уладзімір Караткевіч	Міхась Лынькоў
Якуб Колас	Алесь Адамовіч
Іван Шамякін	Кузьма Чорны
Янка Брыль	Эліза Ажэшка
Кандрат Крапіва	Зьмітрок Бядуля
Васіль Быкаў	Цішка Гартны
Пятрусь Броўка	Максім Гарэцкі
Іван Навуменка	Алесь Савіцкі
Ніл Гілевіч	Цётка
Андрэй Макаёнак	Міхась Чарот
Іван Чыгрынаў	Ядвігін Ш.
Іван Мележ	Віктар Супрунчук

Весь массив текстов был выверен в соответствии с официальными правилами белорусской орфографии 1957 г. (утверждены постановлением Совета министров БССР)². Для этого вначале корпус прошел проверку правописания с помощью мультязычной БояМ, а затем был дополнительно утвержден кандидатом филологических наук, старшим научным сотрудником лаборатории распознавания и синтеза речи ОИПИ НАН Беларуси Н. В. Супрунчуком. Общий объем текстов авторов составил 5 250 182 слова (около 59 МБ).

В то же время литература классических авторов не всегда отражает реалии современного мира и не содержит новые понятийные категории. В связи с этим для расширения лексического разнообразия корпуса и повышения качества и уровня связности текстов [5] будущей малой белорусской языковой модели BelLitGPT в работе были использованы материалы Википедии. Каждая статья прошла процедуру очистки и фильтрации. В результате был получен текстовый файл с отобранными статьями на белорусском языке. Вместе с прозой классиков он сформировал объединенный корпус белорусских текстов объемом 6 333 013 слов (около 75 МБ).

Следует отметить, что сохранение социокультурных ценностей в ответах белорусской генеративной модели в значительной мере достигается через отбор информационного содержания корпуса. Особенность генеративных трансформеров заключается в том, что нормы языка передаются модели через корпус текстов при обучении, поэтому необходимо подготовить тексты, соответствующие языковому стандарту. Это является важной предпосылкой для устойчивого качества генерируемых текстов.

²Отобранные произведения написаны до вступления в силу действующего Закона «О правилах белорусской орфографии и пунктуации» от 23 июля 2008 г. Различия между правилами 1957 и 2008 гг. незначительные.

2. Создание белорусского токенизатора

Предложенный подход к созданию белорусского токенизатора основывается на процедуре расширения словарного запаса (токенизатора) модели ruGPT-3. Русскоязычные токенизаторы могут разбивать слово на несколько токенов, что позволяет адаптировать их к белорусскому языку, в то время как англоязычные одному слову сопоставляют один токен [6].

Белорусский язык отличается от русского по фонетике, лексике, морфологии, орфографии, синтаксису. Между тем базовая механика токенизации – это только разделение слов на токены, поэтому при разработке белорусского токенизатора не требуется учитывать все отличия между языками. Поскольку оба языка используют кириллицу, базовый русскоязычный токенизатор изначально способен обрабатывать белорусскую лексику, разбивая незнакомые слова на более короткие известные участки (субтокены). Однако отсутствие в словаре токенизатора незнакомых букв «ў», «і» и знака «'» (апостроф) приводит к ошибкам в процессе работы. Находя их, он создает токены <unk> (unknown), наличие которых делает дальнейшее обучение невозможным. Добавление указанных символов является минимально необходимым техническим шагом (первым этапом), после которого токенизатор способен обрабатывать белорусские тексты. Для этого авторы создают временный ВРЕ(Byte Pair Encoding)-токенизатор и обучают его брать частые слова целиком и трактовать их как один токен, а редкие или незнакомые разбивать на части [7]. Например, слово 'слоўнік' базовый токенизатор может разбить на неэффективные фрагменты ['сло', 'ў', 'н', 'і', 'к']. В то же время адаптированный токенизатор может распознавать семантическое ядро слова (корень) и его грамматические модификаторы (приставки и суффиксы) как отдельные смысловые единицы. В результате нейросеть обучается не просто запоминать слова целиком, а выводить значение новых или редких терминов, опираясь на знакомые морфемы, что критически важно для белорусского языка с его богатой системой словообразования.

Стоит отметить, что между белорусским и русским языками существуют и другие алфавитные различия, однако они не влияют на работоспособность токенизатора. Так, присутствие в словаре русской буквы «щ» не снижает точности токенизации. Избыточные для белорусского языка символы игнорируются при обработке и не влияют на результат, а при обработке белорусских «дз» и «дж» модель может разделять названные диграфы на разные субтокены, однако это не сказывается на связности и правильности генерируемого текста.

На следующем шаге алгоритм вычисляет разность множеств между словарями исходной модели и новой. Так, он выделяет токены, которые присутствуют в белорусском словаре, но отсутствуют в русском. Далее из этого списка отбирается фиксированное количество наиболее важных токенов, в данной конфигурации 3000. Ограничение в 3000 токенов выбрано как компромисс между полнотой охвата лексики и вычислительной сложностью модели. В список попадают наиболее частотные белорусские лексемы (например, 'гэта', 'ёсць', 'праз') и специфические морфемы, содержащие буквы «і» и «ў». Это дает возможность BellLitGPT воспринимать характерные белорусские слова, например с «ў» и «і», целиком, а не разбивать их на случайные буквы или слоги. Таким образом, представленная методика позволила создать белорусский токенизатор. Его использование повышает качество генерации текста.

В свою очередь, расширение словаря привело к необходимости обновления параметров самой нейросети. Для этого веса из исходной модели ruGPT-3 загрузили в нейросеть и произвели изменение размерности входного и выходного слоев в соответствии с размером нового словаря. Так была получена базовая белорусская модель `rugpt-belarusian-base`, которая готова для обучения на белорусском корпусе текстов.

3. Выстраивание процесса обучения

Процесс обучения построен на базе библиотеки Hugging Face Transformers и включает в себя четыре ключевые стадии: инициализацию, подготовку данных, настройку гиперпараметров и цикл оптимизации.

В начале работы система загружает базовую белорусскую модель `rugpt-belarusian-base` и соответствующий ей токенизатор. Для возможности обучения крупной модели на ограниченных аппаратных ресурсах NVidia RTX 4090 активируется механизм контрольного сохранения градиентов. Это позволяет существенно экономить видеопамять (VRAM) за счет пересчета части градиентов «на лету» вместо их хранения, что в свою очередь дает возможность использовать более крупные пакеты данных.

Далее происходят предварительная обработка и сегментация данных. Действия направлены на оптимизацию эффективного использования контекстного пространства модели и являются наиболее ресурсоемким этапом процесса обучения. Объединенный корпус белорусских текстов подвергается процедуре преобразования в последовательный ряд числовых токенов. Обучение осуществляется не на изолированных предложениях, а на объединенном длинном потоке текста, который разделяется на фрагменты постоянной протяженности (фиксированная длина блока составляет 1024 токена). Это обеспечивает подачу на вход модели максимального объема контекста, который препятствует заполнению пустых позиций, исключает потерю данных на стыках предложений и увеличивает производительность обучения.

Для более эффективной перенастройки модели под новый язык были установлены следующие гиперпараметры обучения:

Скорость обучения. Установлено значение $3 \cdot 10^{-4}$, что несколько выше стандартного для дообучения. Это сделано намеренно для более агрессивной адаптации к новому языковому распределению.

Эпохи. Процесс подразумевает пять полных проходов по всему набору данных.

Смешанная точность. Используются 16-битные числа (FP16) с плавающей запятой для ускорения вычислений и снижения потребления памяти.

Накопление градиента. Позволяет эмулировать большой размер пакетов путем накопления результатов нескольких шагов перед обновлением весов модели.

Затем иницируется итерационный алгоритм оптимизации целевой функции потерь. Автоматизированная система осуществляет управление загрузкой тренировочных данных и регулярное обновление весовых коэффициентов нейронной сети.

По окончании указанного цикла оптимизационного процесса финальная версия обученной модели вместе с откорректированным лексиконом токенизатора сохраняется в каталоге `BellLitGPT`.

Итоговым результатом реализации указанного выше процесса переобучения становится полнофункциональная генеративная модель. Она обладает способностью создавать осмысленные и согласованные тексты на белорусском языке, которые отражают особенности исходного корпуса золотого фонда классической белорусской прозы.

4. Тестирование модели

Для валидации обученной модели BellLitGPT была проведена серия качественных экспериментов. Основной целью тестирования являлась проверка способности модели генерировать текст, соответствующий не только грамматическим и морфологическим нормам, но и семантике белорусского литературного языка.

Тестирование проводилось методом zero-shot-генерации. Его суть состоит в том, что в модель подается начальная фраза (промпт). После генерации ответа оценивается лингвистическая корректность продолжения. Ниже представлены результаты проверки базовых лингвистических компетенций.

- **Тест 1: грамматическое согласование (падежные формы)**

Вход. Сонца павольна схавалася за [...]

Выход: лесам.

Анализ. Модель корректно определила управление предлога *за* (творительный падеж) и сгенерировала существительное в правильной словоформе, что подтверждает усвоение структуры падежных окончаний.

- **Тест 2: морфологическое согласование (род)**

Вход. Пасярод поля цякла глыбокая [...]

Выход: рачулка.

Анализ. Прилагательное *глыбокая* указывает на женский род. Модель успешно согласовала с ним сгенерированное существительное *рачулка*, игнорируя частотные варианты мужского или среднего рода, что говорит о корректном моделировании зависимостей внутри именной группы.

- **Тест 3: логическая и контекстуальная связность**

Вход. Неба раптам пацямнела, і пачаўся моцны [...]

Выход: вецер.

Анализ. В связи с сопряженностью таких природных явлений, как дождь и ветер во время штормовой погоды, продолжением высказывания может быть и «дождь». Однако, если углубиться в физику природных явлений, то резкое потемнение неба (формирование кучево-дождевой облачности) перед грозой чаще всего сопровождается именно шквалистым усилением ветра, которое нередко предшествует самому дождю. Поэтому в рассматриваемом случае модель продемонстрировала понимание причинно-следственных связей.

В контексте оценки генеративных способностей БояМ важно учитывать, что они являются вероятностными, предсказывающими следующий токен на основе распределения вероятностей, сформированного в процессе обучения. Понимание причинно-следственных связей в их работе проявляется в том, что алгоритм снижает вероятность логически несовместимых токенов и выбирает продолжение из семантически согласованного кластера (ветер, дождь, ливень, шторм, ураган). Тот факт, что модель выбирает корректное метеорологическое следствие из вероятностного распределения, свидетельствует

о наличии «здравого смысла» (common sense reasoning) в латентном пространстве модели и, в частности, о том, что она успешно усвоила статистические признаки и контекстуальные зависимости, характерные для описания непогоды. При этом речь идет не о понимании в человеческом смысле, а о согласованном вероятностном прогнозировании на основе выявленных закономерностей в данных.

Приведем дополнительные примеры, которые показывают способность белорусской модели генерировать текст с логической и контекстуальной связностью.

Пример 1.

Вход. Каб адчыніць замкнёныя дзверы, ён дастаў з кішэні [...]

Выход: ключ і пачаў круціць.

Пример 2.

Вход. На вуліцы быў моцны мароз, таму вада ў лужынах хутка [...]

Выход: замярзала.

Пример 3.

Вход. Калі яна пачула гэтую радасную навіну, на яе твары з'явілася [...]

Выход: ўсмешка.

Пример 4.

Вход. Каб купіць свежага хлеба і малака, яна ўзяла торбу і пайшла ў [...]

Выход: мястэчка.

• **Тест 4: лексическая сочетаемость**

Вход. Аўтар у сваім творы хацеў паказаць жыццё простым і [...]

Выход: ясным.

Анализ. Модель восстановила устойчивую литературную коллокацию *простым і ясным*. Это свидетельствует о глубоком влиянии корпуса художественной литературы, так как подобные обороты характерны для публицистики и литературной критики.

Если первая серия тестов подтвердила компетенцию модели на уровне предложения, то следующий этап валидации был направлен на оценку генеративных способностей на макроуровне (уровне связного текста). Проведенные испытания показали способность BelLitGPT развивать контекст через добавление новых персонажей, описание природы и включение монологов в созданном тексте из нескольких предложений. Приведем пример генерации текста с входящим запросом «Стары замак стаяў на гары...» и оценим его.

Промпт. Стары замак стаяў на гары [...]

Результат:

...і з замкавай вежы можна было бачыць увесь горад. А калі ўздымалася сонца над горадам, то яно стаяла высока, і здавалася, што гэта сам Стары Замак відзён з вышыні. У такія хвіліны сэрцу робіцца лёгка ад таго, што ўсё навокал свеціцца залатым бляскам. І ў гэтыя хвіліны яснай летняй раніцы Васіль думаў пра свой родны кут: які ён бедны! Хто б мог падумаць...

Проведем комплексный стилистико-семантический анализ сгенерированного фрагмента:

Семантическое расширение. Модель логично развила пространственную сцену: от замка на горе до панорамы города, видимой с башни.

Эмоциональная окраска. Текст содержит сложные синтаксические конструкции, описывающие внутреннее состояние (*сэрцу робіцца лёгка*), что свойственно художественной прозе.

Стилистическая мимикрия. Появление лирического героя (Васіля) и резкий переход к социальной рефлексии (*які ён бедны!*) ярко демонстрируют влияние обучающего корпуса классиков (Я. Коласа, К. Чорного), для которых характерно переплетение пейзажной лирики с социальными мотивами.

5. Нейросимвольный подход для генерации четверостиший

Для дополнительной иллюстрации способностей BelLitGPT генерировать текст, придерживаясь стилистических и семантических правил, был рассмотрен особый случай – создание стихов. Для этого использовалась архитектура гибридного конвейера, основанная на нейросимвольном подходе и объединяющая нейросетевые модели генерации текста, жесткие алгоритмические ограничения (задающие правила ритма и рифмы) и верификацию с помощью БоЯМ, которая играет роль судьи [8].

Для того чтобы BelLitGPT стала генерировать четверостишия, ее дообучили на поэме Якуба Коласа «Новая Зямля». Для этого из всей поэмы было составлено множество пар «запрос – ответ», где запросом является первая строка четверостишия, а ответом – три остальные. Далее модель BelLitGPT дообучили на всем множестве пар. Говоря про процедуру дообучения моделей, исследователи отмечают, что она базируется на единой методологии непрерывного дообучения (continuous fine-tuning) и является универсальной для всех языков [4]. Отличиям же подвержены архитектурные особенности построения моделей с использованием нейросимвольного подхода [8, 10, 11].

Процесс дообучения строится на подаче структурированных пар данных, обрамленных специальными токенами: `<|startoftext|>` для обозначения начала, `<|sep|>` для разделения запроса и ответа и `<|endoftext|>` для фиксации конца последовательности. На каждой итерации модель генерирует предсказание следующего токена, которое сопоставляется с эталонным текстом для вычисления функции потерь. Через алгоритм обратного распространения ошибки выявленное отклонение корректирует внутренние веса нейросети, что постепенно минимизирует влияние общего корпуса и адаптирует модель к особенностям целевого текста (в данном случае к стихам в размере четырехстопного ямба). С формальной точки зрения вышеуказанный алгоритм минимизирует функцию потерь:

$$L = -\sum_{i=1}^n \log P(y_i | x, y_{<i>}),$$

где x – первая строка, y_i – токены трех последующих строк.

Первая попытка генерации продолжения по одной заданной строчке оказалась неудачной. Соблюдение ритма и рифмы было скорее исключением, чем правилом. Например, на 1000 попыток со строкой «Гляджу ў акно на ціхі сад» нашлась только одна со схемой рифмовки ААВВ³.

Пример 1.

*Гляджу ў акно на ціхі сад,
Нібы пытае воўк цыкад...*

³Схема построения строфы, при которой рифмуются две соседние строки (первая со второй, третья с четвертой). Пример: Мой родны кут, як ты мне мілы! (А) Забыць цябе не маю сілы! (А) Не раз, утомлены дарогай, (В) Жыццём вясны мае убогай... (В)

Поэтому для успешного создания четверостиший было предложено использовать генератор в двух разных режимах: в прямой генерации и обратной.

Прямая генерация используется для создания трех последующих строк по одной данной первой строке. Модель предсказывает продолжение текста на основе контекста с изменяющейся для повышения вариативности температурой воспроизведения от 0,9 до 1,3.

Обратная генерация применяется для строк, которые требуют соблюдения рифмы. Алгоритм сначала выбирает целевое рифмующееся слово из словаря, а затем генерирует предшествующий ему текст, обеспечивая когерентность всего четверостишия.

В случае обратной генерации необходимо следить, чтобы в созданном ответе сохранялся смысл. Для этого используется более совершенная языковая модель Gemma 2 (с 9 млрд параметров) [9]. Она оценивает создаваемые строки с точки зрения смысла $Score_{sense}$ и грамматики $Score_{gram}$ и выбирает ту, у которой лучшие оценки. Такой подход соответствует стандартной для машинного обучения схеме «генератор – дискриминатор» [10].

Между тем для обеспечения ритма и рифмы также требуется специальный детерминированный модуль. Он был построен с применением фонетического анализа, который совершается за счет проставленных ударений в грамматической базе данных ресурса bnkorus.info⁴. Модуль допускает строки длиной восемь или девять слогов с чередованием ударных и безударных позиций (хорей/ямб) [11]. Рифма определяется по совпадению фонетических концовок слов (ударная гласная и последующие звуки) с учетом вариативности гласных, например *ы/і*.

5.1. Алгоритм генерации четверостишия

Генерация строфы по схеме ААВВ реализуется следующим образом:

1. На вход подается иницирующая строка L_1 .
2. Для генерации парной строки L_2 производится поиск множества рифм $R = \{r_1, r_2, \dots, r_n\}$ к последнему слову L_1 .
3. Запускается итеративный процесс (до 100 попыток), в котором генератор создает кандидатов, заканчивающихся на $r_i \in R$.
4. Кандидаты фильтруются по метрике и передаются модулю валидации. Лучший кандидат по сумме баллов ($Score_{gram} + Score_{sense}$) принимается как L_2 .
5. Аналогичный процесс повторяется для пары строк L_3 и L_4 , где L_3 генерируется в режиме прямой генерации для развития сюжета.

Пример 2.

*Гарыць агонь у печы зноў
І бліскае там язычкоў
Над квольым лісцем, над асінай
Ды над кудзелай павучынай*

Пример 3.

*Гучыць жалейка над ракой
І плача бедная святой
Над гэтай гразёй, над балотам
І над табою абармотам*

⁴Нацыянальны корпус беларускай мовы. – URL: <http://bnkorus.info> (дата обращения: 26.02.2026). Грамматическая база создана, поддерживается и обновляется сотрудниками сектора компьютерной лингвистики ГНУ «Центр исследований белорусской культуры, языка и литературы НАН Беларуси».

Пример 4.

*Вятрыска гне стары дубок
І буйны ветрык-вечарок
Над гэтым лесам парыўся
І мёдам росным марыўся*

Пример 5.

*Снягі ўкрылі ціхі сад
І ночкі снегам неўпапад
Па лесе разышліся, хвалі
І белым пухам дабягалі*

Рассмотрим отличия предложенного авторами нейросимвольного подхода от других существующих в литературе. Так, в модели Neural Poetry [11] входными токенами выступают исключительно слоги. В нейросимвольном подходе применяется процедура расширения словарного запаса ВРЕ-токенизатора, что позволяет нейросети распознавать семантическое ядро слова (корень) и его грамматические модификаторы (приставки и суффиксы) как отдельные смысловые единицы. Это критически важно для корректной работы с богатой системой словообразования.

Модель Deep-speare [10] решает задачу полностью нейросетевым способом, обучаясь ритму и рифме без внешних словарей и эвристик. Система Hafez [8] использует конечный автомат, который жестко ограничивает пространство на этапе лучевого поиска. В архитектуре гибридного конвейера используется специальный детерминированный фонетический модуль на основе грамматической базы белорусского языка (bnkorus.info), который отбирает кандидатов с длиной восемь-девять слогов и чередованием ударных (безударных) позиций.

Если Neural Poetry оценивает сгенерированные пакеты с помощью математических штрафов за нарушение формы, а Deep-speare и Hafez полагаются на вероятности своих внутренних рекуррентных моделей, то авторы применяют внешнюю языковую модель в роли независимого судьи. Она оценивает кандидатов (в частности, при обратной генерации) с точки зрения сохранения смысла и грамматики, реализуя соревновательную схему «генератор – дискриминатор».

Указанные отличия делают предложенную архитектуру оптимальной для работы с малыми языковыми моделями на вычислительных устройствах потребительского класса. Разработанный конвейер доказывает, что в условиях указанных ограничений можно успешно генерировать стихи, демонстрируя высокую связность текста и точное сохранение сложной поэтической структуры.

Заключение

В работе представлена методика создания национальной большой языковой модели BelLitGPT, основанная на стратегии трансферного обучения и адаптации токенизатора. Предложенный подход показал, что использование архитектуры модели (ruGPT-3) с родственным языком позволяет эффективно преодолевать нехватку подготовленных текстовых данных на белорусском языке.

Ключевым результатом исследования стала успешно натренированная модель BelLitGPT с 700 млн параметров. Эксперименты подтвердили, что даже при сравнительно небольшом объеме обучающего корпуса (6 333 013 слов или порядка 75 МБ), состоящего из произведений классической белорусской прозы и статей Википедии, модель способна усваивать грамматический строй, морфологические нормы и стилистические особенности белорусского языка. Качественная валидация показала способность

нейросети генерировать связные тексты, поддерживать контекст и демонстрировать элементы «здравого смысла». Кроме того, на базе полученной модели была успешно реализована система генерации четверостиший, которая использует нейросимвольный подход с верификацией ритма и рифмы.

Между тем в ходе исследования были выявлены и существенные ограничения, связанные с масштабированием архитектуры. Параллельный эксперимент по дообучению более крупной версии модели ruGPT-3 с 13 млрд параметров на вычислительном устройстве на базе процессора AMD Ryzen AI Max+ PRO 395 с 128 ГБ оперативной памяти LPDDR5x-8000 показал, что имеющегося объема текстовых данных недостаточно для качественной перенастройки столь массивной нейронной сети. Из-за дисбаланса между количеством обучаемых параметров и размером корпуса модель на 13B не смогла полностью переключиться на целевой язык. В ее генерациях сохранялось сильное влияние исходной модели, что выражалось в частом появлении русских слов и смешении грамматических правил.

Таким образом, для создания компактных моделей (до 1 млрд параметров) критически важными являются качество и чистота данных, в то время как для обучения крупных моделей (больше 10 млрд параметров) первоочередной задачей становится кардинальное увеличение объема корпуса. Собрать 1000 МБ подготовленных аутентичных белорусских текстов становится отдельной задачей, выполнение которой позволит приступить к обучению крупных моделей.

Вклад авторов. А. М. Бондоловский и Д. А. Ляхов внесли вклад в планирование исследования, анализ и интерпретацию данных и подготовку статьи. С. В. Кругликов и К. К. Шульган внесли вклад в подготовку статьи, осуществили редакцию статьи с точки зрения существенного интеллектуального содержания и утвердили окончательную версию статьи для публикации.

References

1. Brown T., Mann B., Ryder N., Subbiah M., Kaplan J. D., ..., Amodei D. Language models are few-shot learners. *Advances in Neural Information Processing Systems*, 2020, vol. 33, pp. 1877–1901.
2. Radford A., Wu J., Child R., Luan D., Amodei D., Sutskever I. Language models are unsupervised multitask learners. *OpenAI*, 2019. Available at: https://cdn.openai.com/better-language-models/language_models_are_unsupervised_multitask_learners.pdf (accessed 03.11.2025).
3. Vaswani A., Shazeer N., Parmar N., Uszkoreit J., Jones L., ..., Polosukhin I. Attention is all you need. *Advances in Neural Information Processing Systems*, 2017, vol. 30, pp. 5998–6008.
4. Artetxe M., Ruder S., Yegorikhin D. On the cross-lingual transferability of monolingual representations. *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics, Online, 5–10 July 2020*, pp. 4623–4637.
5. Jakubíček M., Kilgarriff A., Kovář V., Rychlý P., Suchomel V. The tenten corpus family. *Proceedings of the 7th International Corpus Linguistics Conference (CL2013), Lancaster University, United Kingdom, 22–26 July 2013*, pp. 125–127.
6. Sennrich R., Haddow B., Birch A. Neural machine translation of rare words with subword units. *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics, Berlin, Germany, 7–12 August 2016*, vol. 1, pp. 1715–1725.

7. Imamura K., Sumita E. Vocabulary adaptation for domain adaptation in neural machine translation. *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing, Brussels, Belgium, 31 October – 4 November 2018*, pp. 4623–4637.
8. Ghazvininejad M., Shi X., Choi Y., Knight K. Hafez: an interactive poetry generation system. *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics, Vancouver, Canada, 30 July – 4 August 2017*, pp. 43–48.
9. Mesnard T., Hardin C., Dadashi R., Bhupatiraju S., Pathak S., ..., Eck D. *Gemma: Open models based on Gemini research and technology*, 2024. Available at: <https://arxiv.org/pdf/2403.08295> (accessed 03.11.2025).
10. Lau J. H., Cohn T., Baldwin T., Brooke J., Hammond A. Deep-speare: A joint neural model of poetic language, meter and rhyme. *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics, Melbourne, Australia, 15–20 July 2018*, vol. 1, pp. 1948–1958.
11. Zugarini A., Melacci S., Maggini M. Neural poetry: Learning to generate poems using syllables. *Artificial Neural Networks and Machine Learning – ICANN 2019: Text and Time Series: 28th International Conference on Artificial Neural Networks, Munich, Germany, 17–19 September 2019*, pp. 313–325.

Информация об авторах

Ляхов Дмитрий Александрович, кандидат физико-математических наук, старший научный сотрудник, Объединенный институт проблем информатики Национальной академии наук Беларуси.
E-mail: dlyakhov@newman.bas-net.by

Андрей Михайлович Бондоловский, кандидат экономических наук, заведующий лабораторией распознавания и синтеза речи, Объединенный институт проблем информатики Национальной академии наук Беларуси.
E-mail: a.bandalouski@newman.bas-net.by

Сергей Владимирович Кругликов, доктор военных наук, кандидат технических наук, доцент, главный научный сотрудник, Объединенный институт проблем информатики Национальной академии наук Беларуси.
E-mail: kruglikov_s@newman.bas-net.by

Константин Константинович Шульган, заместитель генерального директора по цифровому развитию, Объединенный институт проблем информатики Национальной академии наук Беларуси.
E-mail: skk@newman.bas-net.by

Information about the authors

Dmitry A. Lyakhov, Cand. Sci. (Phys.-Math.), Senior Researcher, The United Institute of Informatics Problems of the National Academy of Sciences of Belarus.
E-mail: dlyakhov@newman.bas-net.by

Andrei M. Bandalouski, Cand. Sci. (Econ.), Head of Laboratory of Speech Synthesis and Recognition, The United Institute of Informatics Problems of the National Academy of Sciences of Belarus.
E-mail: a.bandalouski@newman.bas-net.by

Sergey V. Kruglikov, Dr. Sci. (Milit.), Cand. Sci. (Eng.), Assoc. Prof., Principal Researcher, The United Institute of Informatics Problems of the National Academy of Sciences of Belarus.
E-mail: kruglikov_s@newman.bas-net.by

Konstantin K. Shulgan, Deputy General Director for Digital Development, The United Institute of Informatics Problems of the National Academy of Sciences of Belarus.
E-mail: skk@newman.bas-net.by

ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ INFORMATION TECHNOLOGIES

УДК 004.728.4
<https://doi.org/10.37661/1816-0301-2026-23-1-39-57>

Поступила в редакцию | Received 20.02.2026
Подписана в печать | Accepted 13.03.2026
Опубликована | Published 31.03.2026

Алгоритмы многокритериальной маршрутизации в беспроводных mesh-сетях

А. К. Шульган[✉], Ю. И. Воротницкий
[✉]E-mail: anton.shulgan@mail.ru

*Белорусский государственный университет,
пр. Независимости, 4, Минск, 220030, Беларусь*

Аннотация

Цели. Современные беспроводные mesh-сети предъявляют высокие требования к адаптивности протоколов маршрутизации. Стандартные алгоритмы не всегда способны обеспечить требуемое качество обслуживания (QoS) из-за изменчивости таких параметров, как отношение сигнал/шум, загрузка канала и мобильность узлов.

