

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

ИНФОРМАТИКА INFORMATICS

TOM 22
VOL.

2 | 2025

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 докторских 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 докторских 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.

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

ОБЪЕДИНЕНИЙ ИНСТИТУТ ПРОБЛЕМ ИНФОРМАТИКИ
НАЦИОНАЛЬНОЙ АКАДЕМИИ НАУК БЕЛАРУСИ

ИНФОРМАТИКА

Informatika

Том 22, № 2, апрель-июнь 2025

Ежеквартальный научный журнал

Издаётся с января 2004 г.

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

Главный редактор

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

Заместитель главного редактора

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

Редакционная коллегия

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

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

Бибило Петр Николаевич, д-р техн. наук, проф., ОИПИ НАН Беларусь (Минск, Беларусь)

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

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

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

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

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

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

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

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

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

Чернявский Александр Федорович, д-р техн. наук, проф., академик НАН Беларусь, НИИ ПФП
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 22, № 2, апрель-июнь 2025

Ответственный за выпуск *Мойсейчик Светлана Сергеевна*

Редактор *Гончаренко Галина Борисовна*

Компьютерная верстка *Бутевич Ольга Борисовна*

Сдано в набор 12.05.2025. Подписано в печать 19.06.2025. Формат 60×84 1/8. Бумага офсетная.
Гарнитура Таймс. Ризография. Усл. печ. л. 12,8. Уч.-изд. л. 12,5. Тираж 40 экз. Заказ 3.

Государственное научное учреждение «Объединенный институт проблем информатики
Национальной академии наук Беларусь».

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. 22, no. 2, April-June 2025

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, D. Sc. (Milit.), Ph. D. (Eng.), Assoc. Prof., UIIP NASB (Minsk, Belarus)

Deputy Editor-in-Chief

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

Editorial Board

Sergey V. Ablameyko, D. Sc. (Eng.), Prof., Academician of NASB, BSU (Minsk, Belarus)

Uladimir V. Anishchanka, Ph. D. (Eng.), Assoc. Prof., SoftClub Ltd. (Minsk, Belarus)

Petr N. Bibilo, D. Sc. (Eng.), Prof., UIIP NASB (Minsk, Belarus)

Mikhail N. Bobov, D. Sc. (Eng.), Prof., BSUIR (Minsk, Belarus)

Alexandre B. Dolgui, D. Sc. (Eng.), Prof., IMT Atlantique (Nantes, France)

Alexander N. Dudin, D. Sc. (Phys.-Math.), Prof., BSU (Minsk, Belarus)

Alexey A. Karpov, D. Sc. (Eng.), Assoc. Prof., SPII RAS (Saint Petersburg, Russia)

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

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

Viktor V. Krasnoproschin, D. Sc. (Eng.), Prof., BSU (Minsk, Belarus)

Alexander M. Krot, D. Sc. (Eng.), Prof., UIIP NASB (Minsk, Belarus)

Nikolai A. Likhoded, D. Sc. (Phys.-Math.), Prof., BSU (Minsk, Belarus)

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

Valery A. Sklyarov, D. Sc. (Eng.), Prof., University of Aveiro (Aveiro, Portugal)

Yuri N. Sotskov, D. Sc. (Phys.-Math.), Prof., UIIP NASB (Minsk, Belarus)

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

Ljudmila D. Cheremisinova, D. Sc. (Eng.), Prof., UIIP NASB (Minsk, Belarus)

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

Vyacheslav N. Yarmolik, D. Sc. (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. 22, no. 2, April-June 2025

Issue Head *Sviatlana S. Maiseichyk*

Editor *Halina B. Hancharenka*

Computer Imposition *Volha B. Butsevich*

Sent for press 12.05.2025. Output 19.06.2025. Format 60×84 1/8. Offset paper. Headset Times. Riesography.
Printed sheets 12,8. Publisher's signatures 12,5. Circulation 40 copies. Order 3.

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, Surganova Str., 220012, Minsk, Belarus.

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

СОДЕРЖАНИЕ

ИНТЕЛЛЕКТУАЛЬНЫЕ СИСТЕМЫ

- Кругликов С. В., Давыдовский А. Г.** Моделирование рисков развития «умного города» как сложной социотехнической системы 7

ОБРАБОТКА СИГНАЛОВ, ИЗОБРАЖЕНИЙ, РЕЧИ, ТЕКСТА И РАСПОЗНАВАНИЕ ОБРАЗОВ

- Wu X., Ablameyko S. V.** Efficient detection of building in remote sensing images using an improved YOLOv10 network 33

МАТЕМАТИЧЕСКОЕ МОДЕЛИРОВАНИЕ

- Дудин А. Н., Дудина О. С., Дудин С. А.** Анализ полуоткрытой сети массового обслуживания с адаптацией скоростей обслуживания к скоростям поступления запросов 48

ЛОГИЧЕСКОЕ ПРОЕКТИРОВАНИЕ

- Ярмолик В. Н., Шевченко Н. А., Петровская В. В.** Масштабирование управляемых вероятностных тестов с применением матриц Адамара 63

БИОИНФОРМАТИКА

- Яцков Н. Н., Сарнацкий Д. Д., Скаакун В. В., Гринев В. В.** Программный комплекс для имитационного моделирования сайтов одноклеточного генетического полиморфизма 81

ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ

- Жидович А. А., Лубенько А. А., Войтешенко И. С.** Прототипирование системы беспарольного атрибутивного доступа к информационным ресурсам с использованием eID-карты Республики Беларусь и FIDO2-аутентификации 95

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

CONTENTS

INTELLIGENT SYSTEMS

- Kruglikov S. V., Davydovskii A. G.** Modelling the risks of "smart city" development as a complex sociotechnical system 7

SIGNAL, IMAGE, SPEECH, TEXT PROCESSING AND PATTERN RECOGNITION

- Wu X., Ablameyko S. V.** Efficient detection of building in remote sensing images using an improved YOLOv10 network 33

MATHEMATICAL MODELING

- Dudin A. N., Dudina O. S., Dudin S. A.** Analysis of a semi-open queueing network with adaptation of service rates to the rates of arriving requests 48

LOGICAL DESIGN

- Yarmolik V. N., Shauchenka M. A., Petrovskaya V. V.** Scaling controlled random tests based on Hadamard matrices 63

BIOINFORMATICS

- Yatskou M. M., Sarnatski D. D., Skakun V. V., Grinev V. V.** Software complex for simulation modelling of single nucleotide genetic polymorphism sites 81

INFORMATION TECHNOLOGIES

- Zhidovich A. A., Lubenko A. A., Vojteshenko I. S.** Prototyping of the system of passwordless attributive access to information resources using eID-card of the Republic of Belarus and FIDO2-authentication 95

ИНТЕЛЛЕКТУАЛЬНЫЕ СИСТЕМЫ INTELLIGENT SYSTEMS



УДК 338.49, 338.24
DOI: 10.37661/1816-0301-2025-22-2-7-32

Оригинальная статья
Original Article

Моделирование рисков развития «умного города» как сложной социотехнической системы

С. В. Кругликов¹, А. Г. Давыдовский^{2✉}

¹Объединенный институт проблем информатики
Национальной академии наук Беларусь,
ул. Сурганова, 6, Минск, 220012, Беларусь

²Институт информационных технологий Белорусского
государственного университета информатики и радиоэлектроники,
ул. Козлова, 28, Минск, 220037, Беларусь

✉E-mail: agd2011@list.ru

Аннотация

Цели. Экспоненциальный рост урбанизированных территорий и численности городского населения потребовал разработки новых моделей и методов описания, оптимального управления рисками современных городов как больших и сложных социальных, технологических и логистических систем. Работа посвящена проблеме анализа рисков развития «умного города» как сложной социотехнической системы с позиций гипотезы семи поколений «умных городов» с использованием элементов математического моделирования.