Методы. Предлагается метод многокритериальной маршрутизации, использующий интегральный критерий качества обслуживания в модифицированном алгоритме Дейкстры.

Результаты. Исследованы результаты маршрутизации в схеме из восьми узлов и в схеме с четырьмя кластерами из пяти устройств. Предложен механизм снижения вероятности ложного отбрасывания маршрута в методе многокритериальной маршрутизации.

Заключение. Разработан итерационный алгоритм настройки весовых коэффициентов, который в сочетании с минимаксным критерием позволяет исключать ложные решения об отсутствии QoS-осуществимого маршрута и получать приемлемые решения для разных профилей трафика. Разработаны двухуровневый подход и алгоритм маршрутизации в кластеризованной сети, обеспечивающие снижение вычислительной сложности и локализацию перерасчета маршрутов при изменении состояния сети – качества радиоканалов или энергосостояния узлов.

Ключевые слова: mesh-сеть, алгоритм Дейкстры, многокритериальная маршрутизация, кластеры, качество обслуживания, интегральный критерий, энергоэффективность

Для цитирования. Шульган, А. К. Алгоритмы многокритериальной маршрутизации в беспроводных mesh-сетях / А. К. Шульган, Ю. И. Воротницкий // Информатика. – 2026. – Т. 23, № 1. – С. 39–57. – <https://doi.org/10.37661/1816-0301-2026-23-1-39-57>.

Конфликт интересов. Авторы заявляют об отсутствии конфликта интересов.

Multicriteria routing algorithms in wireless mesh networks

Anton K. Shulgan[✉], Yury I. Vorotnitsky

[✉]E-mail: anton.shulgan@mail.ru

*Belarusian State University,
av. Nezavisimosti, 4, Minsk, 220030, Belarus*

Abstract

Objectives. Modern wireless mesh networks place high demands on the adaptability of routing protocols. Standard algorithms are not always able to ensure the required quality of service (QoS) due to variability in parameters such as signal-to-noise ratio, channel load, and node mobility.

Methods. A method for multi-criteria routing is proposed using an integral quality of service criterion in a modified Dijkstra's algorithm.

Results. Routing results in an eight-node network and in a network with four clusters of five devices are analyzed. A mechanism for reducing the probability of false route discards in the multi-criteria routing method is proposed.

Conclusion. An iterative algorithm for adjusting weighting coefficients has been developed. In combination with the minimax criterion, it allows to eliminate false positives regarding the absence of a QoS-feasible route and to obtain acceptable solutions for different traffic profiles. A two-level approach and routing algorithm in a clustered network have been developed, ensuring a reduction in computational complexity and localization of route recalculation when the network state changes – specifically the quality of radio channels or the energy state of nodes.

Keywords: mesh, Dijkstra's algorithm, multi-criteria routing, clusters, QoS, integral criterion, energy efficiency

For citation. Shulgan A. K., Vorotnitsky Yu. I. *Multicriteria routing algorithms in wireless mesh networks*. *Informatika [Informatics]*, 2026, vol. 23, no. 1, pp. 39–57 (In Russ.). <https://doi.org/10.37661/1816-0301-2026-23-1-39-57>.

Conflict of interests. The authors declare of no conflict of interest.

Введение

Беспроводные сенсорные сети, имеющие mesh-топологии, строятся как самоорганизующиеся сети, где узлы одновременно выступают источниками данных и маршрутизаторами. Такая архитектура повышает покрытие и отказоустойчивость, но усиливает требования к маршрутизации: каналы подвержены затуханию и интерференции, узлы обладают ограниченными запасами энергии, а приложения предъявляют требования к надежности и параметрам качества обслуживания. Вследствие этого маршрутизация в беспроводных сенсорных сетях рассматривается как задача компромисса между несколькими критериями. Отмечается, что распространены два способа формализации компромисса в многокритериальной задаче оптимизации: сведение метрик к единой функции стоимости (например, взвешенной сумме) и оптимизация при ограничениях, когда один критерий минимизируется при выполнении порогов по другим [1–3]. При этом отмечается (см., например [4, 5]), что скаляризация метрик и «взвешенные» маршруты – наиболее актуальный и практичный для ресурсно-ограниченных узлов подход.

В работах [6–10] рассматривалась задача поиска оптимального маршрута между узлами мультисервисной телекоммуникационной сети при наличии нескольких параметров QoS, в [9] указывались оптимальные решения оптимизационных задач. В качестве параметров QoS каждого канала связи используются: полоса пропускания Y , задержка D , вариация задержки (джиттер) J и вероятность потерь пакетов Z . Путь называется QoS-осуществимым, если для него выполняются ограничения на каждый параметр QoS:

$$Y \geq Y_{min}; Z \leq Z_{max}; D \leq D_{max}; J \leq J_{max}. \quad (1)$$

Для удобства расчета потерь используется преобразование $X = -\ln(1 - Z)$, позволяющее вычислять соответствующий показатель по пути как сумму по ребрам. Затем многокритериальная задача сводится к минимизации скалярного интегрального критерия качества R , получаемого сверткой нормированных параметров QoS с весовыми коэффициентами. Модификация алгоритма Дейкстры заключается в использовании расширенной метки вершины, содержащей значения параметров QoS по текущему пути и величину R , а также в отбрасывании путей, не удовлетворяющих ограничениям QoS. В настоящей работе данный подход, позволяющий объединять два названных выше способа формализации многокритериальной задачи, обобщается для критериев качества и архитектурных решений беспроводных mesh-сетей.

Предложенная модель в настоящей работе адаптируется к беспроводным mesh-сетям. В беспроводных сенсорных mesh-сетях маршрутизация должна учитывать не только классические QoS-параметры, но и энергетические ограничения узлов. В частности, для задач умного города и инфраструктурного мониторинга важно предотвращать деградацию связности сети из-за исчерпания энергии промежуточных ретрансляторов.

1. Интегральный критерий качества обслуживания в модифицированном алгоритме Дейкстры

Для беспроводной mesh-сети рассмотрим граф $G(V, E)$, где V – множество узлов, E – множество двунаправленных радиоканалов (ребер). Каждому ребру e_{ij} сопоставляется вектор параметров QoS: $(Y_{ij}, D_{ij}, J_{ij}, Z_{ij}, T_{ij})$. Для преобразования вероятностей потерь используется соотношение $X_{ij} = -\ln(1 - Z_{ij})$.

Пусть каждому узлу v соответствует запас энергии E_v (в Вт·ч). Для типа трафика k задается усредненная мощность P_k (Вт), характеризующая энергопотребление узла при участии в передаче, приеме и ретрансляции пакетов данного типа. Тогда оценка времени жизни узла для трафика k определяется как

$$T_{v,k} = E_v / P_k, \text{ ч.} \quad (2)$$

Время жизни маршрута p для трафика k предлагается оценивать как минимум по всем узлам, входящим в маршрут:

$$T_{p,k} = \min\{T_{v,k} : v \in p\}, \text{ ч.} \quad (3)$$

Такое определение соответствует «узкому месту» по энергии: маршрут считается устойчивым лишь до момента, когда первым исчерпает энергию какой-либо из его узлов. При расширении пути на новую вершину параметр T обновляется операцией минимума, что обеспечивает монотонность (неулучшаемость) критерия при добавлении ребер.

Для пути $p = (s, \dots, t)$ значения показателей QoS вычисляются по правилам

$$Y_p = \min\{Y_e : e \in p\}; D_p = \Sigma D_e, J_p = \Sigma J_e, X_p = \Sigma X_e, e \in p;$$

$$Z_p = 1 - \exp(-X_p), T_{p,k} = \min\{T_{v,k} : v \in p\}. \quad (4)$$

Путь p считается QoS-осуществимым для трафика типа k , если выполняются ограничения

$$Y_p \geq Y_{min}, \quad (5)$$

$$D_p \leq D_{max}, \quad (6)$$

$$J_p \leq J_{max}, \quad (7)$$

$$X_p \leq X_{max}, \quad (8)$$

$$T_p \geq T_{min}. \quad (9)$$

Для выбора оптимального маршрута среди QoS-осуществимых может использоваться интегральный критерий (минимизируемая свертка нормированных показателей):

$$R_p = w_y \cdot \left(\frac{Y_{min}}{Y_p}\right) + w_D \cdot \left(\frac{D_p}{D_{max}}\right) + w_j \cdot \left(\frac{J_p}{J_{max}}\right) + w_X \cdot \left(\frac{X_p}{X_{max}}\right) + w_T \cdot \left(\frac{T_{min}}{T_p}\right). \quad (10)$$

Здесь w_y, w_D, w_j, w_X, w_T – весовые коэффициенты ($w_y + w_D + w_j + w_X + w_T = 1$), задаваемые в соответствии с требованиями обслуживаемого приложения (класса трафика). Величины, которые необходимо максимизировать (Y_p и T_p), включены в критерий в виде обратных отношений, что позволяет сохранять задачу минимизации.

Чтобы учесть важный для сенсорных сетей параметр – время жизни маршрута, в предлагаемом алгоритме для каждой вершины v в процессе поиска поддерживается расширенная метка $L(v)$, включающая $L(v) = (R_v, Y_v, D_v, J_v, X_v, T_v, pred(v))$. Значения Y_v, D_v, J_v, X_v, T_v описывают параметры QoS для текущего лучшего по значению R пути из источника s в вершину v , а $pred(v)$ – номер предшествующей вершины на этом пути.

Правила обновления метки при релаксации (включении в маршрут) ребра (v, u) имеют вид

$$Y'_u = \min(Y_v, Y_{vu}); D'_u = D_v + D_{vu}; J'_u = J_v + J_{vu};$$

$$X'_u = X_v + X_{vu}; T'_u = \min(T_v, T_{u,k}). \quad (11)$$

Если хотя бы одно из ограничений QoS нарушается, то переход (релаксация) не выполняется (эквивалентно присвоению стоимости бесконечности). В противном случае вычисляется R_u' по формуле критерия и метка $L(u)$ заменяется, если R_u' меньше текущего значения R_u . Поскольку D , J и X накапливаются суммированием, а Y и T обновляются операцией минимума, значение R не может уменьшаться при добавлении ребер. Это обеспечивает применимость стратегии выбора следующей вершины с минимальным значением R в соответствии с алгоритмом Дейкстры. Таким образом, предлагаемый алгоритм может быть описан следующим образом:

Вход: граф $G(V, E)$, источник s , приемник t , профиль трафика k (ограничения и весовые коэффициенты), значения энергии E_v .

Шаг 1. Для всех $v \in V$: $R(v) = \infty$; $\text{pred}(v) = \emptyset$.

Шаг 2. Инициализация: $Y(s) = \infty$; $D(s) = 0$; $J(s) = 0$; $X(s) = 0$; $T(s) = \frac{E_s}{P_k}$; $R(s) = w_T \frac{T_{min}}{T_s}$.

Шаг 3. $Q \leftarrow V$ (множество непосещенных вершин).

Шаг 4. Пока $Q \neq \emptyset$:

Шаг 4.1. $v \leftarrow \text{argmin}_{\{u \in Q\}} R(u)$, удалить v из Q .

Шаг 4.2. Если $v = t$, выход.

Шаг 4.3. Для каждого соседа u вершины v :

Шаг 4.3.1. Вычислить Y', D', J', X', T' по правилам релаксации.

Шаг 4.3.2. Если (Y', D', J', X', T') удовлетворяют ограничениям, положить

$$R' = w_Y \cdot \left(\frac{Y_{min}}{Y'} \right) + w_D \cdot \left(\frac{D'}{D_{max}} \right) + w_J \cdot \left(\frac{J'}{J_{max}} \right) + w_X \cdot \left(\frac{X'}{X_{max}} \right) + w_T \cdot \left(\frac{T_{min}}{T'} \right).$$

Если $R' < R(u)$, обновить метку u , $\text{pred}(u) = v$.

Выход: путь p по $\text{pred}(\cdot)$ и его показатели QoS.

Для анализа работы алгоритма рассмотрим mesh-сеть из восьми узлов с двунаправленными радиоканалами. Параметры каналов (Y, D, J, Z) заданы как усредненные величины (табл. 1), характерные для беспроводной среды (помехи, многолучевость, повторные передачи). На рис. 1 обозначена топология примерной mesh-сети, с которой будет работать алгоритм.

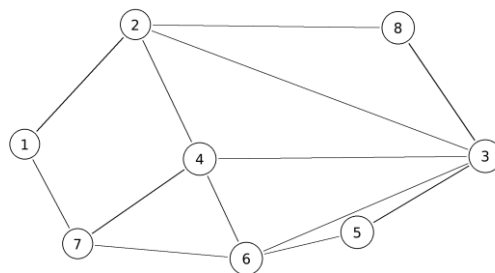


Рис. 1. Топология примерной mesh-сети

Fig. 1. Sample mesh network topology

Таблица 1
Параметры радиоканалов (QoS) для примерной mesh-сети

Table 1
Radio channel parameters (QoS) for a sample mesh network

Ребро <i>Edge</i>	Y, полоса, Мбит/с <i>Bandwidth, Mbps</i>	D, задержка, мс <i>Delay, ms</i>	J, джиттер, мс <i>Jitter, ms</i>	Z, потери <i>Loss</i>
1–2	10	12	3	0,02
1–7	8	8	2	0,005
2–3	5	20	8	0,04
2–4	7	10	4	0,03
2–8	12	15	5	0,02
3–4	8	12	4	0,01
3–5	6	10	3	0,01
3–6	4	18	6	0,02
3–8	10	14	4	0,03
4–6	7	11	3	0,02
4–7	9	9	3	0,01
5–6	7	9	2	0,01
6–7	6	10	3	0,01

Для трех профилей трафика заданы ограничения QoS и веса интегрального критерия (табл. 2). Маршрутизация выполняется из узла 1 (источник) в узел 3 (получатель). Информация о трафике представлена в табл. 3.

Таблица 2
Запасы энергии узлов и оценка времени жизни T для профилей трафика

Table 2
Energy reserves of nodes and lifetime estimate T for traffic profiles

Узел <i>Edge</i>	E, энергия, Вт·ч <i>Energy, W·hrs</i>	T для файлов, ч <i>T for files, hrs</i>	T для голоса, ч <i>T for voice, hrs</i>	T для телеметрии, ч <i>T for telemetry, hrs</i>
1	20	25	28,57	40
2	15	18,75	21,43	30
3	20	25	28,57	40
4	9	11,25	12,86	18
5	10	12,5	14,29	20
6	8	10	11,43	16
7	14	17,5	20	28
8	16	20	22,86	32

Таблица 3
 Ограничения и веса критерия для примеров трафика

Table 3
 Limitations and criteria weights for traffic examples

Профиль Profile	Y_{min} , Mbps	D_{max} , ms	J_{max} , ms	Z_{max}	T_{min} , hrs	w_Y	w_D	w_J	w_X	w_T
Передача файлов (протокол FTP) File Transfer Protocol	6	80	30	0,15	8	0,50	0,15	0,10	0,10	0,15
Голос Voice (VoIP)	1	50	10	0,05	6	0,05	0,35	0,25	0,25	0,10
Телеметрия и мониторинг Telemetry and Monitoring (IoT)	0,5	120	40	0,1	18	0,10	0,15	0,05	0,20	0,50

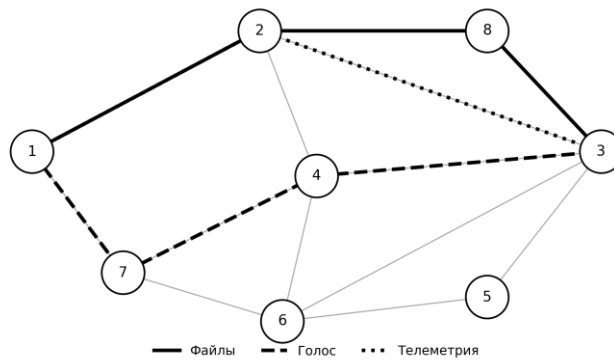


Рис. 2. Оптимальные маршруты для различных типов трафика
 Fig. 2. Optimal routes for different types of traffic

Таблица 4
 Результат маршрутизации: выбранные маршруты и показатели QoS

Table 4
 Routing result: selected routes and QoS indicators

Тип трафика Type of traffic	Оптимальный маршрут Optimal route	Y_p , Mbps	D_p , ms	J_p , ms	Z_p	T_p , hrs	R
Передача файлов (протокол FTP) File Transfer Protocol	1-2-8-3	10	41	12	0,0684	18,75	0,5245
Голос Voice (VoIP)	1-7-4-3	8	29	9	0,0248	12,86	0,6033
Телеметрия и мониторинг Telemetry and Monitoring (IoT)	1-2-3	5	32	11	0,0592	30	0,4796

Результаты работы алгоритма представлены на рис. 2 и в табл. 4. Анализ работы предлагаемого алгоритма показывает, что модификация алгоритма Дейкстры с расширенной меткой позволяет решать задачу многокритериальной QoS-маршрутизации в mesh-сетях с учетом неаддитивных показателей (минимальная полоса пропускания и время жизни маршрута). При этом введение критерия времени жизни T обеспечивает энергоосведомленный выбор маршрута и позволяет избегать путей, проходящих через узлы с ограниченными запасами энергии. Из приведенного примера видно, что различные профили трафика приводят к выбору разных маршрутов за счет изменения ограничений и весов критериев.

2. Устранение ложного отбрасывания QoS-осуществимых путей

При реализации алгоритмов маршрутизации на основе предлагаемой модификации алгоритма Дейкстры для многокритериальной QoS-маршрутизации с расширенной меткой вершины может возникнуть следующая проблема: поскольку алгоритм Дейкстры является алгоритмом поиска без возврата, для каждой вершины фиксируется только один «лучший» по выбранному скалярному критерию путь. В результате возможно ложное решение об отсутствии QoS-осуществимого маршрута между узлами s и t , если выбранная оптимальная траектория к промежуточному узлу r не допускает дальнейшего продолжения до t , хотя альтернативная (хуже по скалярному критерию) траектория может приводить к QoS-осуществимому маршруту.

Для снижения вероятности ложного отбрасывания предлагается использовать два взаимодополняющих приема:

1) переход от аддитивной свертки к минимаксному критерию, который делает решение более устойчивым к «выбросу» одного из нормированных показателей;

2) итерационную настройку весовых коэффициентов: на каждом запуске алгоритма ограничения QoS на этапе релаксаций не проверяются; после получения оптимального по критерию маршрута выполняется проверка QoS-ограничений. При нарушении ограничений веса соответствующих показателей увеличиваются до выполнения ограничений либо до достижения предельных значений.

Минимаксный критерий для пути p (при пяти показателях QoS: Y , D , J , X и T) задается как минимум по максимуму взвешенных нормированных величин:

$$R_p = \max \left\{ w_Y \cdot \left(\frac{Y_{min}}{Y_p} \right), w_D \cdot \left(\frac{D_p}{D_{max}} \right), w_J \cdot \left(\frac{J_p}{J_{max}} \right), w_X \cdot \left(\frac{X_p}{X_{max}} \right), w_T \cdot \left(\frac{T_{min}}{T_p} \right) \right\}. \quad (12)$$

Условия QoS-осуществимости при этом остаются прежними: $Y_p \geq Y_{min}$, $D_p \leq D_{max}$, $J_p \leq J_{max}$, $Z_p \leq Z_{max}$ (эквивалентно $X_p \leq X_{max}$, где $X_{max} = -\ln(1 - Z_{max})$), $T_p \geq T_{min}$.

Итерационный алгоритм выполняет последовательные запуски модифицированного алгоритма Дейкстры с минимаксным критерием. Обозначим через $r_i(p)$ нормированное значение i -го показателя для найденного маршрута p . Для ограничений типа «не более» имеем $r_D = D_p/D_{max}$, $r_J = J_p/J_{max}$, $r_X = X_p/X_{max}$. Для ограничений типа «не менее» используем $r_Y = Y_{min}/Y_p$ и $r_T = T_{min}/T_p$. Маршрут является QoS-осуществимым тогда и только тогда, когда $r_i(p) \leq 1$ для всех i .

Если после очередного запуска получен маршрут, нарушающий ограничения, формируется множество нарушенных показателей $S = \{ i : r_i(p) > 1 \}$. Для них увеличиваются

веса w_i . Предельные значения выбираются как $1/|S|$ (т. е., если нарушен один показатель, его вес может быть увеличен до 1; если нарушены два, – до 0,5 и т. д.). Оставшаяся доля веса распределяется между ненарушенными показателями пропорционально их текущим значениям весов с сохранением условия $w_Y + w_D + w_J + w_X + w_T = 1$.

Алгоритм можно описать следующей последовательностью шагов:

Вход. Задать профиль трафика (ограничения QoS) и значения весов w_Y, w_D, w_J, w_X, w_T .

Шаг 1. Выполнить поиск маршрута p^* , минимизирующего минимаксный критерий R_p , с использованием модифицированного алгоритма Дейкстры. На этапе релаксаций проверка ограничений QoS не выполняется.

Шаг 2. Для найденного маршрута вычислить показатели QoS (Y_p, D_p, J_p, Z_p, T_p) и нормированные значения $r_i(p^*)$.

Шаг 3. Если $r_i(p^*) \leq 1$ для всех i , то маршрут p^* принимается как QoS-осуществимый и оптимальный по критерию R_p .

Шаг 4. Если существует хотя бы один показатель i , для которого $r_i(p^*) > 1$, увеличить веса нарушенных показателей и повторить шаги 1–3. Если веса нарушенных показателей достигли предельных значений и улучшение отсутствует, итерации прекращаются.

Выход. Путь p и его показатели QoS.

В качестве примера рассмотрим mesh-сеть из 10 узлов с двунаправленными радиоканалами. Маршрутизация выполняется из узла 1 (источник) в узел 10 (получатель). Топология сети показана на рис. 3, возможные параметры каналов приведены в табл. 5.

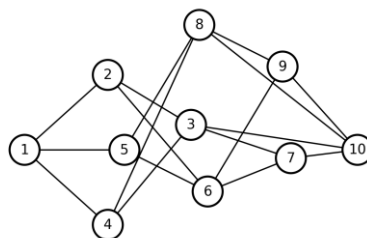


Рис. 3. Топология примерной mesh-сети

Fig. 3. Sample mesh network topology

Таблица 5
 Параметры радиоканалов (QoS) для примерной mesh-сети

Table 5
 Radio channel parameters (QoS) for a sample mesh network

Ребро Edge	Y , Mbps	D , ms	J , ms	Z
1–2	12	10	3	0,02
1–4	8	8	2	0,005
1–5	7	12	3	0,01
2–3	10	12	4	0,02

Окончание табл. 5

End of table 5

Ребро <i>Edge</i>	Y , Mbps	D , ms	J , ms	Z
2–6	9	15	4	0,02
3–4	9	9	3	0,01
3–7	7	9	3	0,02
3–10	9	15	5	0,03
4–8	8	14	5	0,01
5–6	7	10	3	0,01
5–8	5	20	8	0,03
6–7	6	9	2	0,01
6–9	8	12	4	0,02
7–10	6	11	3	0,01
8–9	9	7	2	0,015
8–10	8	18	6	0,01
9–10	8	10	3	0,02

Запасы энергии узлов рассматриваемой сети приведены в табл. 6. Для трех профилей трафика задаются усредненные мощности P_k , характеризующие энергопотребление узла при участии в передаче: $P_{\text{файлы}}=0,8$ Вт, $P_{\text{голос}}=0,7$ Вт, $P_{\text{телеметрия}}=0,5$ Вт. Время жизни узла рассчитывается как $T_{v,k}=E_v/P_k$.

Таблица 6

Запасы энергии узлов и оценка времени жизни T для профилей трафика

Table 6

Energy reserves of nodes and lifetime estimate T for traffic profiles

Узел <i>Edge</i>	E , энергия, Вт·ч <i>Energy,</i> <i>W·hrs</i>	T , файлы, ч <i>Files (FTP),</i> <i>hrs</i>	T , голос, ч <i>Voice (VoIP),</i> <i>hrs</i>	T , телеметрия, ч <i>Telemetry (IoT),</i> <i>hrs</i>
1	20	25	28,57	40
2	18	22,5	25,71	36
3	5	6,25	7,14	10
4	16	20	22,86	32
5	14	17,5	20	28
6	12	15	17,14	24
7	13	16,25	18,57	26
8	22	27,5	31,43	44
9	17	21,25	24,29	34
10	24	30	34,29	48

Ограничения QoS и начальные веса критериев для профилей трафика приведены в табл. 7. В дальнейшем для каждого профиля запускается итерационный алгоритм настройки весов, описанный выше.

Таблица 7
Ограничения QoS и начальные веса критериев для профилей трафика

Table 7
QoS constraints and initial criteria weights for traffic profiles

Профиль Profile	Y_{min}	D_{max}	J_{max}	Z_{max}	T_{min}	w_Y	w_D	w_J	w_X	w_T
Передача файлов File Transfer (FTP)	6	80	30	0,15	8	0,5	0,15	0,1	0,1	0,15
Голос Voice (VoIP)	1	50	10	0,05	6	0,05	0,35	0,25	0,25	0,1
Телеметрия Telemetry (IoT)	0,5	120	40	0,1	18	0,1	0,15	0,05	0,2	0,5

Итоговые маршруты, полученные для различных типов трафика, приведены в табл. 8. Для профиля «Передача файлов (FTP)» требуется итерационная корректировка весов из-за нарушения ограничения по времени жизни маршрута при прохождении через узел 3 с малым запасом энергии.

Таблица 8
Результат маршрутизации: выбранные маршруты и показатели QoS

Table 8
Routing result: selected routes and QoS indicators

Профиль Profile	Оптимальный маршрут Optimal Route	Y_p , Mbps	D_p , ms	J_p , ms	Z_p	T_p , hrs	R_p	Итерации Iterations
Передача файлов File transfer (FTP)	1–4–8–10	8	40	13	0,025	20	0,3	2
Голос Voice (VoIP)	1–4–3–10	8	32	10	0,045	7,14	0,25	0
Телеметрия Telemetry (IoT)	1–4–8–10	8	40	13	0,025	32	0,28	0

Ход итерационного процесса для профиля «Передача файлов (FTP)» приведен в табл. 9. На нулевой итерации выбирается маршрут 1–2–3–10, обладающий высокой полосой пропускания и малой задержкой, однако ограничение $T_p \geq T_{min}$ не выполняется из-за низкой энергии узла 3. После увеличения веса w_T алгоритм выбирает обходной маршрут 1–4–8–10, удовлетворяющий всем ограничениям QoS.

Таблица 9
Итерационная корректировка весов для профиля «Передача файлов (FTP)»

Table 9
Iterative adjustment of weights for the “File Transfer (FTP)” profile

Итерация <i>Iteration</i>	Весы/Weights $w_Y, w_D, w_J, w_X,$ w_T	Маршрут <i>Route</i>	$Y_p,$ Mbps	$D_p,$ ms	$J_p,$ ms	Z_p	$T_p,$ hrs	R_p	Нарушения <i>Infractions</i>
0	0,500, 0,150, 0,100, 0,100, 0,150	1–2–3–10	9	37	2	0,068	6,25	0,33	Т
1	0,451, 0,135, 0,090, 0,090, 0,234	1–2–3–10	9	37	2	0,068	6,25	0,3	Т
2	0,401, 0,120, 0,080, 0,080, 0,318	1–4–8–10	8	40	3	0,025	20	0,3	–

Таким образом, показано, что итерационная настройка весовых коэффициентов в сочетании с минимаксным критерием позволяет исключить ложные решения об отсутствии QoS-осуществимого маршрута и получить приемлемое решение для разных профилей трафика.

3. Двухуровневая маршрутизация в кластеризованной сети

Для масштабируемых беспроводных сенсорных mesh-сетей (WSN-mesh), применяемых в задачах умного города и инфраструктурного мониторинга, характерны большие размеры сети, неоднородность трафика и ограниченность энергетических ресурсов узлов. Одним из практических подходов к повышению масштабируемости и управляемости является кластеризация: узлы объединяются в кластеры, внутри которых выполняется локальная маршрутизация и (при необходимости) агрегация данных, а обмен между кластерами организуется через граничные узлы (шлюзы). Ниже предложен алгоритм двухуровневой маршрутизации, увязывающий кластеризацию с разработанным в разд. 2 алгоритмом многокритериальной QoS-маршрутизации, включающим параметр времени жизни маршрута T .