Методы. Использованы анализ литературных источников, моделирование семи поколений «умного города» как сложной социотехнической системы на основе линейных дифференциальных уравнений, сценарный анализ развития концепции «умного города».

Результаты. Представлена общая характеристика концепции «умного города» и рассмотрены риски дезорганизации сервисов и процессов жизнедеятельности «умного города». Обоснована гипотеза семи поколений «умного города». Разработаны системы линейных дифференциальных уравнений, характеризующие каждое из поколений «умного города». Предложен подход к моделированию и оценке интегрального риска «умного города».

Заключение. Разработаны системы линейных дифференциальных уравнений, характеризующие каждое из семи поколений «умного города», а также соответствующие им риски. Предложена обобщенная модель интегральной оценки рисков «умного города» как сложной социотехнической системы на основе нелинейного дифференциального уравнения. Сформулированы основы методологии управления рисками дезорганизации сервисов и процессов жизнедеятельности для различных поколений «умного города» как эволюционирующей социотехнической системы.

Ключевые слова: математическое моделирование, оценка рисков, социотехническая система, сценарный подход, «умный город»

Для цитирования. Кругликов, С. В. Моделирование рисков развития «умного города» как сложной социотехнической системы / С. В. Кругликов, А. Г. Давыдовский // Информатика. – 2025. – Т. 22, № 2. – С. 7–32. – DOI: 10.37661/1816-0301-2025-22-2-7-32.

Конфликт интересов. Авторы заявляют об отсутствии конфликта интересов.

Поступила в редакцию | Received 08.05.2025

Подписана в печать | Accepted 05.06.2025

Опубликована | Published 30.06.2025

Modelling the risks of "smart city" development as a complex sociotechnical system

Sergey V. Kruglikov¹, Anatoly G. Davydovskii^{2✉}

¹*The United Institute of Informatics Problems of the National Academy of Sciences of Belarus, st. Surganova, 6, Minsk, 220012, Belarus*

²*Institute of Information Technologies of the Belarusian State University of Informatics and Radioelectronics, st. Kozlova, 28, Minsk, 220037, Belarus*

✉E-mail: agd2011@list.ru

Abstract

Objectives. The exponential growth of urbanized territories and urban populations required the development of new models and methods for describing and optimally managing the risks of modern cities as large and complex social, technological, and logistical systems. The paper is devoted to the problem of analyzing the risks of the development of a "smart city" as a complex sociotechnical system from the standpoint of the hypothesis of seven generations of "smart cities" using elements of mathematical modeling.

Methods. The analysis of literary sources, modeling the seven generations of the "smart city" as a complex sociotechnical system based on linear differential equations and a numerical analysis of the development of the "smart city" concept are used.

Results. The general characteristics of the "smart city" concept are presented and the risks of disruption of services and processes of the "smart city" are considered. The hypothesis of seven generations of the "smart city" is substantiated. Systems of linear differential equations have been developed that characterize each of the generations of the "smart city". An approach to modeling and assessing the integrated risk of a "smart city" is proposed.

Conclusion. Systems of linear differential equations have been developed that characterize each of the seven generations of the "smart city", as well as the risks associated with them. A generalized model of integrated risk assessment of a "smart city" as a complex sociotechnical system based on a nonlinear differential equation is proposed. The fundamentals of the methodology for managing the risks of disorganization of services and life processes for different generations of the "smart city" as an evolving sociotechnical system are formulated.

Keywords: mathematical modeling, risk assessment, sociotechnical system, scenario approach, "smart city"

For citation. Kruglikov S. V., Davydovskii A. G. *Modelling the risks of "smart city" development as a complex sociotechnical system*. Informatika [Informatics], 2025, vol. 22, no. 2, pp. 7–32 (In Russ.). DOI: 10.37661/1816-0301-2025-22-2-7-32.