Рассмотрим сеть, в которой множество узлов V разделено на непересекающиеся кластеры $V=V_1 \cup V_2 \cup \dots \cup V_C$. В каждом кластере выделяется голова кластера (Cluster Head, CH), обеспечивающая локальную координацию (например, сбор метрик, распределение расписаний и т. п.), а также один или несколько граничных узлов (Gateway, GW), имеющих радиоканалы с узлами соседних кластеров. Маршрутизация внутри кластера осуществляется по внутрикластерным радиоканалам, а межкластерная – по радиоканалам, соединяющим шлюзы разных кластеров. Качество маршрута на обоих уровнях оценивается по тем же параметрам QoS (Y, D, J, Z) и времени жизни маршрута T , что и в предыдущих подразделах.

Двухуровневый подход состоит в разделении вычислений на два уровня: локальные вычисления внутри каждого кластера и глобальную маршрутизацию между кластерами на основе редуцированного (укрупненного) графа. В каждом кластере определяется множество опорных узлов S_c , включающее СН и все GW данного кластера, а также узлы источника s и назначения t , если они принадлежат кластеру. Для каждой пары опорных узлов $(u, v) \in S_c \times S_c$ ($u \neq v$) рассчитывается оптимальный внутрикластерный маршрут $p_c^*(u, v)$ с использованием разработанного алгоритма из разд. 2 с учетом ограничений QoS и параметра времени жизни маршрута T . Параметры виртуального «канала» $u \rightarrow v$ внутри кластера принимаются равными параметрам маршрута $p_c^*(u, v)$, вычисляемым по правилам (11) (минимум по Y и T , суммирование по D, J и X). На основе внутрикластерных виртуальных каналов и межкластерных радиоканалов строится редуцированный граф $G^*=(V^*, E^*)$, где V^* – объединение всех опорных узлов, E^* включает межкластерные ребра между шлюзами разных кластеров и виртуальные внутрикластерные ребра. Оптимальный маршрут из s в t затем ищется уже на графе G^* тем же алгоритмом, после чего выполняется разворачивание решения: каждое виртуальное ребро заменяется соответствующим внутрикластерным маршрутом $p_c^*(u, v)$.

Алгоритм двухуровневой маршрутизации для профиля трафика k может быть задан следующими шагами:

Вход. Определить кластеризацию сети: для каждого узла v задать номер кластера $cluster(v)$, выделить головы кластеров CH_c и граничные узлы GW_c .

Шаг 1. Сформировать множества опорных узлов S_c для каждого кластера: $S_c = \{CH_c\} \cup GW_c \cup \{s, t\} \cap V_c$.

Шаг 2. Уровень 1 (внутрикластерная маршрутизация): для каждого кластера c и каждой пары (u, v) из S_c вычислить оптимальный маршрут $p_c^*(u, v)$ алгоритмом из разд. 2 с учетом ограничений QoS и времени жизни маршрута T ; сформировать виртуальные ребра $u \rightarrow v$ с параметрами (Y, D, J, Z, T) и сохранить $p_c^*(u, v)$ для последующего разворачивания.

Шаг 3. Построить редуцированный граф G^* : $V^* = \cup_c S_c$; $E^* = E_{inter} \cup E_{virtual}$, где E_{inter} – межкластерные радиоканалы между GW разных кластеров, $E_{virtual}$ – набор виртуальных внутрикластерных ребер между опорными узлами.

Шаг 4. Уровень 2 (межкластерная маршрутизация): выполнить поиск оптимального маршрута $p^*(s, t)$ на графе G^* алгоритмом из разд. 2 (в том числе с использованием минимаксного критерия).

Выход. Развернуть найденный маршрут: заменить каждое виртуальное ребро $u \rightarrow v$ фактическим внутрикластерным маршрутом $p_c^*(u, v)$, получая полный маршрут в исходном графе.

Преимуществами двухуровневого подхода являются снижение вычислительной сложности и локализация перерасчета маршрутов при изменении состояния сети: при изменении качества радиоканалов или энергосостояния узлов перерассчитываются только виртуальные ребра затронутых кластеров и (или) межкластерные связи, после чего выполняется перерасчет маршрута на редуцированном графе G^* .

В качестве примера рассмотрим кластеризованную беспроводную mesh-сеть из 20 узлов (1–20), разделенную на четыре кластера (K1–K4) по пять узлов в каждом. Источник трафика расположен в узле 2 (кластер K1), получатель – узле 19 (кластер K4).

Топология сети показана на рис. 4, состав кластеров и роли узлов приведены в табл. 9, параметры радиоканалов – в табл. 10, 11, энергетические характеристики – в табл. 12. Ограничения QoS и веса интегрального критерия заданы в табл. 13.

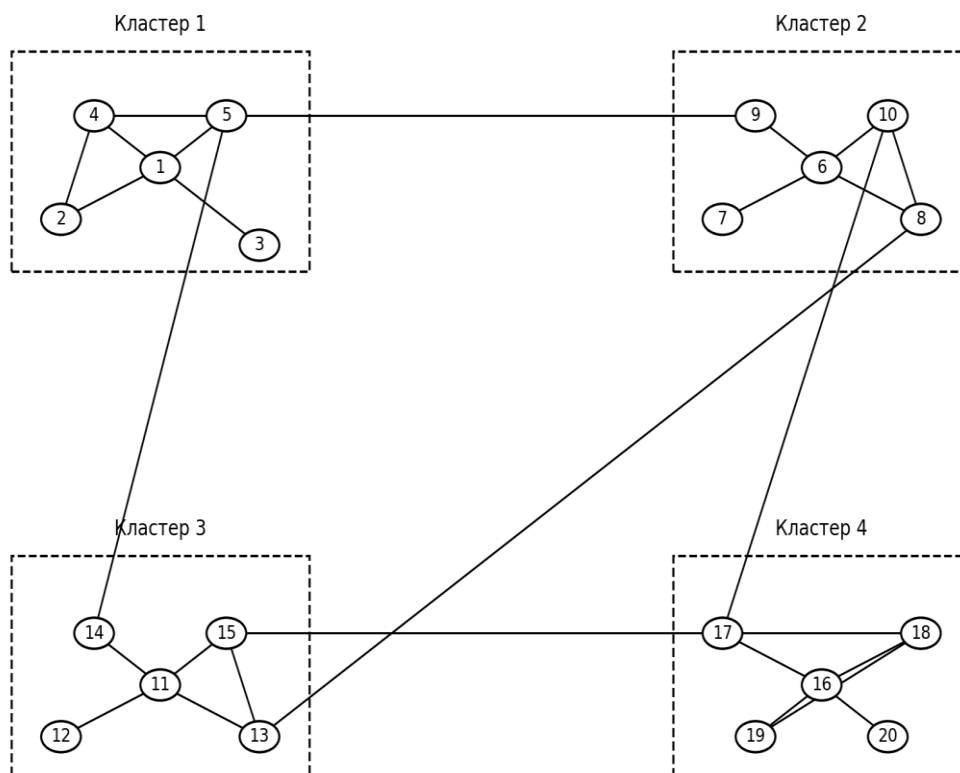


Рис. 4. Кластеризованная mesh-сеть (20 узлов, четыре кластера)

Fig. 4. Clustered mesh network (20 nodes, 4 clusters)

Таблица 9

Состав кластеров и роли узлов

Table 9

Composition of clusters and roles of nodes

Кластер <i>Cluster</i>	Узлы <i>Nodes</i>	Голова кластера <i>Cluster Head</i>	Граничные узлы (шлюзы) <i>Border nodes (gateways)</i>	Комментарий <i>Comment</i>
К1	1, 2, 3, 4, 5	1	4, 5	Шлюзы: 4↔9 (К2), 5↔14 (К3)
К2	6, 7, 8, 9, 10	6	8, 9, 10	Шлюзы: 9↔4 (К1), 8↔13 (К3), 10↔17 (К4)
К3	11, 12, 13, 14, 15	11	13, 14, 15	Шлюзы: 14↔5 (К1), 13↔8 (К2), 15↔18 (К4)
К4	16, 17, 18, 19, 20	16	17, 18	Шлюзы: 17↔10 (К2), 18↔15 (К3)

Таблица 10
 Параметры QoS внутри кластеров

Table 10
QoS parameters within clusters

Кластер <i>Cluster</i>	Ребро <i>Edge</i>	<i>Y</i> , Mbps	<i>D</i> , ms	<i>J</i> , ms	<i>Z</i>
К1	1–2	12	6	1	0,005
К1	1–3	10	7	1	0,010
К1	1–4	9	6	1	0,005
К1	1–5	15	10	2	0,010
К1	2–4	8	5	1	0,005
К2	6–10	9	8	1	0,010
К2	6–7	8	9	2	0,020
К2	6–8	11	7	1	0,010
К2	6–9	10	6	1	0,005
К2	8–10	12	6	1	0,005
К3	11–12	8	8	1	0,010
К3	11–13	12	6	1	0,005
К3	11–14	15	7	1	0,010
К3	11–15	14	7	1	0,005
К3	13–15	16	5	1	0,005
К4	16–17	10	6	1	0,005
К4	16–18	12	6	1	0,005
К4	16–19	11	7	1	0,005
К4	16–20	14	8	2	0,010
К4	18–19	13	5	1	0,005

Таблица 11
 Параметры QoS между кластерами

Table 11
QoS parameters between clusters

Связь <i>Connection</i>	Ребро <i>Edge</i>	<i>Y</i> , Mbps	<i>D</i> , ms	<i>J</i> , ms	<i>Z</i>
К1↔К2	4–9	9	8	1	0,005
К1↔К3	5–14	18	18	2	0,010
К2↔К3	8–13	10	14	2	0,020
К2↔К4	10–17	8	10	1	0,010
К3↔К4	15–18	12	9	1	0,005

Таблица 12
Запасы энергии узлов и оценка времени жизни
для профилей трафика

Table 12
Node energy reserves and lifetime estimation for traffic
profiles

Узел Node	E , W·hr	$T(FTP)$, hr	$T(VoIP)$, hr	$T(IoT)$, hr
1	20	25,0	28,57	40,0
2	18	22,5	25,71	36,0
3	16	20,0	22,86	32,0
4	15	18,75	21,43	30,0
5	14	17,5	20,0	28,0
6	22	27,5	31,43	44,0
7	8	10,0	11,43	16,0
8	18	22,5	25,71	36,0
9	20	25,0	28,57	40,0
10	6	7,5	8,57	12,0
11	24	30,0	34,29	48,0
12	20	25,0	28,57	40,0
13	18	22,5	25,71	36,0
14	16	20,0	22,86	32,0
15	15	18,75	21,43	30,0
16	20	25,0	28,57	40,0
17	18	22,5	25,71	36,0
18	18	22,5	25,71	36,0
19	20	25,0	28,57	40,0
20	22	27,5	31,43	44,0

Таблица 13
Ограничения QoS и веса интегрального критерия

Table 13
QoS limitations and integral criterion weights

Профиль Profile	Y_{min}	D_{max}	J_{max}	Z_{max}	T_{min}	w_Y	w_D	w_J	w_X	w_T
Передача файлов File Transfer (FTP)	6	80	30	0,15	8	0,5	0,15	0,1	0,1	0,15
Голос Voice (VoIP)	1	50	10	0,05	6	0,05	0,35	0,25	0,25	0,1
Телеметрия Telemetry (IoT)	0,5	120	40	0,10	18	0,1	0,15	0,05	0,2	0,5

В результате реализации данного алгоритма для рассматриваемой сети получены опорные внутрикластерные маршруты (табл. 14) и полные маршруты – результаты двух-уровневой маршрутизации (табл. 15).

Таблица 14

Опорные внутрикластерные маршруты (виртуальные каналы) между ключевыми узлами

Table 14

Backbone intra-cluster routes (virtual channels) between key nodes

Виртуальный канал (маршрут внутри кластера) <i>Virtual channel (route inside the cluster)</i>	Профиль <i>Profile</i>	Y , Mbps	D , ms	J , ms	Z	T , hr
2-4 (K1)	FTP	8	5	1	0,005	18,75
2-4 (K1)	VoIP	8	5	1	0,005	21,43
2-4 (K1)	IoT	8	5	1	0,005	30,0
2-1-5 (K1)	FTP	12	16	3	0,015	17,5
2-1-5 (K1)	VoIP	12	16	3	0,015	20,0
2-1-5 (K1)	IoT	12	16	3	0,015	28,0
9-6-10 (K2)	FTP	9	14	2	0,015	7,5
9-6-10 (K2)	VoIP	9	14	2	0,015	8,57
9-6-10 (K2)	IoT	9	14	2	0,015	12,0
14-11-15 (K3)	FTP	14	14	2	0,015	18,75
14-11-15 (K3)	VoIP	14	14	2	0,015	21,43
14-11-15 (K3)	IoT	14	14	2	0,015	30,0
17-16-19 (K4)	FTP	10	13	2	0,01	22,5
17-16-19 (K4)	VoIP	10	13	2	0,01	25,71
17-16-19 (K4)	IoT	10	13	2	0,01	36,0
18-19 (K4)	FTP	13	5	1	0,005	22,5
18-19 (K4)	VoIP	13	5	1	0,005	25,71
18-19 (K4)	IoT	13	5	1	0,005	36,0

Таблица 15

Результаты двухуровневой маршрутизации: выбранные маршруты и показатели QoS

Table 15

Two-tier routing results: selected routes and QoS metrics

Профиль <i>Profile</i>	Кластерный маршрут <i>Cluster Route</i>	Маршрут (узлы) <i>Route (for nodes)</i>	Y_p , Mbps	D_p , ms	J_p , ms	Z_p	T_p , hr	R
Голос <i>Voice (VoIP)</i>	K1→K2→K4	2-4-9-6-10-17-16-19	8	50	7	0,044	8,57	0,8214
Передача файлов <i>File Transfer (FTP)</i>	K1→K3→K4	2-1-5-14-11-15-18-19	12	62	9	0,049	17,5	0,4957
Телеметрия <i>Telemetry (IoT)</i>	K1→K3→K4	2-1-5-14-11-15-18-19	12	62	9	0,049	28,0	0,5096

Анализ примера показывает, что двухуровневый алгоритм обеспечивает согласование кластеризации с многокритериальной маршрутизацией. Для профиля «Голос (VoIP)» выбирается кластерный маршрут $K1 \rightarrow K2 \rightarrow K4$ за счет меньшей суммарной задержки и вариации задержки, а маршрут через $K3$ не удовлетворяет ограничению D_{max} . Для профилей «Передача файлов (FTP)» и «Телеметрия (IoT)» маршрут через кластер $K2$ становится неосуществимым по ограничению времени жизни (узел 10 имеет малый запас энергии: $T(FTP) = 7,50 \text{ ч} < T_{min} = 8 \text{ ч}$ и $T(IoT) = 12 \text{ ч} < T_{min} = 18 \text{ ч}$), поэтому выбирается обходной кластерный маршрут $K1 \rightarrow K3 \rightarrow K4$. Таким образом, кластеризация уменьшает размерность задачи на глобальном уровне, а введение параметра T обеспечивает энергоосведомленный выбор маршрута как внутри кластеров, так и при межкластерной маршрутизации.

Заключение

Предложенная модификация алгоритма Дейкстры с расширенной меткой позволяет решать задачу многокритериальной QoS-маршрутизации в mesh-сетях с учетом неаддитивных показателей (минимальная полоса пропускания и время жизни маршрута). При этом предложенный алгоритм дает возможность учитывать энергетические критерии, например время жизни маршрута T , обеспечивая энергоосведомленный выбор маршрута и позволяя избегать путей, проходящих через узлы с ограниченными запасами энергии, что критично для сенсорных сетей умного города и инфраструктурного мониторинга. Разработан итерационный алгоритм настройки весовых коэффициентов, который в сочетании с минимаксным критерием позволяет исключать ложные решения об отсутствии QoS-осуществимого маршрута и получать приемлемые решения для разных профилей трафика. Разработаны двухуровневый подход и алгоритм маршрутизации в кластеризованной сети, обеспечивающие снижение вычислительной сложности и локализацию перерасчета маршрутов при изменении состояния сети: качества радиоканалов или энергосостояния узлов.

Вклад авторов. *А. К. Шульган* осуществил сбор и анализ результатов работы, написал текст статьи, внес значительный вклад в концепцию. *Ю. И. Воротницкий* осуществил сбор, анализ и оформление результатов работ и внес идею написания статьи.

Список использованных источников

1. Al-Karaki, J. N. Routing techniques in wireless sensor networks: a survey / J. N. Al-Karaki, A. E. Kamal // IEEE Wireless Communications. – 2004. – Vol. 11, no. 6. – P. 6–28.
2. Akkaya, K. A survey on routing protocols for wireless sensor networks / K. Akkaya, M. Younis // Ad Hoc Networks. – 2005. – Vol. 3, no. 3. – P. 325–349.
3. Pantazis, N. A. Energy-efficient routing protocols in wireless sensor networks: a survey / N. A. Pantazis, S. A. Nikolidakis, D. D. Vergados // IEEE Communications Surveys & Tutorials. – 2013. – Vol. 15, no. 2. – P. 551–591.
4. LEA-RPL: lightweight energy-aware RPL protocol for internet of things based on particle swarm optimization / S. Mokrani, M. Belkadi, T. Sadoun [et al.] // Telecommunication Systems. – 2025. – Vol. 88, iss. 1, art. 14. – P. 433–454.
5. Alotaibi, M. A. A novel reliable and trust objective function for RPL-based IoT routing protocol / M. A. Alotaibi, S. S. Alwakeel, A. N. Alyahya // Computers, Materials and Continua. – 2025. – Vol. 82, iss. 2. – P. 3467–3497.
6. Листопад, Н. И. Маршрутизация в мультисервисных сетях телекоммуникаций на основе модифицированного алгоритма Дейкстры / Н. И. Листопад, Ю. И. Воротницкий, А. А. Хайдер // Вестник БГУ. Серия 1. – 2015. – № 1. – С. 70–76.
7. Многокритериальная маршрутизация информационных потоков / Н. И. Листопад, Ю. И. Воротницкий, В. В. Бортновский, А. А. Хайдер // Проблемы физики, математики и техники. – 2017. – № 2(31). – С. 84–90.

8. A scalable approach to QoS-aware self-adaptation in service-oriented architectures / V. Cardellini, E. Casalicchio, V. Grassi [et al.] // Proc. of 6th Intern. ICST Conf. on Heterogeneous Networking for Quality, Reliability, Security and Robustness, Q Shine 2009 and 3rd Intern. Workshop on Advanced Architectures and Algorithms for Internet Delivery and Applications, AAA-IDEA 2009, Las Palmas, Gran Canaria, 23–25 Nov. 2009. – Berlin, Heidelberg, N. Y., 2009. – P. 431–447.
9. Mahmoud, W. A. A proposal algorithm to solve delay constraint least cost optimization problem / W. A. Mahmoud, D. J. Kadhim // Journal of Engineering. University of Baghdad. – 2013. – Vol. 19, no. 1. – P. 155–160.
10. Dijkstra, E. W. A note on two problems in connexion with graphs / E. W. Dijkstra // Numerische Mathematik. – 1959. – Vol. 1. – P. 269–271.

References

1. Al-Karak J. N., Kamal A. E. Routing techniques in wireless sensor networks: a survey. *IEEE Wireless Communications*, 2004, vol. 11, no. 6, pp. 6–28.
2. Akkaya K., Younis M. A survey on routing protocols for wireless sensor networks. *Ad Hoc Networks*, 2005, vol. 3, no. 3, pp. 325–349.
3. Pantazis N. A., Nikolidakis S. A., Vergados D. D. Energy-efficient routing protocols in wireless sensor networks: a survey. *IEEE Communications Surveys & Tutorials*, 2013, vol. 15, no. 2, pp. 551–591.
4. Mokrani S., Belkadi M., Sadoun T., Lloret J., Aoudjit R. LEA-RPL: lightweight energy-aware RPL protocol for internet of things based on particle swarm optimization. *Telecommunication Systems*, 2025, vol. 88, iss. 1, art. 14, pp. 433–454.
5. Alotaibi M. A., Alwakeel S. S., Alyahya A. N. A novel reliable and trust objective function for RPL-based IoT routing protocol. *Computers, Materials and Continua*, 2025, vol. 82, iss. 2, pp. 3467–3497.
6. Listopad N. I., Vorotnitsky Y. I., Hayder A. A. *Routing in multiservice telecommunication networks based on a modified Dijkstra algorithm*. Vestnik Belorusskogo gosudarstvennogo universiteta. Serija 1 [Journal of the Belarusian State University. Series I], 2015, no. 1, pp. 70–76 (In Russ.).
7. Listopad N. I., Vorotnitsky Y. I., Bortnovsky V. V., Hayder A. A. *Multi-criterial routing of information flows*. Problemy fiziki, matematiki i tehniki [Problems of Physics, Mathematics and Technics], 2017, no. 2(31), pp. 84–90 (In Russ.).
8. Cardellini V., Casalicchio E., Grassi V., Presti F. L., Mirandola R. A scalable approach to QoS-aware self-adaptation in service-oriented architectures. *Proceedings of 6th International ICST Conference on Heterogeneous Networking for Quality, Reliability, Security and Robustness, Q Shine 2009 and 3rd International Workshop on Advanced Architectures and Algorithms for Internet Delivery and Applications, AAA-IDEA 2009, Las Palmas, Gran Canaria, 23–25 November 2009*. Berlin, Heidelberg, New York, 2009, pp. 431–447.
9. Mahmoud W. A., Kadhim D. J. A proposal algorithm to solve delay constraint least cost optimization problem. *Journal of Engineering. University of Baghdad*, 2013, vol. 19, no. 1, pp. 155–160.
10. Dijkstra E. W. A note on two problems in connexion with graphs. *Numerische Mathematik*, 1959, vol. 1, pp. 269–271.

Информация об авторах

Шульган Антон Константинович, аспирант, Белорусский государственный университет.
E-mail: anton.shulgan@mail.ru

Воротницкий Юрий Иосифович, кандидат физико-математических наук, доцент, заведующий кафедрой, Белорусский государственный университет.
E-mail: vorotn@bsu.by

Information about the authors

Anton K. Shulgan, Postgraduate Student, Belarusian State University.
E-mail: anton.shulgan@mail.ru

Yury I. Vorotnitsky, Cand. Sci. (Phys.-Math.), Assoc. Prof., Head of Department, Belarusian State University.
E-mail: vorotn@bsu.by

УДК 004.42+004.383.3
<https://doi.org/10.37661/1816-0301-2026-23-1-58-68>

Поступила в редакцию | Received 16.02.2026
Подписана в печать | Accepted 06.03.2026
Опубликована | Published 31.03.2026

Метод повышения пространственного разрешения амбисонического звука на основе разреженного МДКП-представления

Д. С. Лихачёв[✉], Н. А. Петровский, И. С. Азаров
[✉]E-mail: likhachov@bsuir.by

*Белорусский государственный университет
информатики и радиоэлектроники,
ул. П. Бровки, 6, Минск, 220013, Беларусь*

Аннотация

Цели. Решается задача разработки метода повышения пространственного разрешения амбисонического звука на основе разреженного МДКП-представления (МДКП – модифицированное дискретное косинусное преобразование), позволяющего улучшить эффективность существующих подходов представления пространственного звука в цифровом виде.

Методы. Предлагаемый подход основан на временно-частотной декомпозиции аудиосигнала с последующим извлечением направленной плоской волны из каждой частотной составляющей. Данный подход развивает основные идеи методов разложения на плоские волны с высоким разрешением (HARPEX) и аудиокодирования с учетом направления прихода звука, используя преимущества вещественной разреженной декомпозиции.

Результаты. Метод повышения пространственного разрешения амбисонического звука предполагает использование вещественных частотных компонентов, что, в отличие от комплексных, обеспечивает более простую и устойчивую оценку направления прихода звука. Разреженная декомпозиция позволяет реализовать точный и унифицированный подход к описанию звуков различной природы – от транзиентных до тоновых.

Заключение. Практические результаты подтверждают применимость предложенного метода для обработки аудио с повышением разрешения вплоть до амбисоники седьмого порядка. Основным недостатком метода является его сравнительно с другими методами высокая вычислительная сложность, что, однако, некритично для приложений, которые не требуют обработки в реальном времени.

Ключевые слова: пространственный звук, амбисоника, апмиксинг, пространственное разрешение, разреженное представление, быстрое преобразование Фурье, модифицированное дискретное косинусное преобразование

Для цитирования. Лихачёв, Д. С. Метод повышения пространственного разрешения амбисонического звука на основе разреженного МДКП-представления / Д. С. Лихачёв, Н. А. Петровский, И. С. Азаров // Информатика. – 2026. – Т. 23, № 1. – С. 58–68. – <https://doi.org/10.37661/1816-0301-2026-23-1-58-68>.

Конфликт интересов. Авторы заявляют об отсутствии конфликта интересов.

Method for spatial resolution enhancement of ambisonic audio based on sparse MDCT representation

Denis S. Likhachov[✉], Nick A. Petrovsky, Elias S. Azarov

[✉]E-mail: likhachov@bsuir.by

*Belarusian State University
of Informatics and Radioelectronics,
st. P. Brovki, 6, Minsk, 220013, Belarus*

Abstract

Objectives. The problem of developing a method for spatial resolution enhancement of ambisonic audio which improves the efficiency of existing approaches to digital spatial sound representation is being solved.

Methods. The proposed approach is based on time-frequency decomposition of the audio signal with subsequent extraction of a directional plane wave from each frequency component. The approach develops the core ideas of high-resolution plane wave expansion (HARPEX) and directional audio coding (DirAC) methods, utilizing the advantages of real-valued sparse decomposition.

Results. Application of the proposed ambisonic spatial resolution enhancement method involves the use of real-valued frequency components, which, unlike complex ones, provide a simpler and more robust estimation of sound arrival direction. Sparse decomposition enables an accurate and unified approach to describing sounds of various nature – from transient to tonal.

Conclusion. Practical results confirm the method's applicability for audio processing with resolution enhancement up to seventh-order ambisonics. The disadvantage of the method is its high computational complexity; however, it is suitable for applications that do not require real-time processing.

Keywords: spatial audio, ambisonics, upmixing, spatial resolution, sparse representation, fast Fourier transform, modified discrete cosine transform

For citation. Likhachov D. S., Petrovsky N. A., Azarov E. S. *Method for spatial resolution enhancement of ambisonic audio based on sparse MDCT representation*. *Informatika [Informatics]*, 2026, vol. 23, no. 1, pp. 58–68 (In Russ.). <https://doi.org/10.37661/1816-0301-2026-23-1-58-68>.

Conflict of interests. The authors declare of no conflict of interest.

Введение

В настоящее время амбисоника первого порядка является популярным и доступным способом записи пространственного звука. Существует множество удобных портативных устройств, обеспечивающих четырехканальную запись в формате амбисоники, которая может воспроизводиться в приложениях виртуальной реальности как через колонки, так и через наушники. Качество воспроизведения сильно зависит от пространственного разрешения, которое для амбисоники первого порядка относительно низкое. Существуют параметрические методы, позволяющие улучшить пространственный образ и хорошо зарекомендовавшие себя на практике [1, 2]. Общая идея параметрического подхода заключается в выполнении временно-частотной декомпозиции аудиосигнала и обработке каждой частотной составляющей отдельно, представляя ее в виде комбинации направленных плоских волн и (или) ненаправленных компонентов. Оценка направления прихода звука формирует четкий пространственный образ, который может быть закодирован в формате амбисоники более высокого порядка (рис. 1).

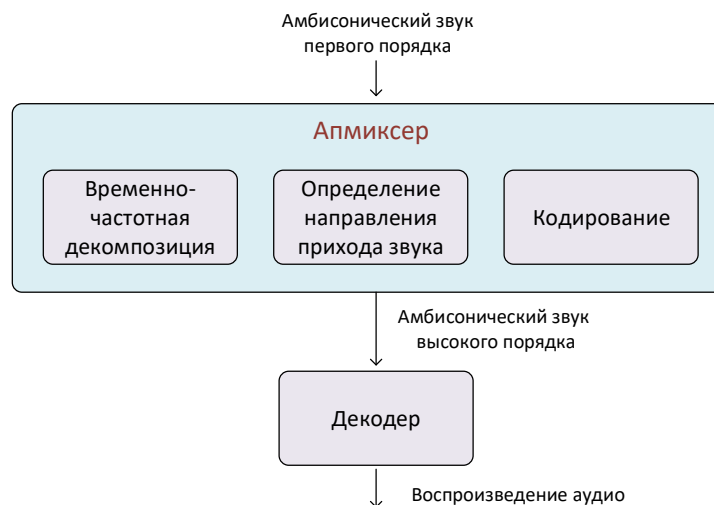


Рис. 1. Схема воспроизведения амбисонического аудио с повышенным пространственным разрешением

Fig. 1. Ambisonic audio playback scheme with improved spatial resolution

Традиционно в задачах подавления шума, уменьшения реверберации и обработки многоканального аудио для частотного представления сигнала используется быстрое преобразование Фурье (БПФ) [3–6]. Преобразование Фурье обеспечивает комплексное представление сигнала, что не всегда удобно применять на практике при обработке аудиосигналов.

В настоящей работе предполагается, что вещественные частотные представления более эффективны для задачи определения направления прихода звука. Так, например, вещественное МДКП не только обеспечивает более интуитивную интерпретацию направленной волны, но и делает решение более устойчивым. Применение МДКП исключает неопределенные состояния, характерные для метода HARPEX [2], который работает с комплексными экспонентами.

Помимо перехода из комплексной частотной области в вещественную предлагается использовать избыточные базисы различной длины для представления частотных составляющих. Основная мотивация заключается в возможности точной обработки как транзитных компонентов, т. е. очень коротких импульсов, так и тоновых компонентов, т. е. длительных периодических сигналов. При использовании линейных преобразований с базисными функциями одинаковой длины невозможно достичь одинаковой точности представления всех компонентов. Как правило, на практике делается компромиссный выбор в пользу тоновых сигналов, что приводит к размытию транзитных составляющих и значительному снижению пространственного разрешения. В данной работе для выполнения более точной временно-частотной декомпозиции применяется оптимизационная процедура, минимизирующая $L1$ -норму и обеспечивающая разреженное представление в пространстве избыточных МДКП базисов различной длины. Субъективные тесты подтверждают высокий потенциал предложенного решения, несмотря на его высокую вычислительную сложность.

Метод повышения пространственного разрешения амбисонического звука HARPEX

HARPEX раскладывает четырехканальное амбисоническое аудио, т. е. амбисонику первого порядка, на перекрывающиеся фреймы, которые преобразуются в частотную область с помощью БПФ [2]. Это позволяет получать комплексные спектральные отсчеты вида $w_r + iw_i$, $x_r + ix_i$, $y_r + iy_i$ и $z_r + iz_i$ для соответствующих каналов, обозначенных как W, X, Y, Z соответственно. Для простоты в данном случае далее рассматривается только одна частота, предполагая, что обработка всех частот выполняется одинаково. Четыре комплексных значения одной частоты декомпозируются в два представляющие направление прихода звука вещественных вектора $[x_1, y_1, z_1]$ и $[x_2, y_2, z_2]$ и комплексные амплитуды a_1, a_2 двух звуковых волн, соответствующих этим направлениям. Таким образом, каждая частота представляется в виде суммы двух плоских волн, каждая из которых имеет свое направление прихода. Декомпозиция описывается матричным уравнением

$$\begin{bmatrix} w_r + iw_i \\ x_r + ix_i \\ y_r + iy_i \\ z_r + iz_i \end{bmatrix} = a_1 \begin{bmatrix} 2^{-\frac{1}{2}} \\ x_1 \\ y_1 \\ z_1 \end{bmatrix} + a_2 \begin{bmatrix} 2^{-\frac{1}{2}} \\ x_2 \\ y_2 \\ z_2 \end{bmatrix}. \quad (1)$$

Решение данной системы может быть представлено в следующем виде:

$$\begin{bmatrix} a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} m_1 & 0 \\ 0 & m_2 \end{bmatrix} \begin{bmatrix} c_1 & s_1 \\ c_2 & s_2 \end{bmatrix} \begin{bmatrix} 1 \\ i \end{bmatrix}, \quad (2)$$

$$c_{1,2} = \sqrt{\frac{2r^2 - pq + p^2 \pm 2r\sqrt{r^2 - pq}}{(q - p)^2 + 4r^2}}, \quad (3)$$

$$s_{1,2} = \frac{(q - p)c_{1,2} + p/c_{1,2}}{2r}, \quad (4)$$

$$r = -2w_r w_i + x_r x_i + y_r y_i + z_r z_i, \quad (5)$$

$$p = -2w_r^2 + x_r^2 + y_r^2 + z_r^2, \quad (6)$$

$$q = -2w_i^2 + x_i^2 + y_i^2 + z_i^2. \quad (7)$$

Некоторые частотные отсчеты попадают в область неопределенности при условии $r^2 - pq < 0$, однако количество таких отсчетов достаточно мало.

Метод повышения пространственного разрешения амбисонического звука на основе разреженной МДКП-декомпозиции с объединением различных временно-частотных разрешений

В отличие от HARPEX предлагается использовать МДКП вместо БПФ. МДКП предоставляет вещественные коэффициенты w_r, x_r, y_r, z_r для каждого соответствующего канала, что заметно упрощает систему. Эти значения декомпозируются в вещественный вектор $[x_1, y_1, z_1]$, представляющий направление прихода, и ненаправленную компоненту с вещественными амплитудами a_1, a_2 соответственно:

$$\begin{bmatrix} w_r \\ x_r \\ y_r \\ z_r \end{bmatrix} = a_1 \begin{bmatrix} 2^{-\frac{1}{2}} \\ x_1 \\ y_1 \\ z_1 \end{bmatrix} + a_2 \begin{bmatrix} 2^{-\frac{1}{2}} \\ 0 \\ 0 \\ 0 \end{bmatrix},$$

$$a_1 = \sqrt{x_1^2 + y_1^2 + z_1^2},$$

$$a_2 = \sqrt{2}w_r - a_1.$$

Данное решение неполно, поскольку теряется знак направления при отрицательном значении w_r . Для устранения этой проблемы вместо вектора $[x_1, y_1, z_1]$ можно использовать вектор $[\bar{x}_1, \bar{y}_1, \bar{z}_1]$, который учитывает знак w_r , и соответствующие знаочувствительные амплитуды \bar{a}_1 и \bar{a}_2 :

$$[\bar{x}_1, \bar{y}_1, \bar{z}_1] = \text{sgn}(w_r)[x_r, y_r, z_r],$$

$$\bar{a}_1 = \text{sgn}(w_r)a_1,$$

$$\bar{a}_2 = \sqrt{2}w_r - \bar{a}_1.$$

Таким образом, для МДКП решение становится значительно более простым по сравнению с БПФ: в методе HARPEX для каждой частотной составляющей извлекаются две плоские волны, тогда как в предлагаемом решении используется только одна плоская волна и ненаправленная составляющая.

Линейные временно-частотные преобразования (БПФ, МДКП и др.) с фиксированной длиной окна не могут обеспечить точную декомпозицию тоновых и транзиентных компонентов звука одновременно. Эти два крайних случая всегда приводят к размытию в частотном представлении. Выполнение разреженной декомпозиции аудио с использованием комбинации базисов МДКП различной длины позволяет решить данную проблему и добиться хорошей временной и частотной локализации как тонов, так и импульсов (рис. 2).

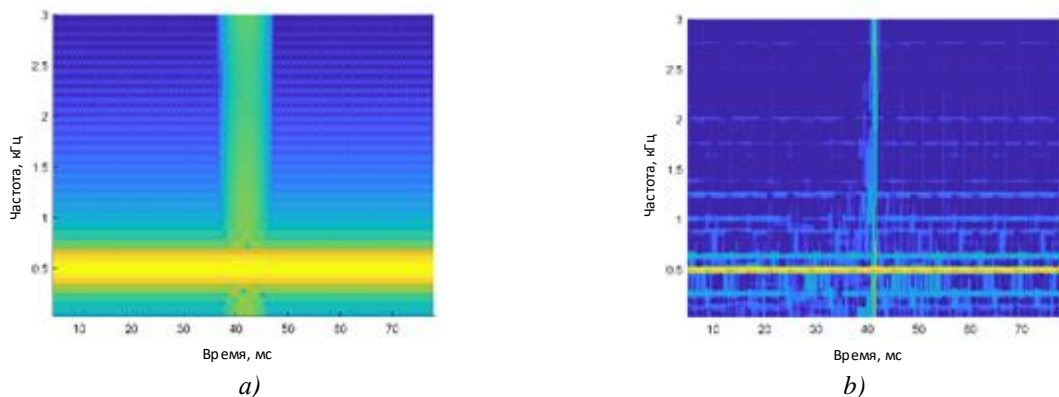


Рис. 2. Временно-частотная декомпозиция: а) БПФ (HARPEX); б) разреженная МДКП-декомпозиция
Fig. 2. Time-frequency decomposition: a) FFT (HARPEX); b) sparse MDCT decomposition

Для представления сигнала в вещественной частотной области используется объединение МДКП-базисов B , каждый из которых имеет различную длину:

$$B = \begin{bmatrix} \text{МДКП}_{32} \\ \text{МДКП}_{64} \\ \dots \\ \text{МДКП}_{2048} \end{bmatrix},$$

где МДКП_N обозначает множество базисных функций МДКП длиной N .

Таким образом, исходный сигнал можно декомпозировать на отдельные слои, обладающие различными временно-частотными характеристиками. Декомпозиция аудиосигнала сводится к следующей задаче оптимизации. Сигнал x восстанавливается при условии минимизации L1-нормы представления X :

$$\arg \min \{ \|X\|_1 \mid x = XB \},$$

где L1-норма $\|X\|_1 = \sum_i |X_i|$ представляет собой сумму абсолютных значений.

Существует множество известных подходов к решению данной задачи с минимизацией L1-нормы, включая метод согласованного преследования (matching pursuit), градиентные методы, жадные алгоритмы и линейное программирование.

Наиболее подходящим с точки зрения практического применения в рассматриваемом случае является метод градиентного спуска. Причина выбора этого метода заключается в следующем. С одной стороны, размерность задачи очень велика и не позволяет использовать сложные вычислительные методы, основанные на вторых производных или обращении матриц. С другой стороны, критически важной частью предлагаемого подхода является способность подавлять искажения, вызванные алиасингом в процессе декомпозиции. Метод градиентного спуска дает возможность добавлять в функцию потерь штрафные члены для контроля алиасинга. Его недостаток – не очень высокая общая производительность, обусловленная медленной сходимостью к решению. Данный недостаток не является критическим, поскольку в этом случае акцент ставится на лучшее качество обработки, а не на производительность технического решения.

Решатель на основе градиентного спуска для проведения экспериментов создан по аналогии с обучением нейронных сетей. Он работает итеративно, приближаясь к решению, делая шаги в направлении убывания функции потерь. Функция потерь состоит из трех компонентов: потерь реконструкции, потерь L1-нормы и потерь от наложения (алиасинга) (рис. 3).

Разреженность представления достигается за счет потерь L1-нормы, которые должны быть достаточно большими, чтобы ускорить процесс оптимизации. Однако одновременно невозможно достичь идеальной реконструкции при ненулевой потере L1-нормы. Чтобы преодолеть это, вклад каждой составляющей параметризуется с помощью параметра α , который начинается с большого значения и постепенно уменьшается на каждой итерации, в конечном итоге приближаясь к нулю.

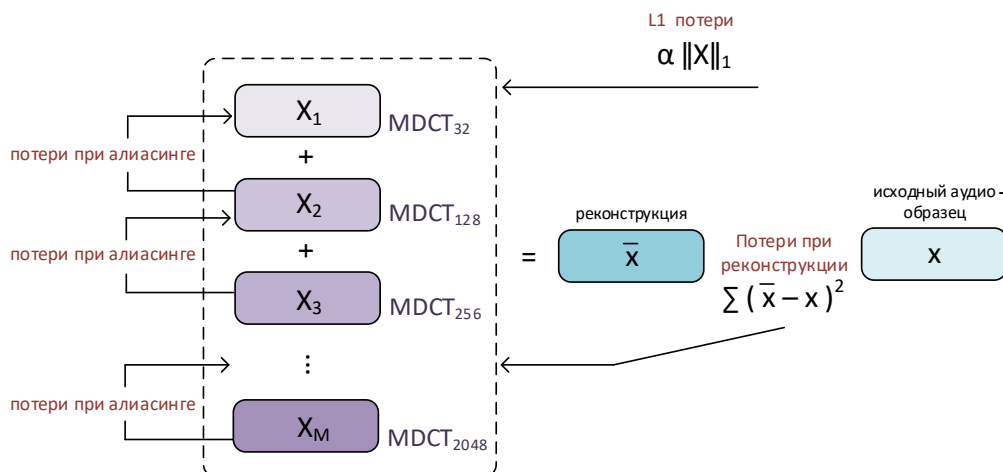


Рис. 3. Разработанный метод получения разреженного представления на основе градиентного спуска

Fig. 3. Developed method for obtaining sparse representation based on gradient descent

Механизм назначения штрафа для борьбы с алиасингом основан на выявлении искажений, возникающих в слоях с более длинными базисными функциями. Используется следующий принцип: слои с более короткими базисными функциями не должны увеличивать энергию выборок слоев с более длинными базисными функциями.

Экспериментальная оценка эффективности предложенного подхода

Оценка эффективности предложенного подхода производилась на основе субъективных оценок качества аудиообразцов, полученных при преобразовании аудио в формате амбисоники первого порядка в аудио в формате амбисоники седьмого порядка с использованием различных схем кодирования.

Обработка аудио включала следующие этапы:

- разреженную декомпозицию каждого канала в базисах МДКП;
- определение направления прихода звука;
- кодирование результата в формат амбисоники седьмого порядка.

Полученное аудио высокого порядка декодировалось в бинауральное аудио с помощью бинаурального декодера SPARTA [7].

Для разреженной декомпозиции использовались пять слоев декомпозиции с размерами МДКП 32, 128, 256, 1024 и 2048 с синусоидальным окном. Для ускорения оптимизационной процедуры в каждом слое не применялась передискретизация с повышенной частотой, хотя это могло бы дополнительно повысить качество декомпозиции. Решатель выполнял 2000 итераций, сформировав пять слоев декомпозиции, охватывающих диапазон от коротких переходных компонентов до длительных тоновых компонентов (рис. 4).

Для детального анализа и субъективных тестов использовались четыре аудиообразца, записанных с помощью устройств амбисоники первого порядка с частотой дискретизации 48 кГц:

- 1) береговая зона¹;
- 2) тяжелые грузовики²;

¹RØDE Ambisonic sound library. – URL: <https://library.soundfield.com> (date of access: 20.11.2025).

²URL: <https://library.soundfield.com>

- 3) лифт³;
- 4) оркестровое произведение⁴.

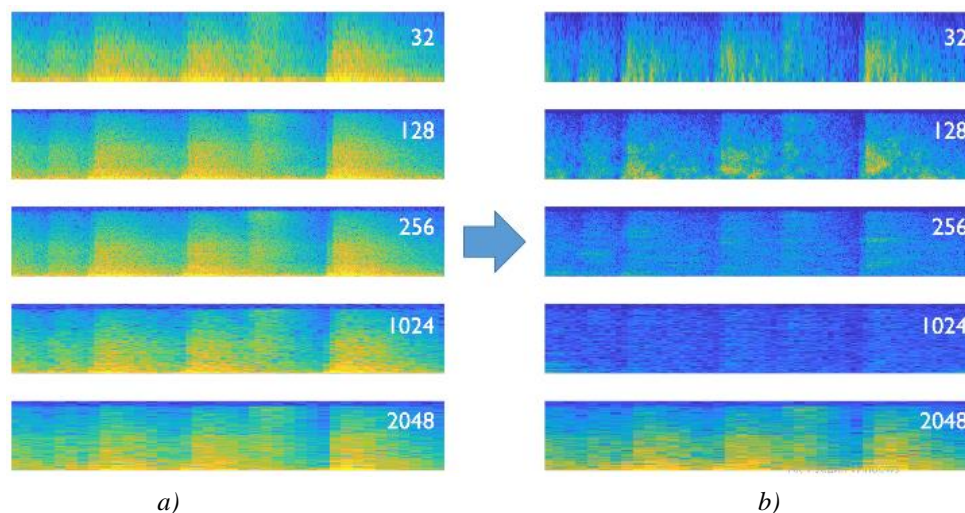


Рис. 4. Разреженная МДКП-декомпозиция после 2000 итераций:
 а) слои МДКП до разреженной оптимизации; б) слои МДКП после разреженной оптимизации
 Fig. 4. Sparse MDCT decomposition after 2000 iterations: a) MDCT layers before sparse optimization; b) MDCT layers after sparse optimization

Образцы 1–3 были записаны с помощью микрофона Rode NT-SF1, а образец 4 – с помощью микрофона Calrec Soundfield MkIV.

Закодированное аудио в формате амбисоники седьмого порядка визуально сравнивалось с оригиналом при помощи плагина для визуализации пространственного звука EnergyVisualizer из IEM Plug-in Suite⁵. Согласно визуализациям звукового поля аудио после обработки явно обладает значительно более высоким пространственным разрешением (рис. 5). Визуальная пространственная локализация обработанного звука устойчива в тестах и реагирует даже на тонкие изменения положения источников звука.

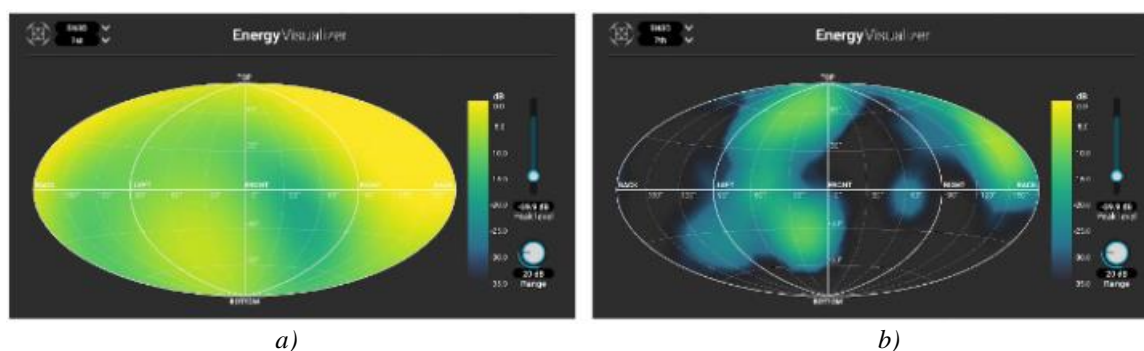


Рис. 5. Визуализация звукового поля исходного и обработанного амбисонического аудио:
 а) исходное аудио (амбисоника первого порядка); б) обработанное аудио (амбисоника седьмого порядка)
 Fig. 5. Visualization of the sound field of the original and processed ambisonic audio: a) original audio (first-order ambisonic); b) processed audio (seventh-order ambisonic)

³URL: <https://library.soundfield.com>

⁴Ambisonia. – URL: <http://www.ambisonia.com> (date of access: 20.11.2025).

⁵IEM Plug-in Suite. – URL: <https://plugins.iem.at> (date of access: 20.11.2025).

Результаты субъективной оценки пространственного разрешения декодированных аудиообразцов представлены в табл. 1, шкала субъективной оценки – в табл. 2. Оценка проводилась путем их прослушивания через наушники.

Таблица 1
Субъективная оценка пространственного разрешения

Table 1
Subjective assessment of spatial resolution

Схема декодирования <i>Decoding circuit</i>	Номер образца <i>Sample number</i>			
	1	2	3	4
Амбисоника первого порядка (без обработки)	3,9	4,0	3,8	3,9
Обработанный звук, амбисоника седьмого порядка, линейное МДКП-2048	4,2	4,2	4,0	4,3
Обработанный звук, амбисоника седьмого порядка, разреженное МДКП, без функции штрафа за алиасинг	4,5	4,4	4,3	3,5
Обработанный звук, амбисоника седьмого порядка, разреженное МДКП, с функцией штрафа за алиасинг	4,6	4,4	4,7	4,6

Таблица 2
Шкала субъективной оценки пространственного разрешения

Table 2
Subjective assessment scale for spatial resolution

Описание пространственного восприятия <i>Description of spatial perception</i>	Качественная характеристика <i>Qualitative characteristics</i>	Оценка в баллах <i>Score in points</i>
Четкая и устойчивая локализация всех источников звука, высокая детализация пространственной структуры, отсутствие артефактов	Отлично	5
Уверенная локализация источников, незначительное размывание пространственного образа, не мешающее восприятию	Хорошо	4
Среднее пространственное разрешение, допустимое размывание источников, общая пространственная картина сохранена	Удовлетворительно	3
Значительное размывание и неустойчивая локализация источников, затрудненное определение направлений	Плохо	2
Отсутствие четкого пространственного образа, полная диффузность звукового поля	Очень плохо	1

Для выявления вклада каждой составляющей предложенного метода приведены оценки для следующих схем декодирования:

- амбисоника первого порядка без обработки;
- амбисоника седьмого порядка, полученная декодированием с помощью обычного линейного МДКП с длиной окна 2048 отсчетов;
- амбисоника седьмого порядка, полученная декодированием с помощью разреженного МДКП с пятью слоями декомпозиции и без функции штрафа за наложение (алиасинг);

– амбисоника седьмого порядка, полученная декодированием с помощью разреженного МДКП с пятью слоями декомпозиции и с функцией штрафа за наложение (алиасинг).

Из табл. 1 видно, что полученные в ходе эксперимента результаты демонстрируют явное улучшение субъективного пространственного разрешения звуковых образов при использовании предложенного подхода.

Между тем следует акцентировать внимание на следующих особенностях и ограничениях проведенного эксперимента. Во-первых, использовалось небольшое количество аудиообразцов и малое разнообразие оборудования для прослушивания. Во-вторых, применение для прослушивания аудиообразцов только наушников может приводить к неточным оценкам, поскольку используемые в декодере встроенные передаточные функции для перевода сферической волны в бинауральный формат могут плохо восприниматься конкретным слушателем.

Несмотря на перечисленные выше ограничения, с высокой степенью уверенности можно утверждать, что по отношению к слышимым артефактам тест является надежным, а реализованный решатель временно-частотной декомпозиции обеспечивает разреженное представление с приемлемо высоким качеством. Функция штрафа за алиасинг вносит существенный вклад в восприятие отдельных звуков, что наглядно проявляется при сравнении образца 4 (оркестровое произведение), обработанного с применением и без применения функции штрафа.

Заметное улучшение пространственного разрешения достигается благодаря разреженному представлению (см. табл. 1). Вероятно, это достигается главным образом за счет отдельной пространственной локализации четких переходных звуков, которые достаточно отчетливо воспринимаются в декодированном звуковом образе.

==== Заключение

В статье рассмотрен метод повышения пространственного разрешения аудио, записанного в формате амбисоники первого порядка. Метод использует вещественную разреженную временно-частотную декомпозицию и обеспечивает более четкий пространственный образ. Практические результаты подтверждают применимость метода для обработки аудио с повышением разрешения вплоть до амбисоника седьмого порядка. Вычислительная сложность предложенного решения более чем на порядок выше, чем у HARPEX, по причине итеративного способа вычислений, что является его основным ограничением. Однако данный подход обеспечивает более точную временно-частотную интерпретацию сигнала и вполне применим для приложений, не требующих обработки в реальном масштабе времени.

Вклад авторов. *Д. С. Лихачёв* исследовал идею повышения пространственного разрешения амбисонического звука с использованием разреженного МДКП и оценил ее эффективность. *Н. А. Петровский* принимал участие в программной реализации системы, предназначенной для постановки экспериментов, и исследовал возможность эффективной аппаратной реализации оптимизационных задач. *И. С. Азаров* определил цели и задачи исследования, принимал участие в обсуждении полученных результатов. Все авторы принимали участие в подготовке текста статьи и интерпретации результатов экспериментов.

References

1. Pulkki V. Spatial sound reproduction with directional audio coding. *Journal of the Audio Engineering Society*, 2007, vol. 55, no. 6, pp. 503–516.
2. Berge S., Barrett N. High angular resolution planewave expansion. *Proceedings of the 2nd International Symposium on Ambisonics and Spherical Acoustics, Paris, France, 6–7 May 2010*. Available at: <https://www.harpex.net/harpex.pdf> (accessed 20.11.2025).
3. Herzog A., Habets E. A. P. Direction and reverberation preserving noise reduction of ambisonics signals. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 2020, vol. 28, pp. 2461–2475. <https://doi.org/10.1109/TASLP.2020.3013979>.
4. Rafaely B. *Fundamentals of Spherical Array Process*. Springer, 2015, 204 p. (Springer Topics in Signal Processing).
5. Schwartz O., Gannot S., Habets E. Multi-microphone speech dereverberation and noise reduction using relative early transfer functions. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 2015, vol. 23, no. 2, pp. 240–251.
6. Meyer J., Elko G. W. Spherical microphone arrays for 3D sound recordings. *Audio Signal Processing for Next-Generation Multimedia Communication Systems*. In Y. Huang, J. Benesty (eds). Norwell, MA, USA, Kluwer Academic, 2004, pp. 67–89.
7. McCormack L., Politis A. SPARTA & COMPASS: Real-time implementations of linear and parametric spatial audio reproduction and processing methods. *2019 AES International Conference on Immersive and Interactive Audio, York, UK, 27–29 March 2019*. Available at: https://acris.aalto.fi/ws/portalfiles/portal/34325781/ELEC_McCormick_Sparta_AES.pdf (accessed 20.11.2025).

Информация об авторах

Лихачёв Денис Сергеевич, кандидат технических наук, доцент кафедры электронных вычислительных средств, Белорусский государственный университет информатики и радиоэлектроники.
E-mail: likhachov@bsuir.by

Петровский Николай Александрович, кандидат технических наук, доцент кафедры электронных вычислительных средств, Белорусский государственный университет информатики и радиоэлектроники.
E-mail: nick.petrovsky@bsuir.by

Азаров Илья Сергеевич, доктор технических наук, профессор, заведующий кафедрой электронных вычислительных средств, Белорусский государственный университет информатики и радиоэлектроники.
E-mail: azarov@bsuir.by

Information about the authors

Denis S. Likhachov, Cand. Sci. (Eng.), Assoc. Prof. of Computer Engineering Department, Belarusian State University of Informatics and Radioelectronics.
E-mail: likhachov@bsuir.by

Nick A. Petrovsky, Cand. Sci. (Eng.), Assoc. Prof. of Computer Engineering Department, Belarusian State University of Informatics and Radioelectronics.
E-mail: nick.petrovsky@bsuir.by

Elias S. Azarov, Dr. Sci. (Eng.), Prof., Head of Computer Engineering Department, Belarusian State University of Informatics and Radioelectronics.
E-mail: azarov@bsuir.by

ОБРАБОТКА СИГНАЛОВ, ИЗОБРАЖЕНИЙ, РЕЧИ, ТЕКСТА И РАСПОЗНАВАНИЕ ОБРАЗОВ

SIGNAL, IMAGE, SPEECH, TEXT PROCESSING AND PATTERN RECOGNITION

УДК 004.94 + 534.77
<https://doi.org/10.37661/1816-0301-2026-23-1-69-87>

Поступила в редакцию | Received 26.01.2026
Подписана в печать | Accepted 23.02.2026
Опубликована | Published 31.03.2026

Распознавание эмоций по речи на основе LSTM-сетей с мультивекторным механизмом внимания

Д. В. Краснопрошин[✉], М. И. Вашкевич
[✉]E-mail: daniil.krasnoproshin@gmail.com

*Белорусский государственный университет
информатики и радиоэлектроники,
ул. П. Бровки, 6, Минск, 220013, Беларусь*

Аннотация

Цели. Целью исследования является повышение точности распознавания эмоций по речевому сигналу с помощью моделей на основе рекуррентных нейронных сетей (РНС) с долгой краткосрочной памятью.

Методы. В работе предложен мультивекторный механизм внимания для РНС на основе ячеек LSTM. Данный механизм представляет собой обобщение классического мягкого внимания и позволяет модели одновременно анализировать различные аспекты временных зависимостей. Предложенные архитектуры РНС применены к задаче распознавания эмоций по речевому сигналу. В качестве входных данных использовались последовательности мел-частотных кепстральных коэффициентов, отражающих частотно-временную структуру речевого сигнала. Эксперименты проводились на общедоступном наборе данных RAVDESS. Для автоматизированного подбора оптимальных гиперпараметров моделей использовался метод байесовской оптимизации.