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

Введение. В первой четверти XXI в. городское население мира составило более 55 %. Особенno активно его численность возрастала в последние 70 лет. Если в 1950 г. в городах проживало около 751 млн человек, то в 2018 г. в мире насчитывалось уже 4,2 млрд горожан. По данным ООН, к 2050 г. доля жителей городов возрастет до 68 %. Наиболее урбанизированными регионами мира являются Северная Америка (82 % городского населения), Латинская Америка и страны Карибского бассейна (81 % городского населения), Европа (74 % городского населения) и Океания (68 % городского населения). В Российской Федерации 74,4 % граждан проживают в городах. При этом именно в городах производится порядка 87 % мирового внутреннего валового продукта [1]. В Республике Беларусь доля городского населения составляет 78,85 % [2]. Бурный рост городов и городского населения является не только результатом повышения эффективности мировой экономики, но и причиной возникновения множества проблем, обусловленных интенсивным развитием урбанизированной среды [3, 4]:

- низкого уровня управления и слабых институтов;
- неполноценной инфраструктуры;

- роста неравенства (1 млрд горожан в мире живут за чертой бедности, в 75 % городов мира ситуация с бедностью ухудшилась);
- дефицита жилья (в городах требуется более 1 млрд новых домов и квартир, свыше 880 млн человек живут в трущобах и за чертой бедности);
- высокого уровня высокотехнологической преступности;
- загрязнения окружающей среды (города, занимая до 3 % общей поверхности суши, потребляют 75 % природных ископаемых и обеспечивают 75 % токсичных выбросов в окружающую среду). Существенными экологическими рисками «умного города» могут быть широкое использование новых синтетических и композиционных материалов, прогрессивно возрастающее электромагнитное загрязнение окружающей среды (электромагнитный смог) вследствие распространения интернета вещей (Internet of Things, IoT), мобильных источников доступа в интернет, персонального и общественного электротранспорта, а также необходимость утилизации и (или) регенерации использованных электрических аккумуляторных элементов;
- новых рисков, включая высокотехнологический терроризм, быстрое распространение угроз кибербезопасности, возникновение и распространение новых штаммов возбудителей инфекционных заболеваний;
- распространения медиавирусов и инфодемий;
- избыточной секьюритизации информационных, технологических и экономических активов;
- рисков для здоровья населения «умного города», обусловленных электромагнитным загрязнением окружающей среды, снижением физической активности, нарушением естественных биологических ритмов жителей «умного города», информационным стрессом, токсическими нагрузками.

Экспоненциальный рост урбанизированных территорий и численности городского населения потребовал разработки новых моделей и методов описания, оптимального управления устойчивостью современных городов как больших и сложных социальных, технологических и логистических систем [2, 3]. В условиях четвертой технологической революции происходит сближение концепции виртуального города, которая позволяет реализовывать городские функции в киберпространстве, обеспечивая доступность городских инфраструктур, и предоставляет широкие каршеринговые услуги [4]. Подобный синтез технологий в условиях «умного города» неизбежно будет сопровождаться формированием новых рисков по мере развития технологий «умного города», иерархической сложности организации и эволюции его форм на основе социотехнического подхода. Вне последнего невозможно проектирование больших саморазвивающихся систем, к которым относятся и «умные города» [5, 6].

В этой связи целью работы является обоснование гипотезы семи поколений «умного города», а также моделирование рисков развития каждого из семи поколений и интегрального риска «умного города» как сложной социотехнической системы (СТС).

Риски «умного города» и их источники. Для постнеклассического типа проектирования и исследования больших и сложных систем с субъект-полисубъектной или метасубъектной парадигмой управления характерны широкое использование компьютерных технологий, эволюционизм и социотехнический подход [6].

Средовая парадигма саморазвивающихся больших и сложных систем является ведущей в постнеклассической науке. Все современные определения «умного города» предполагают интеграцию информационных и коммуникационных технологий, интернета вещей для управления городским имуществом, включая местные информационные системы, больницы, школы, библиотеки, транспорт, электростанции, системы водоснабжения и управления отходами, правоохранительные органы и другие общественные службы. Несмотря на наличие большого количества определений «умного города» (так, только в докладе Международного союза электросвязи (Женева, 2016) приведено 116 определений), «концептуальная неопределенность умных городов не тормозит их развития»¹ [7]. «Умный город» может быть охарактеризован как инфраструктура интегрированных сетей и технологий, обеспечивающая [8, 9]:

¹Trends in Smart City Development // National League of Cities. – URL: <http://www.nlc.org/sites/default/TrendsinSmartCity.pdf> (date of access: 31.07.2024).

«умную среду» – продвижение технологий устойчивого развития, использование возобновляемых источников энергии;

«умную администрацию» – новые модели прогнозирования развития, принятия решений в режиме реального времени, более качественного, быстрого и прозрачного оказания высокотехнологических услуг сервисов «умного города»;

«умную экономику» – поддержку предпринимательства и конкурентной среды, выработку новых форм экономического развития, например в виде экономики мобильных приложений, «шеренговой» экономики и т. п.;

«умное население» – эффективное информирование, поддержку социальных инициатив, инклюзивность и расширение возможности активного участия граждан с особенностями психофизического развития в управлении городом;

«умный образ жизни» – улучшение качества жизни, повышение безопасности и снижение рисков;

«умную мобильность» – саморегулируемые транспортные системы, беспилотные транспортные системы.

Вместе с тем недостаточно разработаны модели структурно-функциональной организации и управления «умным городом».

Анализ рисков «умного города» как СТС позволяет выделить социальные [10], технологические, технические, киберфизические, экономические, управленческие, институциональные риски [11], а также риски, связанные с деградацией человеческого капитала, и экологические риски.

Социальные риски обусловлены угрозами распространения вирусных, бактериальных и паразитарных инфекций, дезорганизацией здравоохранения, образования, безопасности, туризма, культуры [12]. Сюда можно отнести условия дезорганизации взаимодействия между гражданами, бизнесом и администрацией, а также распространение деструктивных субкультур. Важный вклад в социальные риски «умного города» вносит мотивация отдельных субъектов и групп населения к девиантному и делинквентному поведению, недостаточная готовность жителей к использованию современных и перспективных технологий. Очевидно, другой крайностью может быть преобладание идей технократического, транс- и постгуманистического сознания среди населения «умного города».

Технологические риски вызваны взаимодействием сервисов «умного города», использующих большие данные и интернет вещей, однако кибератака может исказить или уничтожить данные, что приведет к сбоям в системах управления и поддержки принятия решений, техногенным авариям. Далеко не всегда сбор и обработка всех доступных данных с датчиков является оптимальной стратегией. При этом большую угрозу кибербезопасности представляет утечка данных из информационной инфраструктуры «умного города» [12].

Технические риски – утечка данных, уязвимость для киберпреступников, зависимость жесткой инфраструктуры от информационной инфраструктуры. Такие риски могут существенно уменьшить эффективность участия граждан в управлении «умным городом», что снижает качество деятельности «умной администрации/правительства». Наиболее частой проблемой, возникающей в «умном городе», является обеспечение кибербезопасности и конфиденциальности данных, хотя круг проблем гораздо шире [13, 14].

Киберфизические риски обусловлены взаимосвязями и функционированием зданий, офисов, предприятий, транспортных магистралей, регионов в «умном городе». «Умный город» генерирует многочисленные потоки больших данных, но повсеместное внедрение цифровых технологий сопряжено с рисками утечек и потерь данных, включая персональные. При этом наиболее уязвимыми являются персональные мобильные смарт-устройства с интернетом вещей. Сетевые компьютерные эпидемии могут вызывать системные сбои в работе сервисов «умного города», вплоть до полной их дезорганизации и дезинтеграции. Учитывая информационно насыщенную коммуникационную среду «умного города», его население будет неизбежно подвергаться влиянию недостоверных медиаресурсов и медиавирусных сообщений, способных спровоцировать инфодемии [10, 15].