Результаты. Результаты экспериментов с LSTM-сетями, имеющими различную размерность скрытого состояния (64, 96, 128), показывают, что применение мультивекторного механизма внимания приводит к статистически значимому улучшению среднего значения точности на величину от 0,88 до 1,56 %.

Заключение. Полученные результаты подтверждают целесообразность использования предложенного механизма мультивекторного внимания в архитектурах LSTM-сетей для задачи классификации эмоций в речи.

Ключевые слова: обработка речи, распознавание эмоций, глубокое обучение, рекуррентные нейронные сети, механизм внимания

Для цитирования. Краснопрошин, Д. В. Распознавание эмоций по речи на основе LSTM-сетей с мультивекторным механизмом внимания / Д. В. Краснопрошин, М. И. Вашкевич // Информатика. – 2026. – Т. 23, № 1. – С. 69–87. – <https://doi.org/10.37661/1816-0301-2026-23-1-69-87>.

Конфликт интересов. Авторы заявляют об отсутствии конфликта интересов.

Speech emotion recognition based on LSTM networks with multi-vector attention

Daniil V. Krasnoproshin[✉], Maxim I. Vashkevich

[✉]E-mail: daniil.krasnoproshin@gmail.com

*Belarusian State University of Informatics and Radioelectronics,
st. Brovki, 6, Minsk, 220013, Belarus*

Abstract

Objectives. Improvement of speech emotion recognition accuracy using Long Short-Term Memory (LSTM) recurrent neural network (RNN) models.

Methods. The paper proposes a multi-vector attention mechanism for LSTM-based RNNs. This mechanism generalizes the classical soft attention and allows the model to simultaneously analyze different aspects of temporal dependencies. The proposed RNN architectures were applied to the task of speech emotion recognition. Input data consisted of sequences of mel-frequency cepstral coefficients (MFCCs), which reflect the time-frequency structure of the speech signal. Experiments were conducted on the publicly available RAVDESS dataset. Bayesian optimization was employed for automated hyperparameter tuning of the models.

Results. The experimental results with LSTM networks having different hidden state dimensions (64, 96, 128) demonstrate that the application of the multi-vector attention mechanism leads to a statistically significant improvement in the average accuracy metric (UAR) by 0.88 to 1.56 %.

Conclusion. The obtained results confirm the effectiveness of using the proposed multi-vector attention mechanism in LSTM-based architectures for speech emotion classification.

Keywords: speech processing, emotion recognition, deep learning, recurrent neural networks, attention mechanism

For citation. Krasnoproshin D. V., Vashkevich M. I. *Speech emotion recognition based on LSTM networks with multi-vector attention*. *Informatika [Informatics]*, 2026, vol. 23, no. 1, pp. 69–87 (In Russ.). <https://doi.org/10.37661/1816-0301-2026-23-1-69-87>.

Conflict of interests. The authors declare of no conflict of interest.

Введение

Распознавание эмоций по речи является одной из важных задач в области обработки естественного языка и разработки человеко-машинных интерфейсов [1]. Развитие этой области обусловлено широким спектром практических приложений – от интеллектуальных голосовых ассистентов до анализа поведения пользователей и оценки эмоционального состояния в системах дистанционного обучения и медицинского обслуживания. Несмотря на существенный прогресс в области обработки естественного языка, достигнутый благодаря применению глубоких нейронных сетей, задача точного и устойчивого распознавания эмоций по речи остается сложной из-за высокой вариативности речевого сигнала, индивидуальных особенностей дикторов и контекстной зависимости эмоциональных проявлений [1–3].

Традиционные подходы к распознаванию эмоций, основанные на экспертных признаках (например, мел-частотных кепстральных коэффициентах, спектральных моментах и просодических характеристиках), демонстрируют приемлемую точность на небольших и «чистых» наборах данных [4]. Однако их обобщающая способность резко

снижается при работе с большими объемами зашумленных данных, что связано с ограниченной гибкостью параметризации и высокой чувствительностью к вариациям акустических условий [5].

С развитием глубокого обучения в задаче распознавания эмоций активно применяются сверточные нейронные сети (СНС), работающие с визуальным представлением речи – спектрограммами. Такие модели эффективно извлекают спектрально-временные признаки, достигая высоких результатов [6, 7]. Вместе с тем использование двумерных представлений сигнала приводит к частичной потере тонкой временной динамики, так как последовательная природа речи искусственно сводится к пространственной структуре, характерной для изображений. Данное ограничение особенно критично для задач, где временные зависимости (например, интонационные переходы) играют важную роль.

Современные исследования фокусируются на применении для задачи распознавания эмоций моделей с архитектурой трансформер, предобученных на аудиоданных [2, 8]. Благодаря способности моделировать долгосрочные зависимости трансформеры позволяют достигать высокой точности. Однако их практическое внедрение сдерживается высокой вычислительной сложностью: модели с десятками и сотнями миллионов параметров требуют значительных ресурсов для обучения и исполнения (англ. inference), что делает их непригодными для реального времени и устройств с ограниченным вычислительным ресурсом.

Учитывая указанные ограничения, перспективным направлением для разработки практических систем распознавания эмоций являются РНС. Различные архитектуры РНС изначально ориентированы на обработку последовательных данных, что позволяет эффективно учитывать временную динамику речевого сигнала [9]. Параметризируя аудиопоток в виде последовательности векторов, РНС сохраняют информацию о структуре и динамике речевого сигнала. Ключевое преимущество данного подхода – значительно меньшее количество параметров по сравнению с трансформерами при сохранении способности к моделированию контекстных зависимостей. Это делает РНС оптимальным выбором для сценариев, где важны скорость обработки и возможность запуска модели на мобильном или портативном устройстве.

В данной работе исследуется система распознавания эмоций на основе РНС с долгой краткосрочной памятью (англ. LSTM, Long Short-Term Memory) [10]. LSTM-сеть представляет собой ячейку памяти, которая посредством механизма управляющих затворов (англ. gate) изменяет во времени свое состояние в зависимости от текущего входного значения и предыдущего состояния. Наличие в структуре нескольких управляющих затворов позволяет LSTM-сети моделировать долгосрочные зависимости, продуктивно решая проблему «исчезающего» градиента, которая характерна для классических РНС [9]. Тем не менее, несмотря на способность LSTM-сетей моделировать временные зависимости, такие архитектуры не всегда эффективно фокусируются на наиболее значимых участках сигнала, что важно для правильного распознавания эмоций. Механизмы внимания (англ. attention mechanism) позволяют решать эту проблему, предоставляя модели возможность адаптивно выделять релевантные части входной последовательности. В частности, механизм мягкого внимания (англ. soft attention) показал свою эффективность в задаче классификации эмоций [11]. В настоящей работе предлагается новый ме-

ханизм мультивекторного внимания, который используется для формирования улучшенного вектора контекста из последовательности выходов LSTM-сети. Механизм мультивекторного внимания является обобщением механизма мягкого внимания, который также ранее использовался при решении задачи классификации эмоций с применением РНС [3, 11].

Целью исследования являются разработка и количественная оценка мультивекторного механизма внимания для LSTM-сетей в контексте задачи распознавания эмоций в речи. Предполагается, что увеличение числа векторов внимания позволит модели формировать улучшенное представление эмоционального содержания аудиосигнала, что приведет к повышению точности классификации эмоциональных состояний при сохранении умеренной вычислительной сложности.

Извлечение признаков

Для параметризации речевого сигнала в работе используются мел-частотные кепстральные коэффициенты (МЧКК). Основная идея метода расчета МЧКК заключается в приближенном моделировании особенностей восприятия звука в слуховой системе человека. Данный метод относится к кратковременному спектральному анализу речевых сигналов, при котором исходный сигнал разбивается на фреймы длительностью 10–30 мс. Такой временной интервал выбран исходя из предположения о квазистационарности речевого сигнала на указанных отрезках, что позволяет применять спектральные методы анализа. Преобразования речевого сигнала в процессе расчета МЧКК показаны на рис. 1.

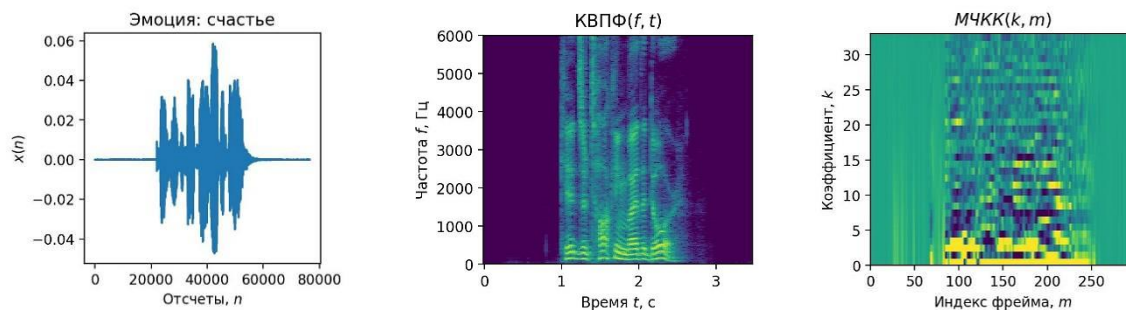


Рис. 1. Преобразования речевого сигнала в последовательность векторов с использованием МЧКК
Fig. 1. Transformations of speech signal into sequence of vectors using MFCC

Согласно рис. 1 процесс извлечения МЧКК включает следующие шаги [12]:

- 1) сигнал разделяется на короткие фреймы длины L с перекрытием h_{size} ;
- 2) для каждого фрейма вычисляется кратковременное преобразование Фурье (КВПФ) и находится квадрат модуля КВПФ;
- 3) выполняется переход от КВПФ к мел-спектрограмме (энергия сигнала из шкалы герц переводится в мел-шкалу, отражающую свойства человеческого слуха);
- 4) рассчитывается логарифм от энергии сигнала в мел-частотных полосах;
- 5) применяется декоррелирующее преобразование, в качестве которого выступает дискретное косинусное преобразование II типа (ДКП-II).

В результате параметризации речевого сигнала при помощи МЧКК образуется компактное представление речевого сигнала в последовательность векторов:

$$\mathbf{X} = [\mathbf{x}_0, \mathbf{x}_1, \dots, \mathbf{x}_{T-1}], \mathbf{x}_t \in \mathbb{R}^n, \quad (1)$$

где T – длина последовательности (количество фреймов сигнала).

В данной работе обрабатывались речевые сигналы с частотой дискретизации 48 кГц, размер фрейма выбирался равным $L=1024$, а перекрытие – $h_{size} = 512$ отсчетов. Размерность вектора МЧКК выбиралась равной $n = 34$.

===== LSTM-сеть с механизмом внимания

В настоящем исследовании разрабатывался мультивекторный механизм внимания, который основывается на способности LSTM-сети моделировать временные зависимости в последовательных данных, таких как речь. LSTM является улучшенной разновидностью РНС, которая решает проблему исчезающего градиента. LSTM-сети моделируют временные зависимости в последовательности \mathbf{X} с помощью скрытого состояния \mathbf{h}_t , которое обновляется на каждом временном шаге и передается для вычисления следующего состояния, что позволяет сохранять информацию о предыдущих элементах последовательности. Математически LSTM-ячейку можно описать набором уравнений для вычисления состояния и выхода на каждом временном шаге t :

$$\begin{aligned} \mathbf{f}_t &= \sigma(\mathbf{W}_f \cdot [\mathbf{h}_{t-1}, \mathbf{x}_t] + \mathbf{b}_f), \\ \mathbf{i}_t &= \sigma(\mathbf{W}_i \cdot [\mathbf{h}_{t-1}, \mathbf{x}_t] + \mathbf{b}_i), \\ \hat{\mathbf{C}}_t &= \tanh(\mathbf{W}_c \cdot [\mathbf{h}_{t-1}, \mathbf{x}_t] + \mathbf{b}_c), \\ \mathbf{C}_t &= \mathbf{f}_t \odot \mathbf{C}_{t-1} + \mathbf{i}_t \odot \hat{\mathbf{C}}_t, \\ \mathbf{o}_t &= \sigma(\mathbf{W}_o \cdot [\mathbf{h}_{t-1}, \mathbf{x}_t] + \mathbf{b}_o), \\ \mathbf{h}_t &= \mathbf{o}_t \odot \tanh(\mathbf{C}_t), \end{aligned} \quad (2)$$

где $\mathbf{f}_t, \mathbf{i}_t, \mathbf{o}_t \in \mathbb{R}^n$ – забывающий, входной и выходной затворы соответственно; $\mathbf{C}_t \in \mathbb{R}^n$ – внутреннее состояние LSTM-ячейки; $[\mathbf{h}_{t-1}, \mathbf{x}_t]$ – конкатенация предыдущего вектора скрытого состояния ячейки на шаге $t-1$ и входного вектора на шаге t ; $\tanh(\cdot)$ – функция активации гиперболический тангенс; $\sigma(\cdot)$ – функция активации логистического сигмоида; $\mathbf{W}_f, \mathbf{W}_i, \mathbf{W}_c, \mathbf{W}_o \in \mathbb{R}^{n \times 2n}$ – обучаемые матрицы весов; $\mathbf{b}_f, \mathbf{b}_i, \mathbf{b}_c, \mathbf{b}_o \in \mathbb{R}^n$ – обучаемые векторы смещений; \odot – операция поточечного умножения векторов.

Рассмотрим кратко функционирование модели (2). Ее сердцевиной является вектор состояния ячейки \mathbf{C}_t – долгосрочная память, которая хранит данные на протяжении всех временных шагов. Вектор внутреннего состояния \mathbf{h}_t – это текущая (или краткосрочная) память, значение которой подается на выход модели. Таким образом, \mathbf{h}_t хранит информацию, которая актуальна для текущего шага. В структуре LSTM-ячейки важную роль играет понятие затворов. Под затворами понимаются векторные переменные, которые определяют важность или, наоборот, неважность информации, хранимой в векторе-состоянии ячейки \mathbf{C}_t . Например, затвор \mathbf{f}_t определяет, какую часть информации своего

предыдущего состояния LSTM-ячейка должна «забыть», а затвор o , определяет, какая часть информации долгосрочной памяти важна на текущем шаге.

Таким образом, классическая LSTM-сеть принимает на вход последовательность входных векторов x_t и генерирует на выходе последовательность выходных векторов h_t . Требуется доопределить ее структуру, чтобы сеть была способна решать задачу классификации эмоций. С точки зрения классификации задач [9], которые решаются при помощи РНС, распознавание эмоций может быть отнесено к классу «один ко многим» (рис. 2).

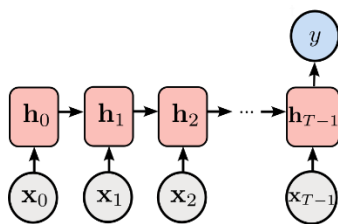


Рис. 2. Классификация эмоций при помощи РНС

Fig. 2. Emotion classification by RNN using many-to-one scheme

Диаграмма на рис. 2 показывает, что метка эмоции может быть получена путем классификации с использованием последнего скрытого состояния, полученного в LSTM-сети. Для этого достаточно вектор h_{T-1} подать на полносвязный слой с функцией классификации $\text{softmax}(\cdot)$. Ошибки, полученные в данном слое, затем распространяются обратно к началу последовательности x_t . В литературе данный подход известен как обучение по последнему фрейму (англ. final-frame training) [2]. Его основная идея заключается в предположении, что последнее скрытое состояние РНС содержит достаточную для корректной классификации эмоций информацию, извлеченную из всей речевой последовательности. При таком подходе РНС рассматривается в качестве адаптивного средства, преобразующего переменную по длине последовательность векторов X в вектор фиксированной размерности h_{T-1} .

Однако, как показано в работе [3], практическая эффективность данного подхода оказывается ограниченной. Это объясняется тем, что, хотя LSTM-архитектура и обладает улучшенной способностью к сохранению долгосрочных зависимостей по сравнению с обычными РНС, механизм забывания все же приводит к постепенной потере информации. В результате в последнем скрытом состоянии данные, относящиеся к начальным сегментам последовательности, могут быть представлены недостаточно полно. Для решения этой проблемы был предложен метод временного усреднения выходов LSTM-сети [11, 13], который позволяет агрегировать информацию со всех временных шагов. Под агрегацией в данном случае понимается глобальное усреднение всех выходных состояний LSTM-сети.

Таким образом, можно сделать вывод, что LSTM-сети обладают способностью учитывать долговременные зависимости в последовательных данных, однако их внутренняя структура не предполагает механизма приоритизации отдельных элементов временного контекста [3, 11].

Улучшить работу LSTM-сети можно за счет внедрения в ее структуру механизма внимания. Его добавление позволяет модели вычислять веса «важности» для различных временных шагов, что способствует получению более информативного выходного вектора, на основании которого выполняется классификация эмоции.

В настоящем исследовании в качестве базовой модели рассматривается LSTM-сеть с механизмом мягкого внимания, предложенная в работе [3]. В данной модели на основании последовательности выходных состояний \mathbf{h}_t рассчитывается взвешенная сумма скрытых состояний, которая называется вектором контекста:

$$\mathbf{h}_{wp} = \sum_{t=0}^{T-1} \alpha_t \mathbf{h}_t, \quad (3)$$

где α_t – весовые коэффициенты, отражающие значимость вектора состояния \mathbf{h}_t в формировании вектора контекста. Обозначение *wp* берется от англ. *weighted pooling* – взвешенное усреднение.

На основании вектора контекста \mathbf{h}_{wp} выполняется классификация эмоции, т. е. \mathbf{h}_{wp} подается на полносвязный слой с активационной функцией softmax. Общая схема описанного подхода изображена на рис. 3.

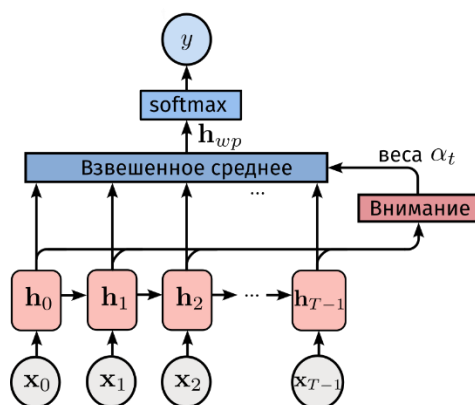


Рис. 3. Классификация эмоций при помощи РНС с механизмом внимания
 Fig. 3. Emotion classification using RNN with attention mechanism

В оригинальной работе [3] для формирования весов использовался механизм мягкого внимания (англ. *soft attention*):

$$\alpha_t = \text{softmax}(e_t) = \frac{\exp(e_t)}{\sum_{k=0}^{T-1} \exp(e_k)}, \quad (4)$$

где $e_t = \mathbf{u}^T \mathbf{h}_t$ – оценка внимания (англ. *attention score*), \mathbf{u} – вектор внимания (обучаемый параметр).

Добавление механизма мягкого внимания существенно улучшает работу LSTM-сети за счет того, что вектор контекста \mathbf{h}_{wp} адаптивно выделяет из выходной последовательности состояний те компоненты, которые наиболее важны для задачи распознавания эмоций.

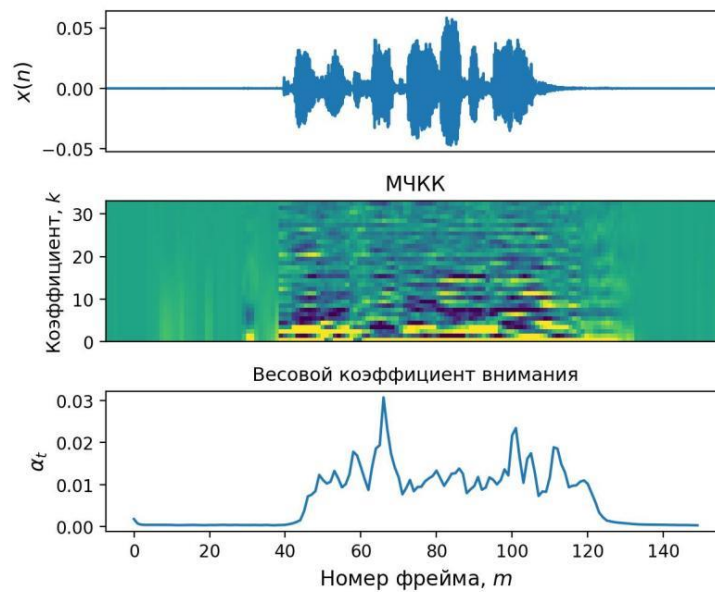


Рис. 4. Пример вычисления весов внимания
 Fig. 4. Example of local attention weights calculation

На рис. 4 показаны речевой сигнал, последовательность коэффициентов МЧКК, а также веса внимания, рассчитанные обученной LSTM-сетью. Видно, что после обучения механизм внимания приобретает ожидаемую селективность. Например, фреймам, которые относятся к паузам, присваиваются очень низкие значения весов внимания, а фреймам, несущим эмоциональную окраску, – высокие значения.

Мультивекторный механизм внимания

Рассмотренный выше механизм мягкого внимания можно интерпретировать следующим образом. Механизм в неявном виде предполагает, что в векторном пространстве, к которому принадлежат векторы состояний \mathbf{h}_t , имеется некоторое направление, соответствующее вектору внимания \mathbf{u} . При этом значимость вектора \mathbf{h}_t тем больше, чем больше его проекция на вектор внимания \mathbf{u} . Таким образом, в процессе обучения модель стремится «вытянуть» все релевантные векторы состояний \mathbf{h}_t вдоль направления вектора внимания \mathbf{u} .

В настоящей работе предлагается мультивекторный механизм мягкого внимания, согласно которому оценка внимания рассчитывается по формуле

$$e_t = \max \left(\left(\begin{array}{c} \mathbf{u}_1^T \\ \mathbf{u}_2^T \\ \vdots \\ \mathbf{u}_{N_{att}}^T \end{array} \right) \mathbf{h}_t \right), \quad (5)$$

где N_{att} – число векторов внимания; \mathbf{u}_i – набор векторов внимания, каждый из которых отвечает за независимое направление в пространстве векторов скрытых состояний.

Следует отметить, что мультивекторный механизм внимания включает в себя обычный механизм мягкого внимания как частный случай, когда $N_{att} = 1$.

Предполагается, что различные векторы внимания \mathbf{u}_i могут отвечать за различные проявления эмоций в последовательности \mathbf{h}_t . Таким образом, мультивекторный механизм внимания более эффективно использует пространство векторов скрытых состояний \mathbf{h}_t , формируя в нем не одно, а несколько базовых направлений, которые способны повлиять на получение более информативного контекстного вектора \mathbf{h}_{wp} .

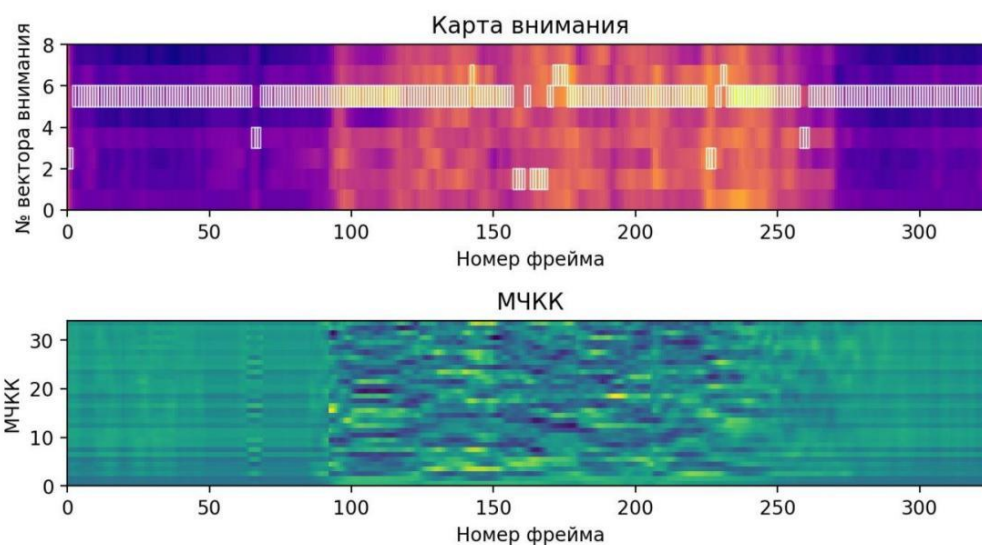


Рис. 5. Иллюстрация работы механизма мультивекторного внимания
 Fig. 5. Illustration of the multi-head attention mechanism

На рис. 5 показана работа обученного механизма мультивекторного внимания. На верхней панели рис. 5 изображена карта внимания, вычисленная для модели с восемью векторами внимания. Белым контуром выделены оценки внимания, которые имеют максимальное значение на текущем временном шаге. Рис. 5 демонстрирует, что чаще всего максимальное значение оценки внимания дает проекция на вектор \mathbf{u}_6 . Однако в отдельные моменты времени максимальные оценки дают проекции \mathbf{h}_t на другие векторы внимания. Таким образом, применение механизма мультивекторного внимания позволяет LSTM-сети формировать более информативно насыщенное представление об эмоции, содержащейся в речевом сигнале.

Описание базы данных

При проведении исследования использовался набор данных Ryerson Audio-Visual Database of Emotional Speech and Song (RAVDESS). RAVDESS содержит записи 24 актеров (12 мужчин и 12 женщин). Для каждого актера имеется 104 различных сообщения (60 речевых высказываний и 44 песенных). В рамках данной работы использована только часть базы RAVDESS, содержащая речевые высказывания. В этой части содержатся 1440 файлов в формате wav (16 бит, 48 кГц) – 60 записей на каждого из 24 актеров. Речевые эмоции включают: нейтральность, спокойствие, счастье, грусть, гнев, страх,

удивление и отвращение. Все эмоциональные состояния, кроме нейтрального, озвучивались на двух уровнях эмоциональной громкости (нормальная и повышенная). Актеры повторяли каждую вокализацию дважды.

Постановка эксперимента

Для экспериментальной оценки мультивекторного механизма внимания была реализована однослойная РНС на основе LSTM-ячеек. В качестве входных данных использовались нормализованные МЧКК размерностью $n = 34$. Исследование проводилось для различных размерностей скрытого состояния РНС: $h_{size} \in \{64, 96, 128\}$. Количество векторов внимания варьировалось в дискретном множестве значений $N_{att} \in \{1, 2, 4, 8, 16, 32, 64\}$. Для модели со скрытой размерностью 128 дополнительно были протестированы конфигурации с 128 и 256 векторами внимания.

Разработанные LSTM-сети с механизмом внимания обучались с помощью метода оптимизации Adam [9]. В качестве функции потерь использовалась перекрестная энтропия, которая оптимальна для задач многоклассовой классификации и обеспечивает эффективное вычисление градиентов при работе с вероятностными распределениями:

$$CE = \sum_{n=0}^{N-1} \sum_{i=0}^{N_c-1} y_{true}^{(i,n)} \cdot \ln(y_{pred}^{(i,n)}), \quad (6)$$

где N – число обучающих примеров, N_c – количество классов, y_{true} – истинное распределение меток, y_{pred} – предсказанные вероятности классов.

Инициализация весов модели выполнялась с учетом архитектурных особенностей LSTM-сети: веса инициализировались методом Ксавье с нормальным распределением [9], что обеспечивает сохранение дисперсии сигналов при прямом и обратном распространении через слой; смещения всех затворов LSTM-ячейки инициализировались нулевыми значениями, за исключением затвора забывания, смещения которого устанавливались равными единице для обеспечения начального сохранения информации в ячейке памяти.

Оптимизация гиперпараметров проводилась с использованием фреймворка Optuna, который обеспечивает эффективный автоматизированный поиск конфигураций модели [14]. Этот подход динамически формирует пространство поиска в процессе оптимизации. В качестве алгоритма сэмпирования значений гиперпараметров Optuna использует метод древовидных оценок Парзена (англ. TPE, Tree-Structured Parzen Estimator). TPE – это байесовский метод оптимизации, который моделирует плотности вероятности высокоэффективных и низкоэффективных конфигураций гиперпараметров по отдельности [15]. Сравнивая эти распределения, TPE интеллектуально направляет поиск в перспективные области пространства гиперпараметров, значительно ускоряя сходимость по сравнению со случайными методами или методами поиска по сетке.

Оптимизация гиперпараметров включала пять основных параметров: скорость обучения (выборка в логарифмическом масштабе в диапазоне $\eta = [3 \cdot 10^{-5}, 2 \cdot 10^{-4}]$), коэффициент затухания весов (выборка в логарифмическом масштабе в диапазоне $\lambda = [2 \cdot 10^{-5}, 2 \cdot 10^{-2}]$), отсечение (англ. dropout) в полносвязных слоях (равномерная выборка в диапазоне $p_{drop} = [10^{-1}, 5 \cdot 10^{-1}]$), количество циклов отжига в планировщике ко-

синусного отжига $T_0 \in \{1, 2, 3, 5, 10\}$ и размер пакета $batch_size \in \{16, 32, 64\}$. Каждая модель обучалась в течение 200 эпох, поскольку точность на валидационном наборе после этого практически не улучшалась.

Оптимизация выполнялась для базовой модели с одним вектором внимания. Полученные оптимальные параметры использовались далее для всех экспериментов с различным числом векторов внимания. Для каждой конфигурации модели обучение повторялось 10 раз с разными начальными значениями весов. Для итоговой оценки качества модели вычисляли среднее невзвешенное значение полноты (англ. unweighted average recall, UAR):

$$UAR = \frac{1}{N_c} \sum_{i=1}^{N_c} \frac{A_{i,i}}{\sum_{j=1}^{N_c} A_{i,j}}, \quad (7)$$

где A – матрица ошибок (англ. confusion matrix), N_c – количество классов.

Метрика UAR применяется для оценки общей эффективности моделей многоклассовой классификации. В UAR полнота, полученная для каждого класса, вносит равный вклад в итоговую оценку независимо от распределения данных. Значения UAR нормированы в диапазоне $[0, 1]$, где значения, приближающиеся к единице, соответствуют более высокой производительности классификатора. Для статистически надежной оценки модели применялась процедура перекрестной проверки по пяти блокам (англ. 5-fold cross-validation). В исследовании использовалась схема разбиения дикторов на блоки, предложенная в работе [2], для обеспечения воспроизводимости результатов и возможности их сопоставления с предыдущими исследованиями [4, 5, 12].

===== Результаты экспериментов

В данной работе для обозначения исследуемых моделей принята нотация LSTM-hX-vY, где X соответствует размерности скрытого состояния, а Y – количеству векторов механизма внимания.

Результаты, полученные в ходе экспериментов для моделей LSTM-h64, приведены в табл. 1.

Таблица 1

Результаты экспериментального исследования LSTM-сети со скрытым слоем размерности 64

Table 1

Experimental study results for LSTM network with hidden layer dimension of 64

Модель <i>Model</i>	Число параметров <i>Number of parameters</i>	Точность <i>UAR</i>	Пиковое значение точности <i>UAR_{max}</i>
LSTM-h64-v1	26 312	0,5510 ± 0,0071	0,5625
LSTM-h64-v2	26 376	0,5566 ± 0,0104	0,5742
LSTM-h64-v4	26 504	0,5548 ± 0,0065	0,5618
LSTM-h64-v8	26 760	0,5593 ± 0,0079	0,5710
LSTM-h64-v16	27 272	0,5586 ± 0,0112	0,5775
LSTM-h64-v32	28 296	0,5581 ± 0,0109	0,5729
LSTM-h64-v64	30 344	0,5598 ± 0,0083	0,5742

Анализ табл. 1 позволяет сделать вывод, что добавление механизма мультивекторного внимания повышает точность модели. В частности, среднее значение метрики UAR для модели с 64 векторами внимания на 0,88 % превышает аналогичный показатель модели с одним вектором внимания. Полученные наборы значений UAR были проанализированы с использованием парного t-критерия Стьюдента. Статистический тест показал, что различие в средних значениях между моделью с одним и 64 векторами внимания не достигает общепринятого уровня статистической значимости ($p = 0,052$). Однако, несмотря на формальное отсутствие статистической значимости на уровне $\alpha = 0,05$, наблюдаемые различия могут иметь практическую важность. В частности, из полученных данных следует, что добавление дополнительных векторов внимания приводит к повышению максимальной достигнутой точности модели в серии экспериментов. Например, пиковое значение UAR для модели с 16 векторами внимания составляет 0,5775, что на 1,5 % выше аналогичного показателя базовой модели с одним вектором вниманием.

Результаты, полученные в ходе экспериментов для моделей LSTM-h96, приведены в табл. 2.

Таблица 2

Результаты экспериментального исследования LSTM-сети со скрытым слоем размерности 96

Table 2

Experimental study results for LSTM network with hidden layer dimension of 96

Модель <i>Model</i>	Число параметров <i>Number of parameters</i>	Точность <i>UAR</i>	Пиковое значение точности <i>UAR_{max}</i>
LSTM-h96-v1	51 752	0,5485 ± 0,0059	0,5573
LSTM-h96-v2	51 848	0,5515 ± 0,0068	0,5605
LSTM-h96-v4	52 040	0,5508 ± 0,0150	0,5729
LSTM-h96-v8	52 424	0,5538 ± 0,0093	0,5716
LSTM-h96-v16	53 192	0,5515 ± 0,0079	0,5684
LSTM-h96-v32	54 728	0,5566 ± 0,0091	0,5788
LSTM-h96-v64	57 800	0,5641 ± 0,0112	0,5859

Полученные результаты также показывают, что увеличение количества векторов внимания улучшает качество классификации. Так, при помощи парного t-критерия была выявлена статистически значимая разница между моделью с одним и 64 векторами внимания ($p < 0,02$). При этом среднее значение UAR для модели с 64 векторами внимания превышает аналогичный показатель модели с одним вектором внимания на 1,56 %. Отметим также, что пиковое значение UAR для модели LSTM-h96-v64 составляет 0,5859, что на 2,86 % выше, чем у базовой модели с одним вектором внимания.

Результаты, полученные в ходе экспериментов для моделей LSTM-h128, приведены в табл. 3. Они демонстрируют устойчивую тенденцию, наблюдаемую в экспериментах: увеличение количества векторов внимания положительно влияет на качество классификации модели. Во-первых, анализ средних значений метрики UAR показывает прогрессирующий рост точности. Во-вторых, при помощи парного t-критерия была вы-

явлена еще большая статистически значимая разница между моделью с одним и 256 векторами внимания ($p < 0,0008$). Можно отметить, что средний UAR для модели с 256 векторами внимания превышает показатель базовой модели (с одним вектором) на 1,45 %. Это указывает на то, что мультивекторное внимание позволяет модели более эффективно извлекать и использовать релевантную информацию из последовательностей данных.

Таблица 3
 Результаты экспериментального исследования LSTM-сети со скрытым слоем размерности 128

Table 3
 Experimental study results for LSTM network with hidden layer dimension of 128

Модель <i>Model</i>	Число параметров <i>Number of parameters</i>	Точность <i>UAR</i>	Пиковое значение точности <i>UAR_{max}</i>
LSTM-h128-v1	85 384	0,5600 ± 0,0097	0,5794
LSTM-h128-v2	85 512	0,5636 ± 0,0078	0,5801
LSTM-h128-v4	85 768	0,5727 ± 0,0117	0,5840
LSTM-h128-v8	86 280	0,5708 ± 0,0111	0,5859
LSTM-h128-v16	87 304	0,5658 ± 0,0062	0,5768
LSTM-h128-v32	89 352	0,5716 ± 0,0126	0,5964
LSTM-h128-v64	93 448	0,5754 ± 0,0092	0,5859
LSTM-h128-v128	101 640	0,5766 ± 0,0117	0,5996
LSTM-h128-v256	118 024	0,5781 ± 0,0069	0,5892

Как и в предыдущих сериях экспериментов, максимальная точность модели также значительно возрастает. Например, пиковое значение UAR для конфигурации с 128 векторами внимания достигает 0,5996, что на 1,95 % выше, чем у базового варианта с одним вектором внимания.

На рис. 6 изображена матрица ошибок, полученная для модели LSTM-h128-v128, у которой UAR=0,5996.

Анализ матрицы ошибок показывает, что модель для распознавания эмоций демонстрирует наилучшие результаты для классов «отвращение» (73 %), «страх» (69 %) и «удивление» (67 %). Это свидетельствует о наличии у них уникальных и легко выделяемых закономерностей, отражаемых в данных. Наибольшую сложность представляет идентификация «нейтральности» (44 %) и «счастья» (46 %), которые часто путают как друг с другом, так и с другими эмоциями, что указывает на размытость границ этих состояний в признаковом пространстве. Взаимные ошибки часто возникают при классификации «счастья» и «грусти» (счастье принимается за грусть в 12,5 %, а грусть за счастье в 10,4 %), что, возможно, связано с неярко выраженной этими эмоциями в наборе данных RAVDESS. Обращает на себя также внимание тот факт, что 20,8 % нейтральных образцов классифицированы как «счастье» – это достаточно распространенная проблема в системах распознавания эмоций, поскольку спокойная речь может восприниматься за слабовыраженную радость.

Таким образом, модель демонстрирует большую эффективность при различении эмоций по степени их интенсивности и яркости проявления (отвращение, страх, удивление), однако меньшую эффективность при точной идентификации их положительной или отрицательной окраски, что особенно заметно для состояний со слабой выраженностью (грусть, счастье).

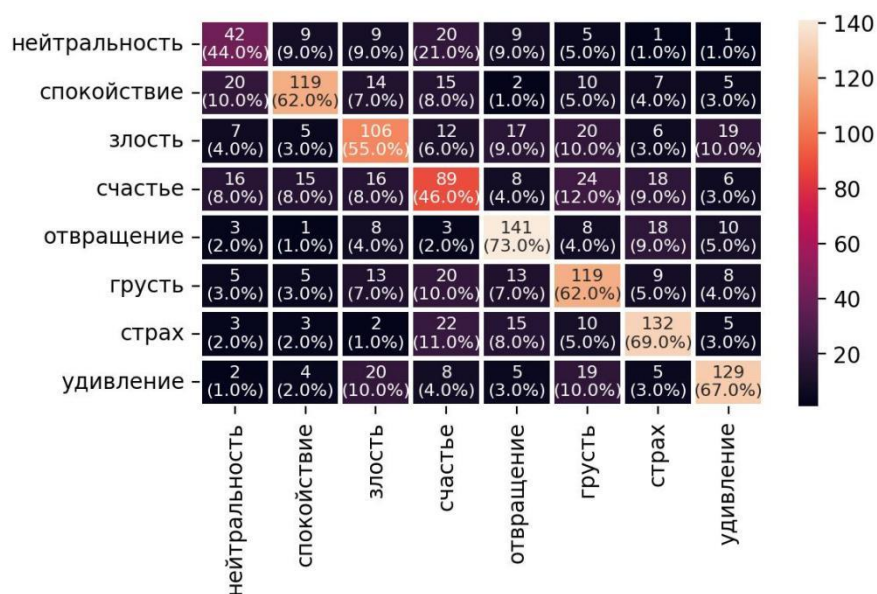


Рис. 6. Матрица ошибок для модели LSTM-h128-v128

Fig. 6. Confusion matrix for LSTM-h128-v128

Из вышесказанного можно сделать вывод, что результаты для различных конфигураций модели на основе архитектуры LSTM согласуются с гипотезой, проверяемой в рамках данного исследования: переход от мягкого внимания с одним вектором к мультивекторному вниманию является эффективным методом, статистически значимо повышающим как среднюю, так и пиковую точность модели в задаче классификации.

На рис. 7 показаны сглаженные распределения значений UAR для моделей с 64, 96 и 128 векторами внимания. Представленные графики позволяют сделать вывод, что распределение значений UAR для моделей с несколькими векторами внимания сдвинуто вправо относительно распределения модели с одним вектором. Полученные распределения наглядно демонстрируют, что при одних и тех же условиях модель с несколькими векторами внимания показывает более высокие показатели качества, чем модель с одним вектором внимания.

В целом рост точности можно объяснить тем, что каждый вектор внимания учится фокусироваться на различных аспектах эмоциональных признаков – интонации, тембре, спектральной энергии и т. д. Таким образом, использование нескольких таких векторов создает более богатое представление входного сигнала, повышая способность модели выделять эмоционально релевантные сегменты речи.

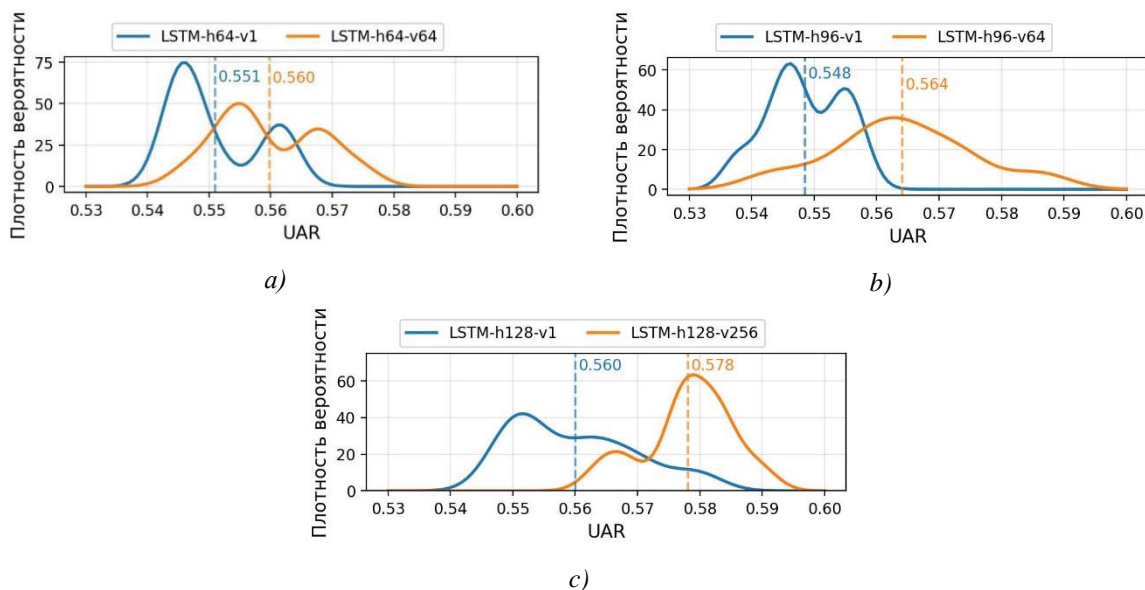


Рис. 7. Сглаженные распределения точностей (UAR) моделей: а) LSTM-h64-v1 и LSTM-h64-v64; б) LSTM-h96-v1 и LSTM-h96-v64; в) LSTM-h128-v1 и LSTM-h128-v256

Fig. 7. Smoothed distributions of model UAR: а) LSTM-h64-v1 and LSTM-h64-v64; б) LSTM-h96-v1 and LSTM-h96-v64; в) LSTM-h128-v1 and LSTM-h128-v256

Результаты, полученные в рамках настоящего исследования, были сопоставлены с известными из литературы методами классификации эмоций по речи (табл. 4). Поскольку в опубликованных работах, как правило, приводятся точечные оценки производительности моделей, а не интервальные, в сводной табл. 4 для моделей указаны их максимальные достигнутые значения метрики UAR.

Таблица 4

Сравнение производительности различных моделей классификации эмоций с использованием набора данных RAVDESS

Table 4

Comparison of the performance of various speech emotion classification models using the RAVDESS database

Модель <i>Model</i>	Число параметров <i>Number of parameters</i>	Точность <i>UAR</i>
SVM с 4096-мерным вектором признаков, полученных из модели AlexNet [2]	61 000 тыс.	0,4580
SVM с 304-мерным вектором надсегментных признаков [12]	–	0,4820
LDA со 190-мерным вектором признаков [4]	–	0,5380
GResNet+S [16]	–	0,5970
LSTM-h128-v128 [предлагаемая]	101,6 тыс.	0,5996
AlexNet [2]	61 000 тыс.	0,6167
CNN14 [2]	81 000 тыс.	0,7658
Модифицированная LARGE xlsr-Wav2Vec2.0 [17]	317 000 тыс.	0,8182

Анализ табл. 4 позволяет выделить некоторые тенденции в эволюции методов распознавания эмоций по речи. Традиционные методы машинного обучения показывают ограниченную эффективность: SVM (англ. support vector machine) с признаками, извлеченными предобученной AlexNet, имеет точность $UAR = 0,4580$, а его модификация из работы [12] достигает $UAR = 0,4820$. Примечателен результат классификатора на основе линейного дискриминантного анализа (англ. LDA, linear discriminant analysis) [4] с $UAR = 0,5380$, учитывая простоту данного метода.

Более высокие показатели демонстрируют специализированные архитектуры нейронных сетей. Модель GResNet+S [16], основанная на блоках со стробируемыми остаточными связями (англ. gated residual networks blocks), имеет достаточно высокую точность ($UAR = 0,5970$) и извлекает признаки непосредственно из спектрограмм. Предобученные модели, дообученные (англ. fine tuned) на целевом наборе данных, показывают прирост точности. Так, в работе [2] предобученная модель AlexNet дообучалась на базе RAVDESS, в результате достигнута точность $UAR = 0,6167$. Хотя эта точность превосходит результаты, которые достигаются при помощи рекуррентных и сверточных сетей, ее можно считать ограниченной, учитывая большой размер модели – порядка 61 млн параметров. Авторы статьи [2] связывают такой результат с тем, что сеть была предобучена не на речевых данных, а на базе изображений ImageNet.

Переход к моделям, предобученным непосредственно на аудиоданных, обеспечивает качественный скачок. CNN14 [8], дообученная в работе [2], демонстрирует $UAR = 0,7658$. Наивысший результат ($UAR = 0,8182$) достигается при использовании крупной трансформерной модели xlsr-Wav2Vec2.0 [17], хотя ее сложность (317 млн параметров) ограничивает практическое применение.

Таким образом, использование больших предобученных моделей подтверждает возможность построения высокоточных систем распознавания эмоций. Однако для практических задач необходимы архитектуры с умеренной вычислительной сложностью. В этом контексте предложенные модели с мультивекторным механизмом внимания представляют перспективное направление, сочетающее приемлемую точность с вычислительной эффективностью.

===== Заключение

В работе проведено исследование влияния количества векторов внимания на качество распознавания эмоций в речи с использованием LSTM-сетей. Основное внимание уделялось разработке и экспериментальной оценке мультивекторного механизма внимания, предназначенного для более гибкого выделения информативных фрагментов акустического сигнала.

Эксперименты, выполненные на наборе данных RAVDESS, показали, что увеличение числа векторов внимания приводит к статистически значимому росту метрики точности UAR по сравнению с базовой моделью, использующей один вектор внимания. Это свидетельствует о том, что предложенный механизм способствует более точному моделированию эмоциональных состояний в речи. Несмотря на умеренное абсолютное улучшение метрики, статистическая значимость и воспроизводимость указывают на устойчивое положительное влияние разработанного подхода.

Перспективными направлениями дальнейших исследований представляются: интеграция разнородных акустических признаков, включая мел-спектрограммы и хромограммы, для формирования более полного описания речевого сигнала; использование сверточных слоев для автоматического извлечения иерархических признаков непосредственно из спектральных представлений аудиосигнала [7].

Вклад авторов. Д. В. Краснопрошин занимался решением задач исследования, анализом полученных результатов и формированием структуры статьи; М. И. Вашкевич определил цель и задачи, которые необходимо было решить в ходе проведения исследования, руководил исследованием и принял участие в разработке метода интерпретации результатов эксперимента, редактировании текста статьи и подготовке графического материала.

Список использованных источников

1. A review of affective computing: From unimodal analysis to multimodal fusion / S. Poria, E. Cambria, R. Bajpai, A. Hussain // *Information Fusion*. – 2017. – Vol. 37. – P. 98–125.
2. Multimodal emotion recognition on RAVDESS dataset using transfer learning / C. Luna-Jiménez, D. Griol, Z. Callejas [et al.] // *Sensors*. – 2021. – Vol. 21. – P. 1–29.
3. Mirsamadi, S. Automatic speech emotion recognition using recurrent neural networks with local attention / S. Mirsamadi, E. Barsoum, C. Zhang // *Proc. of IEEE Intern. Conf. on Acoustics, Speech and Signal Processing (ICASSP)*, New Orleans, LA, USA, 05–09 Mar. 2017. – New Orleans, 2017. – P. 2227–2231.
4. Краснопрошин, Д. В. Отбор признаков на основе техники переноса обучения для классификации эмоций в речи с помощью полносвязной нейронной сети прямого распространения / Д. В. Краснопрошин, М. И. Вашкевич // *Системный анализ и прикладная информатика*. – 2025. – № 1. – С. 38–43.
5. Краснопрошин, Д. В. Анализ подходов к построению систем распознавания эмоций по речи с использованием методов глубокого обучения / Д. В. Краснопрошин, М. И. Вашкевич // *Big Data and Advanced Analytics* : сб. науч. ст. XI Междунар. науч.-практ. конф., Минск, 23–24 апр. 2025 г. – Мн., 2025. – С. 343–353.
6. Dal Rí, F. A. Speech emotion recognition and deep learning: an extensive validation using convolutional neural networks / F. A. Dal Rí, F. C. Ciardi, N. Conci // *IEEE Access*. – 2023. – Vol. 11. – P. 116638–116649.
7. Waleed, G. T. Speech emotion recognition on MELD and RAVDESS datasets using CNN / G. T. Waleed, S. H. Shaker // *Information*. – 2025. – Vol. 16, no. 7. – P. 518.
8. PANNs: Large-scale pretrained audio neural networks for audio pattern recognition / Q. Kong, Y. Cao, T. Iqbal [et al.] // *IEEE/ACM Transactions on Audio, Speech, and Language Processing*. – 2020. – Vol. 28. – P. 2880–2894.
9. Николенко, С. Глубокое обучение / С. Николенко, А. Кадури, Е. Архангельская. – СПб. : Питер, 2019. – 480 с.
10. Hochreiter, S. Long short-term memory / S. Hochreiter, J. Schmidhuber // *Neural Computation*. – 1997. – Vol. 9, no. 8. – P. 1735–1780.
11. Context-aware attention mechanism for speech emotion recognition / G. Ramet, P. N. Garner, M. Baeriswyl, A. Lazaridis // *2018 IEEE Spoken Language Technology Workshop (SLT)*, Athens, Greece, 18–21 Dec. 2018. – Athens, 2018. – P. 126–131.
12. Краснопрошин, Д. В. Метод распознавания эмоций в речевом сигнале с использованием машины опорных векторов и надсегментных акустических признаков / Д. В. Краснопрошин, М. И. Вашкевич // *Доклады БГУИР*. – 2024. – Т. 22, № 3. – С. 93–100.

13. Bahdanau, D. Neural machine translation by jointly learning to align and translate / D. Bahdanau, K. Cho, Y. Bengio // 3rd Intern. Conf. on Learning Representations, ICLR 2015, San Diego, CA, USA, 7–9 May 2015. – San Diego, 2015. – URL: <https://arxiv.org/abs/1409.0473> (date of access: 13.11.2025).
14. Optuna: A next-generation hyperparameter optimization framework / T. Akiba, S. Sano, T. Yanase [et al.] // Proc. of the 25th ACM SIGKDD Intern. Conf. on Knowledge Discovery & Data Mining (KDD'19), Anchorage, AK, USA, 4–8 Aug. 2019. – Anchorage, 2019. – P. 2623–2631.
15. Algorithms for hyper-parameter optimization / J. S. Bergstra, R. Bardenet, Y. Bengio, B. Kégl // NIPS'11: Proc. of the 25th Intern. Conf. on Neural Information Processing Systems, Granada, Spain, 12–15 Dec. 2011. – Granada, 2011. – P. 2546–2554.
16. Spectrogram based multi-task audio classification / Y. Zeng, H. Mao, D. Peng, Z. Yi // Multimedia Tools and Applications. – 2019. – Vol. 78, no. 3. – P. 3705–3722.
17. A proposal for multimodal emotion recognition using aural transformers and action units on RAVDESS dataset / C. Luna-Jiménez, R. Kleinlein, D. Griol [et al.] // Applied Sciences. – 2022. – Vol. 12, no. 1. – P. 1–23.

References

1. Poria S., Cambria E., Bajpai R., Hussain A. A review of affective computing: From unimodal analysis to multimodal fusion. *Information Fusion*, 2017, vol. 37, pp. 98–125.
2. Luna-Jiménez C., Griol D., Callejas Z., Kleinlein R., Montero J. M., Fernández-Martínez F. Multimodal emotion recognition on RAVDESS dataset using transfer learning. *Sensors*, 2021, vol. 21, pp. 1–29.
3. Mirsamadi S., Barsoum E., Zhang C. Automatic speech emotion recognition using recurrent neural networks with local attention. *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), New Orleans, LA, USA, 05–09 March 2017*. New Orleans, 2017, pp. 2227–2231.
4. Krasnoproshin D. V., Vashkevich M. I. *Transfer learning based feature selection for feedforward neural network for speech emotion classifier*. Sistemnyj analiz i prikladnaja informatika [System Analysis and Applied Information Science], 2025, no. 1, pp. 38–43 (In Russ.).
5. Krasnoproshin D. V., Vashkevich M. I. *Analysis of approaches to building speech emotion recognition systems using deep learning methods*. Big Data and Advanced Analytics: sbornik nauchnyh statej XI Mezhdunarodnoj nauchno-prakticheskoy konferencii, Minsk, 23–24 aprelja 2025 g. [Big Data and Advanced Analytics: Collection of Scientific Articles of the XI International Scientific and Practical Conference, Minsk, 23–24 April 2025]. Minsk, 2025, pp. 343–353 (In Russ.).
6. Dal Rí F. A., Ciardi F. C., Conci N. Speech emotion recognition and deep learning: an extensive validation using convolutional neural networks. *IEEE Access*, 2023, vol. 11, pp. 116638–116649.
7. Waleed G. T., Shaker S. H. Speech emotion recognition on MELD and RAVDESS datasets using CNN. *Information*, 2025, vol. 16, no. 7, p. 518.
8. Kong Q., Cao Y., Iqbal T., Wang Y., Wang W., Plumbley M. D. PANNs: Large-scale pretrained audio neural networks for audio pattern recognition. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 2020, vol. 28, pp. 2880–2894.
9. Nikolenko S., Kadurin A., Archangelskaya E. Glubokoe obuchenie. *Deep Learning*. Saint Petersburg, Piter, 2019, 480 p. (In Russ.).
10. Hochreiter S., Schmidhuber J. Long short-term memory. *Neural Computation*, 1997, vol. 9, no. 8, pp. 1735–1780.
11. Ramet G., Garner P. N., Baeriswyl M., Lazaridis A. Context-aware attention mechanism for speech emotion recognition. *2018 IEEE Spoken Language Technology Workshop (SLT), Athens, Greece, 18–21 December 2018*. Athens, 2018, pp. 126–131.

12. Krasnoproshin D. V., Vashkevich M. I. *Speech emotion recognition method based on support vector machine and suprasegmental acoustic features*. Doklady BGUIR [BGUIR Proceedings], 2024, vol. 22, no. 3, pp. 93–100 (In Russ.).
13. Bahdanau D., Cho K., Bengio Y. Neural machine translation by jointly learning to align and translate. *3rd International Conference on Learning Representations, ICLR 2015, San Diego, CA, USA, 7–9 May 2015*. San Diego, 2015. Available at: <https://arxiv.org/abs/1409.0473> (accessed 13.11.2025).
14. Akiba T., Sano S., Yanase T., Ohta T., Koyama M. Optuna: A next-generation hyperparameter optimization framework. *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining (KDD'19), Anchorage, AK, USA, 4–8 August 2019*. Anchorage, 2019, pp. 2623–2631.
15. Bergstra J. S., Bardenet R., Bengio Y., Kégl B. Algorithms for hyper-parameter optimization. *NIPS'11: Proceedings of the 25th International Conference on Neural Information Processing Systems, Granada, Spain, 12–15 December 2011*. Granada, 2011, pp. 2546–2554.
16. Zeng Y., Mao H., Peng D., Yi Z. Spectrogram based multi-task audio classification. *Multimedia Tools and Applications*, 2019, vol. 78, no. 3, pp. 3705–3722.
17. Luna-Jiménez C., Kleinlein R., Griol D., Callejas Z., Montero J. M., Fernández-Martínez F. A proposal for multimodal emotion recognition using aural transformers and action units on RAVDESS dataset. *Applied Sciences*, 2022, vol. 12, no. 1, pp. 1–23.