Экономические риски включают снижение эффективности внедряемых инноваций и усиление цифрового неравенства [16].

Управленческие риски включают техническую и психологическую неготовность властей, что связано с нежеланием администрации делегировать полномочия социальным сообществам «умного города», а также потерей городской аутентичности [17].

Институциональные риски включают отставание законодательства и правовой системы от технологий и сервисов, которые разрабатываются и инсталлируются в «умной» урбанизированной среде, а также наличие конфликта интересов между общественными и частными коммерческими кругами, между «умной администрацией» и «умным населением», отдельными потребителями и «умным бизнесом». Это требует контроля получения прибыли, в частности, от услуг шеринга.

Риски, связанные с деградацией человеческого капитала, включают неготовность населения к использованию технологий и негативным последствиям внедрения и освоения инноваций. При этом необходимо помнить об устойчивом, комфортном для жизни, экологичном, декарбонизированном развитии «умного города». Цифровизация и внедрение современных технологий не должны быть самоцелью. Здесь возникает риск «умного города» как технологической ловушки для его разработчиков, а затем и населения.

Экологические риски обусловлены высокими темпами распространения и наращивания мощностей информационных технологий и скорости процессов обработки данных. Цифровой сектор является потребителем значительных объемов электроэнергии. Так, в Дании центры обработки данных к 2030 г. будут потреблять более 15 % всей производимой электроэнергии, а сектор банковских услуг потребляет электроэнергии больше, чем сектор золотодобывающей промышленности. Развитие блокчейн-технологий вначале привело к росту потребления электроэнергии для проведения расчетов, а затем к излишку видеокарт, который обострил проблему электронных отходов. В 2021 г. мировой объем электронных отходов (e-waste) составил около 57 млн метрических тонн, из них только 2 % потоков твердых отходов, представляли собой около 70 % опасных отходов, попадающих в окружающую среду [14].

Биотоксикологические риски обусловлены биодеструкцией синтетических и композиционных материалов, используемых в бытовых, производственных и транспортных конструкциях и устройствах, а также в имплантируемых приборах и изделиях биомедицинского назначения.

Биосоциальные риски связаны с вероятностью возникновения и распространения инфекций, вызванных вирусными, бактериальными, грибковыми и паразитарными возбудителями, которые могут легко распространяться по канализационным и вентиляционным системам «умного города» в отсутствие фильтрационных и дезинфицирующих средств и устройств.

Информационно-психологические риски возникают при вирусном медиаконтенте, быстро распространяющимся по медиаканалам «умного города». В «умном городе» в условиях равномерного и практически мгновенного распространения информационных потоков чрезвычайно высок риск инфодемии. Кроме того, угрозы дезинтеграции подсистем «умного города» могут быть обусловлены как целенаправленными хакерскими и социоинженерными атаками, интенсивной продукцией потоков слабоструктурированных больших данных, превышающей возможности их обработки, так и случайными отказами оборудования или сбоями программного обеспечения, а также человеческим фактором [18].

Психофизиологические риски обусловлены нарушениями здоровья, вызванными гиподинамией; информационным стрессом, вызванным интенсивным воздействием медиаисточников, и дезорганизацией биологических ритмов вследствие возможности круглогодичной активности.

К инклюзивным рискам относятся события, связанные с осложнениями участия жителей из числа лиц с особенностями психофизического развития в социально-экономической, культурной, медийно-сетевой и социально-политической жизни сообщества «умного города».