Информация об авторах

Краснопрошин Даниил Вадимович, магистр технических наук, аспирант кафедры электронных вычислительных средств, Белорусский государственный университет информатики и радиоэлектроники.
E-mail: daniil.krasnoproshin@gmail.com

Вашкевич Максим Иосифович, доктор технических наук, доцент, профессор кафедры электронных вычислительных средств, Белорусский государственный университет информатики и радиоэлектроники.
E-mail: vashkevich@bsuir.by

Information about the authors

Daniil V. Krasnoproshin, M. Sci. (Eng.), Postgraduate Student of Computer Engineering Department, Belarusian State University of Informatics and Radioelectronics.
E-mail: daniil.krasnoproshin@gmail.com

Maxim I. Vashkevich, Dr. Sci. (Eng.), Assoc. Prof., Prof. of Computer Engineering Department, Belarusian State University of Informatics and Radioelectronics.
E-mail: vashkevich@bsuir.by

УДК 004.67
<https://doi.org/10.37661/1816-0301-2026-23-1-88-104>

Поступила в редакцию | Received 16.01.2026
Подписана в печать | Accepted 06.02.2026
Опубликована | Published 31.03.2026

Тональные пространства векторных языковых моделей

К. М. Черников, И. А. Суров[✉]

✉E-mail: ilya.a.surov@itmo.ru

*Университет ИТМО,
Кронверкский пр., 49А, Санкт-Петербург, 197101, Россия*

Аннотация

Цели. Тональность как положительно-отрицательное настроение – ключевой параметр алгоритмов анализа и генерации текстов. Существующие методы машинного обучения кодируют ее вычислительно трудоемким и неинтерпретируемым образом, что затрудняет развитие этого направления. Целью работы является решение данных проблем для русского языка.

Методы. Используются предобученные векторные языковые модели, кодирующие слова векторами в многомерных пространствах. В таких пространствах тональность соответствует направлению, наилучшим образом разделяющему положительный и отрицательный прототипы. Тональность слова определяется его проекцией на это направление. Добавление тонального вектора к ключевому слову задает одномерное пространство, позволяющее находить в языковой модели его положительные и отрицательные ассоциации.

Результаты. Алгоритм апробирован на машинных моделях семейств GloVe и FastText, кодирующих отдельные слова и морфемы русского языка векторами в 300-мерном пространстве. В качестве ключевых слов использовались частоупотребимые глаголы и существительные. Средняя достоверность найденных тональных ассоциаций составила 80 %.

Заключение. Результаты свидетельствуют о применимости предобученных векторных языковых моделей для быстрой и интерпретируемой работы с тональной информацией. Разработанный подход востребован в задачах объектно-ориентированного сентимент-анализа, а также в задачах машинной генерации объектно-ориентированных текстов нужной тональности. Обобщение тональной оси до тройки семантических факторов Осгуда позволяет расширить представленный метод для работы с полным спектром эмоционально-смысловой информации.

Ключевые слова: векторные языковые модели, сентимент-анализ, тональность, семантика, пространство, генерация текстов

Благодарности. Исследование выполнено за счет гранта Российского научного фонда № 23-71-01046, <https://rscf.ru/project/23-71-01046/> (код ГРНТИ 28.23.29).

Для цитирования. Черников, К. М. Тональные пространства векторных языковых моделей / К. М. Черников, И. А. Суров // Информатика. – 2026. – Т. 23, № 1. – С. 88–104. – <https://doi.org/10.37661/1816-0301-2026-23-1-88-104>.

Конфликт интересов. Авторы заявляют об отсутствии конфликта интересов.

Tonal spaces of vector language models

Kirill M. Chernikov, Илья А. Сузов✉

✉E-mail: ilya.a.surov@itmo.ru

*ITMO University,
av. Kronverkskiy, 49A, Saint Petersburg, 197101, Russia*

Abstract

Objectives. Tonality as a positive-negative mood is a key parameter of text analysis and generation algorithms. Existing machine learning methods encode tonality in computationally extensive and uninterpretable ways which hampers the development of the corresponding applications. The work aims to solve this problem for Russian language.

Methods. Pre-trained vector language models are used to encode words as vectors in multidimensional spaces. In these spaces, the tonality corresponds to a specific direction which optimally discriminates the positive and negative prototypes. The tonality of word is then determined by its projection onto this direction. Adding a tonal vector to a key word defines a one-dimensional subspace, containing its positive and negative associations.

Results. The algorithm is tested on GloVe and FastText machine language models, encoding individual Russian words and morphemes with vectors in 300-dimensional space. Commonly used verbs and nouns served as key words. The average reliability of the found tonal associations estimates as 80 %.

Conclusion. The results indicate the applicability of pre-trained vector language models for fast and interpretable working with tonal information. The developed approach is applicable for the tasks of aspect-based sentiment analysis, as well as for the machine generation of object-oriented texts with a required tonality. Generalization of the tonal axis to the triple of Osgood's semantic factors allows expanding the method to a full range of affectively-semantic information.

Keywords: vector language models, sentiment analysis, tonality, semantics, space, affective text generation

Acknowledgments. The research was funded by the Russian Science Foundation grant No. 23-71-01046, <https://rscf.ru/project/23-71-01046/> (GRNTI code 28.23.29).

For citation. Chernikov K. M., Surov I. A. *Tonal spaces of vector language models*. Informatika [Informatics], 2026, vol. 23, no. 1, pp. 88–104 (In Russ.). <https://doi.org/10.37661/1816-0301-2026-23-1-88-104>.

Conflict of interests. The authors declare of no conflict of interest.

Введение

Тональность – положительное или отрицательное настроение – важнейший параметр письменной и устной речи. Располагаясь на уровне эмоций и аффектов, тональность слова действует на слушателя или читателя в обход медленного и затратного логического мышления [1]. Поэтому тональность является важнейшим параметром социологических, экономических, культурологических, политических исследований и технологий [2]. При этом анализ тональности (сентимент-анализ) востребован не только для массивов текста в целом, но и по отношению к конкретной целевой сущности-объекту [3, 4].

Тональность письменной речи можно определять на основе частотности ключевых слов с положительной и отрицательной окраской. В последнее время такой простой и в ряде случаев весьма эффективный подход уступает место методам на основе машинного обучения. При этом обычно используются языковые модели общего назначения, дообучаемые под задачу классификации фраз, предложений и текстов на положительные, отрицательные либо нейтральные [5]. В связке с порождающими (генеративными) языковыми моделями аналогичный подход применяется для создания текстов с нужной тональностью [6–8].

Общий недостаток нейросетевых алгоритмов – непрозрачность и необъяснимость результата, характерен для больших языковых моделей. Кроме того, обучение таких алгоритмов требует большого количества данных и вычислительных мощностей. В этой связи констатируется целесообразность разработки «объяснимого ИИ», однако результаты в данном направлении пока скромны [9].

В задаче сентимент-анализа ситуация с объяснимостью теоретически значительно лучше. Пространство эмоциональных состояний имеет вполне интерпретируемую структуру, в основе которой лежат семантические факторы Осгуда: *оценка*, *сила* и *активность*, общие для большинства языков и культур [10, 11]. В алгоритмах анализа текста, однако, эта структура до недавнего времени не использовалась из-за отсутствия связи с машинными моделями языка. Такая связь недавно установлена для модели word2vec, представляющей слова английского языка векторами в 300-мерном пространстве, в котором удалось найти взаимно ортогональные направления, соответствующие всем трем факторам Осгуда [12]. Это позволило использовать модель для исследования эмоций [13], а также для интерпретируемого анализа процессной семантики [14].

Настоящая работа развивает методы интерпретируемой работы с векторными языковыми моделями для применения к конкретным сущностям, что востребовано, в частности, в задачах объектно-ориентированного сентимент-анализа. Задача состоит в построении тональных пространств, населенных положительными и отрицательными ассоциациями ключевого слова-объекта.

1. Теория и метод

Предложенный метод опирается на машинные модели естественного языка, которые представляют слова, фразы и другие языковые блоки наборами из 100–1000 вещественных чисел, образующими векторы в пространстве соответствующей размерности. Чем ближе смыслы слов, тем больше похожи соответствующие им векторы. Благодаря этому поиск ближайших векторов в окрестности данного слова позволяет находить его синонимы и семантически близкие конструкции.

1.1. Смысловые отношения и вектор тональности

Отдельные компоненты векторов определенных интерпретаций обычно не имеют, однако алгебраические отношения между векторами иногда выражают смысловые соотношения кодируемых ими слов. Пример таких отношений показан на рис. 1, где разности векторных представлений слов «дядя» и «тетя», «мужчина» и «женщина», «царь» и «царица» оказываются примерно одинаковыми [15]. Это значит, что усредненный вектор такой разности (толстая черная стрелка) соответствует понятию *мужественности*, а векторная разность слов «царь», «мужчина» и «царица», «женщина» (толстые серые стрелки) кодирует понятие *властности*.

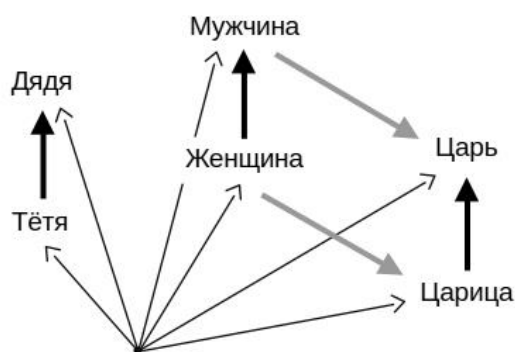


Рис. 1. Пример смысловых отношений в пространстве машинной языковой модели. Тонкие стрелки – векторные представления слов, толстые стрелки – разности словарных векторов, кодирующие понятия *мужественность* (черные) и *властность* (серые)

Fig. 1. Example of semantic relations in space of machine language model. Thin arrows – vector representation of words, thick arrows – differences of word vectors, encoding concepts "masculinity" (black) and "authority" (gray)

Знание векторов *мужественности* и *властности* позволяет подбирать слова, находящиеся в смысловых отношениях с любым данным словом. Например, если требуется властный вариант слова «самец», то к его векторному представлению следует прибавить вектор властности, после чего найти векторы, близкие к полученному результату, а если требуется его женский вариант, то из того же вектора следует вычесть вектор мужественности:

самец + властность = вожак,
 самец – мужественность = самка.

Схожие соотношения имеют место между названиями государств и их столиц, когда векторная разность имеет смысл центральности или периферийности [16]. Аналогичной алгебры можно ожидать для других качеств, достаточно полно представленных в библиотеке текстов, которая используется для обучения машинной модели.

Из множества смысловых характеристик языка задачи сентимент-анализа выделяют характеристику *тональности*, которая также должна кодироваться определенным вектором в пространстве языковой модели. Аналогично характеристике *мужественности* на рис. 1 этот вектор можно найти как разность положительных и отрицательных прототипов (рис. 2, а):

$$\mathbf{Z} = \mathbf{W}_+ - \mathbf{W}_-. \quad (1)$$

При этом векторы прототипов \mathbf{W}_+ и \mathbf{W}_- соответствуют средним от векторных представлений входящих в них слов используемой языковой модели:

$$\mathbf{W}_\pm = \sum \mathbf{w}_\pm, \quad (2)$$

где словами \mathbf{w}_+ и \mathbf{w}_- могут быть *добро, радость, благо, свет и зло, горе, ущерб, тьма* и т. п. соответственно [12].

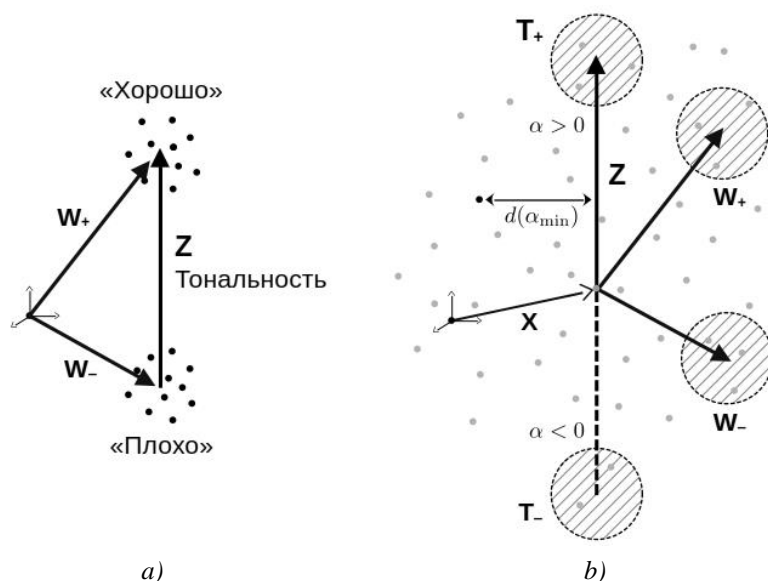


Рис. 2. Схема метода: *a*) ось тональности Z как разность положительного и отрицательного прототипов W_{\pm} (1), (2); *b*) построение тонального пространства T объекта X путем прибавления к нему тонального вектора Z , взвешенного коэффициентом α (3). Вместо Z можно использовать векторы прототипов W_{\pm} . Штриховками отмечены варианты объектно-ориентированных тональных областей

Fig. 2. Scheme of the method: a) Tonality axis Z as the difference between the positive and negative prototypes W_{\pm} (1), (2); b) construction of the tonal space T of the object X by adding to it the tonal vector Z , weighted by the coefficient α (3). Instead of Z , one can use the prototype vectors W_{\pm} . The shading marks the variants of object-oriented tonal areas

Недостатком метода нахождения Z является произвол выбора разностных пар, который может привести к искажению вектора Z и систематическим ошибкам при его последующем применении. Поэтому в качестве w_{\pm} целесообразно использовать большие массивы тонально-размеченных слов, причем вместо разности (1) удобно искать Z как направление, наилучшим образом разделяющее положительный и отрицательный классы. Для этого подходит метод линейного дискриминантного анализа [17].

1.2. Тональное пространство объекта

Положительные и отрицательные прототипы W_{\pm} описывают абсолютное добро и абсолютное зло. Объектно-ориентированная генерация или сентимент-анализ, напротив, работают с настроением текста по отношению к ключевому объекту – событию, личности или предмету [3, 4]. Например, объект *земля* в положительном настроении может описываться как плодородная почва и родина-мать, а в отрицательном – как территория, грязь и пустыня. Такой переход от нейтральной к положительной или отрицательной тональности объекта X соответствует прибавлению к нему вектора тональности Z с положительным или отрицательным знаком (рис. 2, *b*). При этом варьировать можно не только знак, но и длину прибавляемого вектора. Таким образом, тональные смещения ключевого понятия X формируются в виде

$$T = X + \alpha Z, \quad (3)$$

где Z – вектор тональности, α – положительный или отрицательный коэффициент произвольной величины.

Располагая вектором \mathbf{Z} , прямую (3) можно построить для любого объекта \mathbf{X} в используемой языковой модели. Варьирование величины α перемещает вектор \mathbf{T} от объекта \mathbf{X} по направлению \mathbf{Z} . При этом в его окрестности оказываются векторы других слов, содержащиеся в используемой машинной модели (серые точки). Такие соседи в областях положительных и отрицательных α и являются тональными ассоциациями ключевого объекта \mathbf{X} .

1.3. Поиск тональных соседей

Зная прямую (3) в пространстве языковой модели, для нахождения тональных соседей объекта \mathbf{X} можно использовать два способа.

В первом способе для каждого слова V языковой модели вычисляется его расстояние до прямой (3), которое определяется как минимум функции

$$d(\mathbf{V}, \alpha) = | \mathbf{V} - (\mathbf{X} + \alpha \mathbf{Z}) | \quad (4)$$

по параметру α (см. рис. 2, *b*). Слова V с наименьшими значениями $d(\mathbf{V}, \alpha_{\min})$ в областях $\alpha_{\min} > 0$ и $\alpha_{\min} < 0$ являются соседями слова X положительной и отрицательной тональности.

На практике данный способ дает не лучшие результаты, так как большинство слов модели получают значения α_{\min} вблизи нуля, что соответствует нейтральной тональности. Выделить на этом фоне искомые слова положительной и отрицательной тональности затруднительно. Поэтому расстояние от слов модели целесообразно вычислять не до прямой (3), а до пары точек на ее положительной и отрицательной частях (второй способ):

$$\mathbf{T}_+ = \mathbf{X} + \alpha_+ \mathbf{Z}, \quad \mathbf{T}_- = \mathbf{X} - \alpha_- \mathbf{Z}, \quad (5)$$

где $\alpha_{\pm} > 0$ есть некоторые константы. Отсечка по расстояниям от точек \mathbf{T}_{\pm} выделяет вокруг них две окрестности, как показано на рис. 2, *b*. Попадающие в эти окрестности слова V являются тональными соседями объекта X . Вместо \mathbf{Z} можно использовать векторы прототипов \mathbf{W}_{\pm} , также дающие пару окрестностей.

2. Эксперимент

Описанный алгоритм реализован для двух векторных моделей русского языка: модели семейства GloVe [18], тренированной на массиве русской художественной литературы [19], и модели семейства FastText [20], тренированной на текстах Википедии и базы Common Crawl¹. Обе модели являются 300-мерными, т. е. кодируют слова (GloVe) и морфемы (FastText) векторами из 300 вещественных чисел.

В качестве ключевых слов использовались:

- существительные *человек, собака, машина, земля*;
- глаголы *делать, говорить, смотреть, ехать*.

Векторное представление данных слов машинными моделями однотипно, что позволяет находить близкие к глаголам существительные и прилагательные, и наоборот.

¹AI at Meta. fasttext-ru-vectors // Huggingface. – 2023. – URL: <https://huggingface.co/facebook/fasttext-ru-vectors> (date of access: 07.12.2025).

2.1. Вектор тональности

Массивы положительных и отрицательных слов русского языка w_{\pm} для нахождения оси Z были взяты из словаря, размеченного по тональности для sentiment-анализа текстов [21]. При этом из общего числа 7545 слов были выбраны 2404 слова с максимальной по модулю оценкой ± 1 . Искомый вектор Z находился методом линейного дискриминантного анализа как ось, наилучшим образом разделяющая эти два класса в 300-мерном пространстве использованных моделей.

Результат проекции использованной выборки на полученную ось Z для модели FastText показан на рис. 3. Разделяемость классов, измеряемая как точность классификации, составила 67 %. Для модели GloVe вид распределений похожий, точность 75 %. Эти величины подтверждают гипотезу о наличии в 300-мерном пространстве машинных моделей оси благоприятности Z и возможности ее использования для анализа тональности.

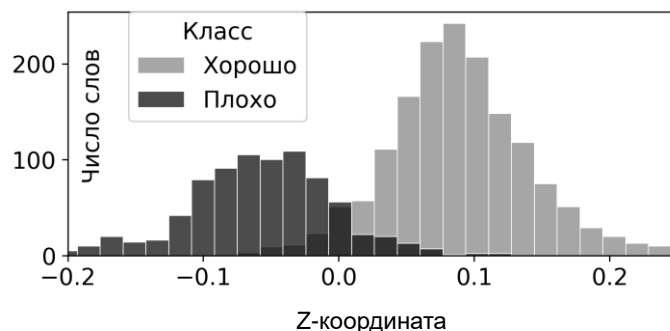


Рис. 3. Z -координаты 2404 слов с тональностью ± 1 [21], использованных для нахождения оси тональности Z в 300-мерном пространстве языковой модели FastText

Fig. 3. Z -coordinates of 2404 words with tonality ± 1 [21], used to find tonality axis Z in the 300-dimensional space of the FastText language model

2.2. Результаты

Для нахождения тональных соседей использовался показанный на рис. 2, *b* метод тональных фокусов со значениями $\alpha_{\pm} = 6$. По 15 найденных таким образом ближайших тональных вариаций к фокусам W_{\pm} и T_{\pm} для слов-объектов *смотреть*, *делать*, *собака* и *земля* приведены в таблице. Слова, попадающие и в положительные и в отрицательные классы, исключались. Средняя доля верных тональных вариаций по всем восьми ключевым словам (оценивалась на основе субъективно-экспертной оценки) составила около 80 %.

Каждый тональный сосед любого объекта характеризуется расстоянием (4) и значением α , определяющим, напротив какой части прямой (3) находится это слово. Расположение приведенных в таблице тональных соседей в осях (d, α) показано на рис. 4–7.

15 тональных вариаций, ближайших к фокусам W_{\pm} и T_{\pm} (см. рис. 2, *b*)
 в пространстве модели FastText для слов *смотреть*, *делать*, *собака* и *земля*

*15 tonal variations closest to the foci W_{\pm} and T_{\pm} (see fig. 2, *b*) in the space
 of the FastText model for the words look, do, dog and earth*

Плохой W_{-} <i>Bad W_{-}</i>	Тональность T_{-} Key T_{-}	Хороший W_{+} Good W_{+}	Тональность T_{+} Key T_{+}
<i>Смотреть</i>			
игнорировать	проклинать	посмотреть	отличный
ненавидеть	ругаться	великолепный	чудесный
жаловаться	провоцировать	просматривать	замечательно
оправдываться	глотать	увидеть	безупречный
возмущаться	подозревать	поглядеть	красивый
закрывать	давить	глядеть	сохранять
обвинять	выслушивать	просмотреть	благодарить
терпеть	вздрагивать	любоваться	послушать
упрекать	раздражать	взглянуть	удивительный
бессмысленно	запрещать	замечательный	прекрасный
срываться	всматриваться	разглядывать	поблагодарить
оглядываться	противно	превосходный	познакомиться
отбрасывать	ругать	наслаждаться	бесплатно
установиться	разрывать	приятно	ценить
покоситься	преследовать	приветствовать	оценить
<i>Делать</i>			
исправлять	поделать	сделать	обеспечивать
оправдываться	извиняться	творить	развивать
провоцировать	ругаться	создавать	сотворить
игнорировать	возмущаться	осуществлять	сохранить
убирать	заставлять	проводить	обеспечить
жаловаться	мучиться	производить	приготовить
запрещать	пытаться	научиться	осуществить
срываться	удалять	готовить	рассчитывать
валить	срывать	выполнять	реализовать
предпринимать	исправить	стараться	поддерживать
терпеть	предпринять	сохранять	помогать
отменять	прекращать	получать	практиковать
выбрасывать	бросать	организовывать	благодарить
дергаться	наказывать	совершать	ценить
закрывать	издеваться	организовать	даровать

Окончание таблицы

End of table

Плохой W_- <i>Bad W_-</i>	Тональность T_- <i>Key T_-</i>	Хороший W_+ <i>Good W_+</i>	Тональность T_+ <i>Key T_+</i>
<i>Собака</i>			
кошка	ворона	щенок	природа
крыса	мамаша	девочка	преданный
собачка	васька	красавица	награда
свинья	старуха	малышка	молодец
скотина	старушка	девушка	красота
обезьяна	тварь	хозяйка	команда
сволочь	депрессия	черепаха	порода
лошадь	помойка	женщина	скульптура
аллергия	пьяница	котенок	отличный
инфекция	бедняга	животное	мудрость
морда	табуретка	корова	мамочка
агрессия	бешенство	игрушка	доброта
истерика	сука	птичка	гармония
курица	свалка	семья	замечательный
лихорадка	колбаса	преданность	лягушка
<i>Земля</i>			
почва	империя	природа	гармония
пустыня	стройка	родина	мудрость
планета	свинья	поляна	усадьба
свалка	древесина	вселенная	культура
равнина	тайга	страна	обитель
скотина	провинция	собственность	атмосфера
картошка	война	вода	семья
помойка	бездна	недвижимость	долина
оккупация	лихорадка	радуга	инфраструктура
степь	колония	столица	энергетика
окраина	мостовая	основа	квартира
стихия	темнота	красота	дорога
трава	коленка	архитектура	площадка
катастрофа	тюрьма	материя	красавица
территория	местность	энергия	цивилизация

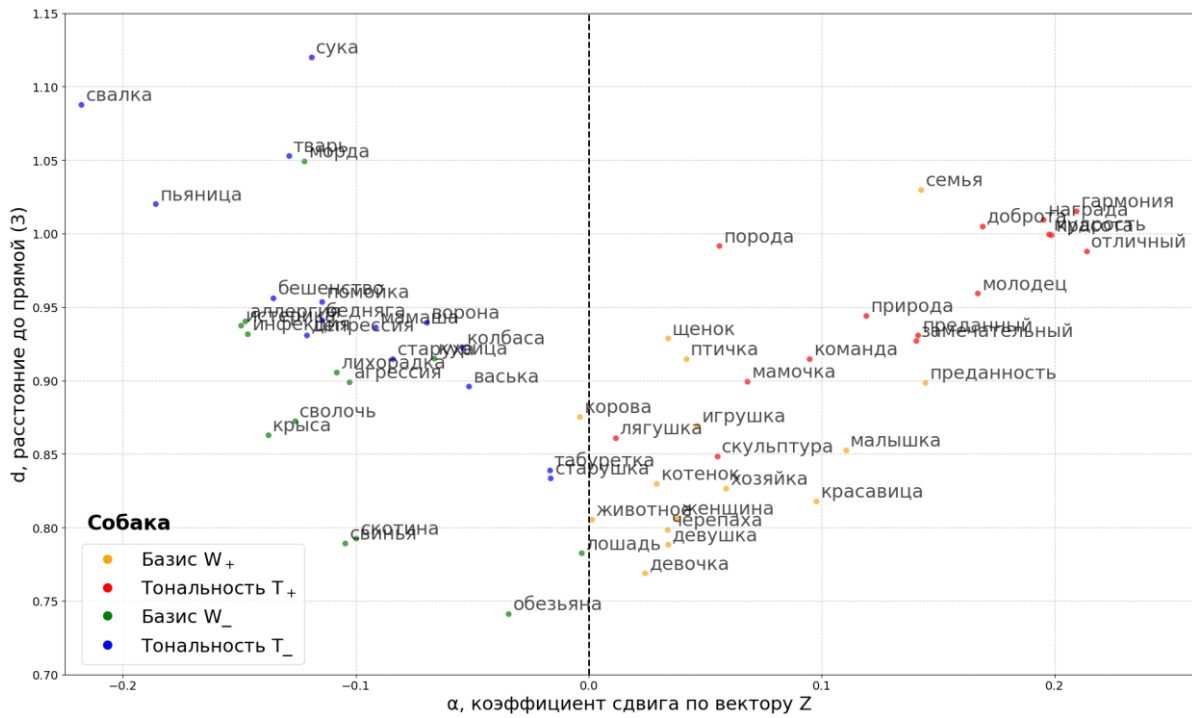


Рис. 6. Тональные соседи существительного *собака* (таблица) в осях (d, α)

Fig. 6. Tonal neighbors of the noun *dog* (table) in the axes (d, α)

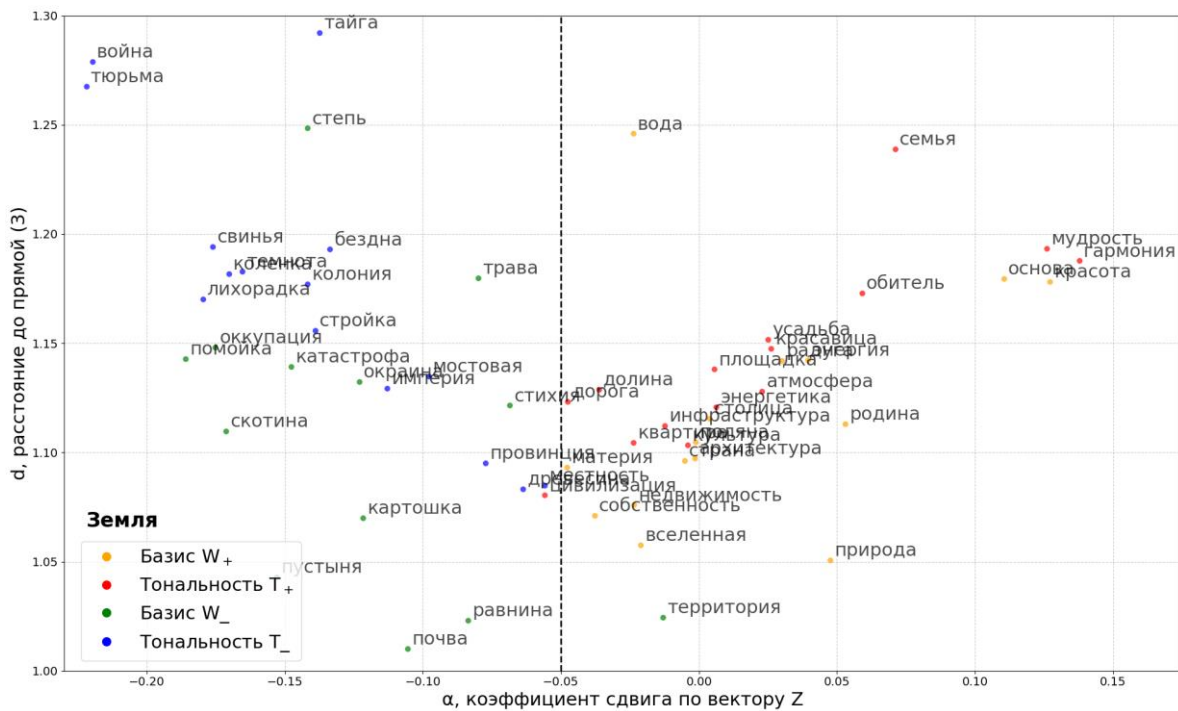


Рис. 7. Тональные соседи существительного *земля* (таблица) в осях (d, α)

Fig. 7. Tonal neighbors of the noun *earth* (table) in the axes (d, α)

На рис. 4–7 видно, что найденные тональные вариации располагаются напротив положительной и отрицательной частей прямой (3), как и ожидалось теоретически. Граница между тональными соседями, однако, не всегда проходит на значении $\alpha=0$. Из восьми рассмотренных объектов этот эффект наблюдается для объекта *земля* на рис. 7, а также для глагольного объекта *ехать*. В первом случае граница проходит около $\alpha \approx -0,05$, во втором – около $\alpha \approx 0,05$.

Таблица и рис. 4–7 показывают, что соседи к тональным фокусам \mathbf{W}_{\pm} и \mathbf{T}_{\pm} одного знака (красный и желтый, зеленый и синий) практически эквивалентны. Устойчивых смысловых различий между ними не выявлено. Геометрически это значит, что векторы \mathbf{W}_{+} и \mathbf{W}_{-} параллельны и антипараллельны вектору тональности \mathbf{Z} и могут быть использованы вместо него, что несколько упрощает представленный метод. Такой эффект имеет место в пространствах обеих рассмотренных моделей.

Результаты, аналогичные представленным в таблице и на рис. 4–7, получены для примерно 10 других частоупотребляемых существительных и глаголов русского языка (*яблоко, работа, относиться, искать* и др.) с похожими показателями качества, что свидетельствует о достаточно высокой надежности разработанного метода. Трудность более масштабной проверки обусловлена отсутствием размеченных нужным образом словарей и надежных методов автоматического определения объектно-ориентированной тональности (чем и определяется новизна поставленной задачи). Создание таких разметок, а также сравнение с другими методами сентимент-анализа [3, 4] могут стать объектами дальнейших исследований.

===== Заключение

Полученный результат подтверждает гипотезу о наличии в пространствах машинных языковых моделей объектно-ориентированных тональных подпространств. В задачах сентимент-анализа [3, 4] эти подпространства могут быть использованы для получения простых метрик объектно-ориентированной тональности. Кроме того, характерные экземпляры таких пространств могут выполнять функцию признаков, присутствие которых указывает на соответствующую тональность текста по отношению к ключевому объекту. Тональные подпространства также предоставляют вспомогательное средство для генерации эмоционально окрашенных текстов [22, 23]. Близость к объектно-ориентированным подпространствам не означает требуемой при этом логической связности, однако позволяет сформировать или уточнить исходный словарь для составления текста желаемой тональности по отношению к целевому объекту.

В обоих случаях основным достоинством разработанного подхода является его вычислительная эффективность. В отличие от больших языковых моделей, требующих значительных вычислительных ресурсов, предложенный метод ограничивается простой векторной арифметикой и поиском ближайших соседей. В этой связи разработанный подход применим в условиях жестких ограничений по вычислительной мощности или в задачах, требующих сверхбыстрой предварительной обработки текста (например, в реальном времени). Кроме того, подход обеспечивает интерпретируемость создаваемых на его основе алгоритмов, которой часто не хватает методам глубокого обучения.

Основным недостатком представленного подхода является его ограниченность статическими векторными моделями, не учитывающими контекстуальную многозначность слов естественного языка. Для работы с динамическими моделями (семейств

BERT, GPT и др.) целесообразно построение контекстно-зависимых тональных пространств. Возможность такого развития может стать объектом дальнейших исследований.

Другим направлением развития разработанного метода является задача обобщенного сентимент-анализа, различающего не только положительность и отрицательность, но и другие факторы эмоциональных состояний. Для этого кроме фактора оценки можно использовать остальные два семантических фактора Осгуда, а именно *активность* и *потенциал* [10, 11]. Соответствующие векторы и их линейные комбинации в пространстве машинной языковой модели [12, 13] прибавляются к ключевому слову аналогично схеме на рис. 2, *b*. В результате вокруг ключевого слова формируется не одномерное, а трехмерное подпространство, охватывающее весь спектр его эмоционально-смысловых отношений [13, 24]. Психологически интерпретируемая алгебра этих отношений могла бы дополнить методы векторного логического вывода [25–27], в том числе с использованием падежно-смысловых структур [28].

Помимо отмеченных областей представленный подход востребован в психологии, где машинные языковые модели позволяют во многом заменить и (или) дополнить дорогостоящие и трудно воспроизводимые натурные эксперименты [13, 29, 30]. Кроме того, объектно-ориентированные семантические пространства могут стать основой для информативных и достоверных измерений общественного мнения по конкретному вопросу [31]. Для приложения к этим задачам разработанный подход целесообразно адаптировать для работы с современными архитектурами языковых моделей.

Вклад авторов. К. М. Черников осуществил разработку метода, компьютерную реализацию алгоритма, проведение экспериментов, визуализацию и интерпретацию результатов. И. А. Суров придумал концепцию исследования, разработал метод и реализовал интерпретацию результатов, подготовил текст к публикации.

Список использованных источников

1. Clynes, M. The communication of emotion: Theory of sentics / M. Clynes ; ed.: R. Plutchik, H. Kellerman // *Theories of Emotion*. – Academic Press, 1980. – P. 271–301. – <https://doi.org/10.1016/B978-0-12-558701-3.50017-X>.
2. Smetanin, S. The applications of sentiment analysis for Russian language texts: Current challenges and future perspectives / S. Smetanin // *IEEE Access*. – 2020. – Vol. 8. – P. 110693–110719. – <https://doi.org/10.1109/access.2020.3002215>.
3. Brauwers, G. A survey on aspect-based sentiment classification / G. Brauwers, F. Frasincar // *ACM Computing Surveys*. – 2021. – Vol. 55. – P. 1–37. – <https://doi.org/10.1145/3503044>.
4. A summary of aspect-based sentiment analysis / S. Fan, J. Yao, Y. Sun, Y. Zhan // *Journal of Physics: Conference Series*. – 2020. – Vol. 1624, no. 2. – URL: <https://iopscience.iop.org/article/10.1088/1742-6596/1624/2/022051> (date of access: 07.12.2025). – <https://doi.org/10.1088/1742-6596/1624/2/022051>.
5. Tang, D. Deep learning for sentiment analysis: successful approaches and future challenges / D. Tang, B. Qin, T. Liu // *WIREs Data Mining and Knowledge Discovery*. – 2015. – Vol. 5, no. 6. – P. 292–303. – <https://doi.org/10.1002/widm.1171>.
6. van der Sluis, I. Affective text: Generation strategies and emotion measurement issues / I. van der Sluis, C. Mellish, G. Doherty // *Proc. of the Twenty-Fourth Intern. Florida Artificial Intelligence Research Society Conf.*, Palm Beach, Florida, USA, 18–20 May 2011. – Palm Beach, 2011. – P. 123–128.

7. Adapting a language model for controlled affective text generation / I. Singh, A. Barkati, T. Goswamy, A. Modi // Proc. of the 28th Intern. Conf. on Computational Linguistics, Barcelona, Spain (Online), Dec. 2020. – Barcelona, 2020. – P. 2787–2801. – <https://doi.org/10.48550/arXiv.2011.04000>.
8. Nie, G. A review of affective generation models / G. Nie, Y. Zhan. – 2022. – URL: <https://arxiv.org/pdf/2202.10763> (date of access: 07.12.2025). – <https://doi.org/10.48550/arXiv.2202.10763>.
9. Explainable AI / V. C. Storey, R. Lukyanenko, W. Maass, J. Parsons // Communications of the ACM. – 2022. – Vol. 65, no. 4. – P. 27–29. – <https://doi.org/10.1145/3490699>.
10. Tanaka, Y. Cross-culture, cross-concept, and cross-subject generality of affective meaning systems / Y. Tanaka, C. E. Osgood // Journal of Personality and Social Psychology. – 1965. – Vol. 2, no. 2. – P. 143–153. – <https://doi.org/10.1037/h0022392>.
11. Osgood, C. E. On the whys and wherefores of E, P, and A / C. E. Osgood // Journal of Personality and Social Psychology. – 1969. – Vol. 12, no. 3. – P. 194–199. – <https://doi.org/10.1037/h0027715>.
12. Сузов, И. А. Открытие черного ящика: извлечение семантических факторов Осгуда из языковой модели word2vec / И. А. Сузов // Информатика и автоматизация. – 2022. – Т. 21, № 5. – С. 916–936. – <https://doi.org/10.15622/ia.21.5.3>.
13. Груздева, А. С. Машинно-семантический дифференциал: картирование эмоций посредством векторных языковых моделей / А. С. Груздева, И. А. Сузов // Ученые записки Института психологии РАН. – 2025. – Т. 5, № 4. – С. 86–99. – https://doi.org/10.38098/proceedings_2025_05_04_10.
14. Surov, I. A. Process-semantic analysis of words and texts / I. A. Surov // Artificial Intelligence in Models, Methods and Applications / ed.: O. Dolinina [et al.]. – Cham : Springer, 2023. – С. 247–260. – https://doi.org/10.1007/978-3-031-22938-1_17.
15. Mikolov, T. Linguistic regularities in continuous space word representations / T. Mikolov, W. Yih, G. Zweig // Proc. of the 2013 Conf. of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Atlanta, Georgia, 9–14 June 2013. – Atlanta, 2013. – P. 746–751.
16. Distributed representations of words and phrases and their compositionality / T. Mikolov, I. Sutskever, K. Chen [et al.] // NIPS'13: Proc. of the 26th Intern. Conf. on Neural Information Processing Systems, Lake Tahoe, Nevada, 5–10 Dec. 2013. – Lake Tahoe, 2013. – Vol. 2. – P. 3111–3119.
17. McLachlan, G. J. Discriminant Analysis and Statistical Pattern Recognition / G. J. McLachlan. – N. Y. : Wiley, 2004. – 526 p.
18. Pennington, J. Glove: Global vectors for word representation / J. Pennington, R. Socher, C. D. Manning // Proc. of the 2014 Conf. on Empirical Methods in Natural Language Processing (EMNLP), Doha, Qatar, 25–29 Oct. 2014. – Doha, 2014. – P. 1532–1543.
19. Kukushkin, A. Navec_hudlit_v1_12B_500K_300d_100q.tar / A. Kukushkin. – 2023. – URL: <https://github.com/natasha/navec> (date of access: 07.12.2025).
20. Enriching word vectors with subword information / P. Bojanowski, E. Grave, A. Joulin, T. Mikolov // Transactions of the Association for Computational Linguistics. – 2017. – Vol. 5. – P. 135–146. – <https://doi.org/10.48550/ARXIV.1607.04606>.
21. Koltsova, O. Y. An opinion word lexicon and a training dataset for Russian sentiment analysis of social media / O. Y. Koltsova, S. V. Alexeeva, S. N. Kolcov // Computational Linguistics and Intellectual Technologies. – 2016. – Vol. 2016. – P. 277–287.
22. From extraction to generation: Multimodal emotion-cause pair generation in conversations / H. Ma, J. Yu, F. Wang [et al.] // IEEE Transactions on Affective Computing. – 2025. – Vol. 16, no. 2. – P. 586–597. – <https://doi.org/10.1109/TAFFC.2024.3446646>.

23. Text-based fine-grained emotion prediction / G. Singh, D. Brahma, P. Rai, A. Modi // *IEEE Transactions on Affective Computing*. – 2024. – Vol. 15, no. 2. – P. 405–416. – <https://doi.org/10.1109/TAFFC.2023.3298405>.
24. Surov, I. A. Quantum core affect. Color-emotion structure of semantic atom / I. A. Surov // *Frontiers in Psychology*. – 2022. – Vol. 13. – <https://doi.org/10.3389/fpsyg.2022.838029>.
25. Widdows, D. Should semantic vector composition be explicit? Can it be linear? / D. Widdows, K. Howell, T. Cohen // *Proc. of the Workshop on Semantic Spaces at the Intersection of NLP, Physics, and Cognitive Science, Groningen, The Netherlands, 14–18 June 2021*. – Groningen, 2021. – P. 76–86. – <https://doi.org/10.48550/arXiv.2104.06555>.
26. Schlegel, K. A comparison of vector symbolic architectures / K. Schlegel, P. Neubert, P. Protzel // *Artificial Intelligence Review*. – 2022. – Vol. 55, no. 6. – P. 4523–4555. – <https://doi.org/10.1007/s10462-021-10110-3>.
27. Vector symbolic architectures as a computing framework for emerging hardware / D. Kleyko, M. Davies, E. P. Frady [et al.] // *Proceedings of the IEEE*. – 2022. – Vol. 110, no. 10. – P. 1538–1571. – <https://doi.org/10.1109/JPROC.2022.3209104>.
28. Суров, И. А. Геометрическая семиотика падежей русского языка / И. А. Суров // *Вестник Московского государственного университета. Гуманитарные науки*. – 2026. (В печати)
29. Predicting survey responses: How and why semantics shape survey statistics on Organizational Behaviour / J. K. Arnulf, K. R. Larsen, Ø. L. Martinsen, C. H. Bong // *PLoS ONE*. – 2014. – Vol. 9, no. 9. – <https://doi.org/10.1371/journal.pone.0106361>.
30. Arnulf, J. K. Measuring the menu, not the food: “psychometric” data may instead measure “lingometrics” (and miss its greatest potential) / J. K. Arnulf, U. H. Olsson, K. Nimon // *Frontiers in Psychology*. – 2024. – Vol. 15. – P. 1308098. – <https://doi.org/10.3389/fpsyg.2024.1308098>.
31. Lukyanenko, R. Integrating LLMs and psychometrics: Global construct validity / R. Lukyanenko, K. R. Larsen // *Forty-Fifth Intern. Conf. on Information Systems, Bangkok, Thailand, 15–18 Dec. 2024*. – Bangkok, 2024. – URL: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=5341306 (date of access: 07.12.2025).

References

1. Clynes M. The communication of emotion: Theory of sentics. *Theories of Emotion*. In R. Plutchik, H. Kellerman (eds.). Academic Press, 1980, pp. 271–301. <https://doi.org/10.1016/B978-0-12-558701-3.50017-X>.
2. Smetanin S. The applications of sentiment analysis for Russian language texts: Current challenges and future perspectives. *IEEE Access*, 2020, vol. 8, pp. 110693–110719. <https://doi.org/10.1109/access.2020.3002215>.
3. Brauwerters G., Frasincar F. A survey on aspect-based sentiment classification. *ACM Computing Surveys*, 2021, vol. 55, pp. 1–37. <https://doi.org/10.1145/3503044>.
4. Fan S., Yao J., Sun Y., Zhan Y. A summary of aspect-based sentiment analysis. *Journal of Physics: Conference Series*, 2020, vol. 1624, no. 2. Available at: <https://iopscience.iop.org/article/10.1088/1742-6596/1624/2/022051> (accessed 07.12.2025). <https://doi.org/10.1088/1742-6596/1624/2/022051>.
5. Tang D., Qin B., Liu T. Deep learning for sentiment analysis: successful approaches and future challenges. *WIREs Data Mining and Knowledge Discovery*, 2015, vol. 5, no. 6, pp. 292–303. <https://doi.org/10.1002/widm.1171>.

6. van der Sluis I., Mellish C., Doherty G. Affective Text: Generation Strategies and Emotion Measurement Issues. *Proceedings of the Twenty-Fourth International Florida Artificial Intelligence Research Society Conference, Palm Beach, Florida, USA, 18–20 May 2011*. Palm Beach, 2011, pp. 123–128.
7. Singh I., Barkati A., Goswamy T., Modi A. Adapting a language model for controlled affective text generation. *Proceedings of the 28th International Conference on Computational Linguistics, Barcelona, Spain (Online), December 2020*. Barcelona, 2020, pp. 2787–2801. <https://doi.org/10.48550/arXiv.2011.04000>.
8. Nie G., Zhan Y. *A review of affective generation models*, 2022. Available at: <https://arxiv.org/pdf/2202.10763> (accessed 07.12.2025). <https://doi.org/10.48550/arXiv.2202.10763>.
9. Storey V. C., Lukyanenko R., Maass W., Parsons J. Explainable AI. *Communications of the ACM*, 2022, vol. 65, no. 4, pp. 27–29. <https://doi.org/10.1145/3490699>.
10. Tanaka Y., Osgood C. E. Cross-culture, cross-concept, and cross-subject generality of affective meaning systems. *Journal of Personality and Social Psychology*, 1965, vol. 2, no. 2, pp. 143–153. <https://doi.org/10.1037/h0022392>.
11. Osgood C. E. On the whys and wherefores of E, P, and A. *Journal of Personality and Social Psychology*, 1969, vol. 12, no. 3, pp. 194–199. <https://doi.org/10.1037/h0027715>.
12. Surov I. A. *Opening the black box: finding Osgood's semantic factors in word2vec space*. *Informatika i avtomatizacija [Informatics and Automation]*, 2022, vol. 21, no. 5, pp. 916–936 (In Russ.). <https://doi.org/10.15622/ia.21.5.3>.
13. Gruzdeva A. S., Surov I. A. *Machine-semantic differential: emotion mapping through vector language models*. *Uchenye zapiski Instituta psihologii Rossijskoj akademii nauk [Proceedings of the Institute of Psychology of the Russian Academy of Sciences]*, 2025, vol. 5, no. 4, pp. 86–99 (In Russ.). https://doi.org/10.38098/proceedings_2025_05_04_10.
14. Surov I. A. Process-semantic analysis of words and texts. *Artificial Intelligence in Models, Methods and Applications*. In O. Dolinina, I. Bessmertny, A. Brovko, V. Kreinovich, V. Pechenkin, ..., V. Zhmud (eds.). Cham, Springer, 2023, pp. 247–260. https://doi.org/10.1007/978-3-031-22938-1_17.
15. Mikolov T., Yih W., Zweig G. Linguistic regularities in continuous space word representations. *Proceedings of the 2013 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Atlanta, Georgia, 9–14 June 2013*. Atlanta, 2013, pp. 746–751.
16. Mikolov T., Sutskever I., Chen K., Corrado G., Dean J. Distributed representations of words and phrases and their compositionality. *NIPS'13: Proceedings of the 26th International Conference on Neural Information Processing Systems, Lake Tahoe, Nevada, 5–10 December 2013*. Lake Tahoe, 2013, vol. 2, pp. 3111–3119.
17. McLachlan G. J. *Discriminant Analysis and Statistical Pattern Recognition*. New York, Wiley, 2004, 526 p.
18. Pennington J., Socher R., Manning C. D. Glove: Global vectors for word representation. *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP), Doha, Qatar, 25–29 October 2014*. Doha, 2014, pp. 1532–1543.
19. Kukushkin A. *Navec_hudlit_v1_12B_500K_300d_100q.tar*, 2023. Available at: <https://github.com/natasha/navec> (accessed 07.12.2025).
20. Bojanowski P., Grave E., Joulin A., Mikolov T. Enriching word vectors with subword information. *Transactions of the Association for Computational Linguistics*, 2017, vol. 5, pp. 135–146. <https://doi.org/10.48550/ARXIV.1607.04606>.

21. Koltsova O. Y., Alexeeva S. V., Kolcov S. N. An opinion word lexicon and a training dataset for Russian sentiment analysis of social media. *Computational Linguistics and Intellectual Technologies*, 2016, vol. 2016, pp. 277–287.
22. Ma H., Yu J., Wang F., Cao H., Xia R. From extraction to generation: Multimodal emotion-cause pair generation in conversations. *IEEE Transactions on Affective Computing*, 2025, vol. 16, no. 2, pp. 586–597. <https://doi.org/10.1109/TAFFC.2024.3446646>.
23. Singh G., Brachma D., Rai P., Modi A. Text-based fine-grained emotion prediction. *IEEE Transactions on Affective Computing*, 2024, vol. 15, no. 2, pp. 405–416. <https://doi.org/10.1109/TAFFC.2023.3298405>.
24. Surov I. A. Quantum core affect. Color-emotion structure of semantic atom. *Frontiers in Psychology*, 2022, vol. 13. <https://doi.org/10.3389/fpsyg.2022.838029>.
25. Widdows D., Howell K., Cohen T. Should semantic vector composition be explicit? Can it be linear? *Proceedings of the Workshop on Semantic Spaces at the Intersection of NLP, Physics, and Cognitive Science, Groningen, The Netherlands, 14–18 June 2021*. Groningen, 2021, pp. 76–86. <https://doi.org/10.48550/arXiv.2104.06555>.
26. Schlegel K., Neubert P., Protzel P. A comparison of vector symbolic architectures. *Artificial Intelligence Review*, 2022, vol. 55, no. 6, pp. 4523–4555. <https://doi.org/10.1007/s10462-021-10110-3>.
27. Kleyko D., Davies M., Frady E. P., Kanerva P., Kent S. J., Olshausen B. A. Vector symbolic architectures as a computing framework for emerging hardware. *Proceedings of the IEEE*, 2022, vol. 110, no. 10, pp. 1538–1571. <https://doi.org/10.1109/JPROC.2022.3209104>.
28. Surov I. A. *Geometrical semiotics of Russian cases*. Vestnik Moskovskogo gosudarstvennogo universiteta. Gumanitarnye nauki [*Vestnik of Moscow State Linguistic University. Humanities*], 2026. (In press). (In Russ.).
29. Arnulf J. K., Larsen K. R., Martinsen Ø. L., Bong C. H. Predicting survey responses: How and why semantics shape survey statistics on Organizational Behaviour. *PLoS ONE*, 2014, vol. 9, no. 9. <https://doi.org/10.1371/journal.pone.0106361>.
30. Arnulf J. K., Olsson U. H., Nimon K. Measuring the menu, not the food: “psychometric” data may instead measure “lingometrics” (and miss its greatest potential). *Frontiers in Psychology*, 2024, vol. 15, p. 1308098. <https://doi.org/10.3389/fpsyg.2024.1308098>.
31. Lukyanenko R., Larsen K. R. Integrating LLMs and Psychometrics: Global Construct Validity. *Forty-Fifth International Conference on Information Systems, Bangkok, Thailand, 15–18 December 2024*. Bangkok, 2024. Available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=5341306 (accessed 07.12.2025).

Информация об авторах

Черников Кирил Михайлович, студент факультета технологий искусственного интеллекта, Университет ИТМО.
E-mail: kmchernikov@itmo.ru

Суров Илья Алексеевич, кандидат физико-математических наук, доцент, старший научный сотрудник факультета технологий искусственного интеллекта, Университет ИТМО.
E-mail: ilya.a.surov@itmo.ru

Information about the authors

Kirill M. Chernikov, Undergraduate of Department of Artificial Intelligence Technologies, ITMO University.
E-mail: kmchernikov@itmo.ru

Ilya A. Surov, Cand. Sci. (Phys.-Math.), Assoc. Prof., Senior Researcher of Department of Artificial Intelligence Technologies, ITMO University.
E-mail: ilya.a.surov@itmo.ru

Правила для авторов

С подробной версией и примерами оформления можно ознакомиться на сайте журнала <http://inf.grid.by> при подаче статьи.

Параметры форматирования: основной текст – Times New Roman, кегль основного шрифта – 12 пт, кегль подрисованных подписей и названия таблиц – 10 пт, междустрочный интервал – точно 15 пт, абзац – 1 см, поля – по 2,5 см со всех сторон, текст с переносами, формулы, рисунки, таблицы отбиваются пробелами от основного текста.

Тип статьи:	Объем	Аннотация	Количество литературных ссылок
Оригинальная	не больше 20 стр.	структурирована	не менее 10
Обзорная	не меньше 20 стр.	структурировать не надо	не более 30

Авторы статей о результатах работ, проведенных в научных учреждениях, должны предоставить в редакцию акт экспертизы.

Титульная страница статьи:

УДК	Индекс по Универсальной десятичной классификации
Название на англ. и рус. языках	- не используются аббревиатуры; - без переносов
Сведения об авторах	- инициалы и фамилии всех авторов – на рус. языке; - имя, инициал и фамилия – на англ. языке; - e-mail ответственного лица; - полное наименование учреждений с указанием почтового адреса
Аннотация Abstract	- 150–250 слов; - включает подразделы: Цели / Objectives, Методы / Methods, Результаты / Results, Заключение / Conclusion; - аббревиатуры расшифровываются
Ключевые слова Keywords	- 5–10 слов и несложных словосочетаний; - нежелательно использовать аббревиатуры
Благодарности Acknowledgements	- признательность коллегам за помощь; - благодарность за финансовую поддержку
Конфликт интересов Conflict of interest	- информация о конфликте интересов. При отсутствии такового формулировка следующая: Автор заявляет об отсутствии конфликта интересов

Текст статьи:

Разделы и подразделы в тексте оформляются отдельной строкой полужирным очертанием.

Формулы, рисунки, таблицы в статье нумеруются в соответствии с порядком их упоминания в тексте, ссылки на них обязательны. Рисунки должны быть хорошего качества с четкими надписями и обозначениями. Цветные иллюстрации печатаются только в том случае, когда это необходимо для понимания излагаемого материала. Перевод подрисованной подписи и пояснений к рисунку, а также перевод названия таблицы, заголовки строк или столбцов располагаются курсивом после русскоязычной версии. Для рисунков и таблиц указывается источник (ссылка на материалы других статей, составлено автором и др.).

Набор формул выполняется в формульном редакторе Math Type. Прямым шрифтом набираются: греческие и русские буквы; математические символы (\sin , \lg , ∞); символы химических элементов (C, Cl, CHCl₃); цифры (римские и арабские); индексы (верхние и нижние), являющиеся сокращениями слов. Курсивом набираются латинские буквы, символы физических величин (в том числе и в индексе).

В статье обязателен раздел Вклад авторов, если автор не один.

Список использованной литературы оформляется в соответствии с требованиями Высшей аттестационной комиссии Республики Беларусь (ГОСТ 7.1–2024). Номер литературной ссылки в тексте указывается по порядку. Ссылаться на неопубликованные работы не допускается. Ссылки на учебно-методическую литературу, ГОСТы, авторефераты, статистические отчеты в список не включаются, а оформляются в виде сносок (с подробными рекомендациями можно ознакомиться на сайте журнала в разделе Правила для авторов).

Отдельно оформляется References со следующей структурой: авторы (транслитерация), транслитерированное название монографии, *Перевод названия монографии на английский язык*. Выходные данные с обозначениями на английском языке. От транслитераций названий статей можно отказаться.

В разделе Информация об авторах (Information about the authors) приводятся ФИО авторов полностью, ученая степень, звание, должность, название организации, ORCID (при наличии).

Все поступающие в редакцию рукописи проходят предварительную проверку на соответствие Правилам для авторов. Статья может быть возвращена автору на доработку с просьбой устранить недостатки или дополнить информацию.

Редакция оставляет за собой право на редакционные изменения, не искажающие основное содержание статьи.

ИНДЕКСЫ

00827

для индивидуальных
подписчиков

008272

для предприятий
и организаций

ISSN 1816-0301 (Print)



9 771816 030000