Цифровые технологии обеспечивают организацию и поддержку взаимодействия в таких системах отношений, как администрация – администрации, администрация – гражданам, администрация – бизнесу, бизнес – гражданам, бизнес – бизнесу, бизнес – администрации, граждане – гражданам, граждане – бизнесу, граждане – администрации. Такие взаимоотношения являются источниками рисков «умного города» и служат основанием для разработки

перспективных классификаций (таксономий) рисков. В этой связи предложен перечень рисков «умного города», включающий:

- возникновение аварий беспилотного и шерингового транспорта;
- возникновение и распространение деструктивных цифровых субкультур, сетевых культов и сект;
- возникновение инфодемий;
- возникновение технологических инцидентов;
- высокую вероятность повторной идентификации предположительно анонимизированных или псевдоанонимизированных данных, повторного использования персональных данных и больших данных;
- деградацию культурных, политических и социальных особенностей «умного города»;
- комплексные ментально-психологические атаки на население «умного города» с последующим развитием инфодемий и массовых панических атак;
- массовое распространение недостоверной информации (фейков) и панических состояний;
- непрозрачное использование персональных данных и больших данных о функционировании сервисов «умного города» как большой и сложной СТС;
- несанкционированное использование технологий блокчейна [19], дополненной и виртуальной реальности, искусственного интеллекта, беспилотных транспортных систем;
- несанкционированное использование персональных данных жителей «умного города»;
- ошибочность выводов, основанных на случайных корреляциях между вариационными рядами больших данных, которые характеризуют деятельность сервисов «умного города»;
- повторное использование больших данных, генерируемых сервисами «умного города» (угрозу «двойной бухгалтерии» больших данных);
- подверженность психологическим манипуляциям по медиальным каналам;
- подмену социальных, экономических и социально-политических аспектов реализации проекта «умного города» исключительно техническими вопросами;
- полную зависимость от технологий, порождающую криминальные риски;
- развитие «цифрового слабоумия»;
- распространение бактериальных и вирусных эпидемий;
- распространение эпидемии компьютерных и сетевых вирусов;
- реализацию проекта «умного города» в интересах отдельных элитарных групп населения;
- сбои функционирования логистических инфраструктур;
- сетевые кибератаки;
- снижение уровня коммуникативной культуры населения;
- снижение уровня образованности населения;
- социоинженерные атаки на граждан, учреждения, предприятия и информационные активы «умного города»;
- тотальный контроль над жителями и отдельными социальными группами населения «умного города»;
- феномены цифрового неравенства и цифрового разрыва между различными социальными группами граждан.

Гипотеза семи поколений «умных городов». Концепция «умного города» была разработана в рамках постнеклассической науки и на основе метасубъектной парадигмы управления. В процессе реализации она стала одним из маркеров перехода к постнеклассическому этапу проектирования и исследования больших и сложных СТС [4]. Их основу составляет интерсубъектная парадигма управления – «умный город» как большая и сложная информационная СТС формируется уже не вокруг субъектов, а вокруг множества объектов, связанных между собой по радиоканалам в рамках интернета вещей, интернета людей и интернета ботов. Это свидетельствует о переходе от парадигмы «умный город для людей» и «умный город для ботов» с преобладанием интернета людей, интернета вещей, интернета процессов и «слабого» искусственного интеллекта в виде голосовых помощников и консультантов наподобие chatGPT к парадигме «умный город для не людей» с господством интернета вещей, интернета процессов и интернета

ботов. Таким образом, происходит трансформация «умного города» из социально-ориентированной в социально-деструктивную большую систему. В большой и сложной СТС «умного города» активно используются непрерывно достраиваемые информационно-технологические платформы, обеспечивающие реализацию всех сервисов, направленных на повышение эффективности управления ресурсами, улучшение качества жизни граждан и решение задач устойчивого социально-экономического и эколого-социотехнического развития. При этом основными условиями реализации «умного города» как развивающейся СТС выступают наличие развитой инфраструктуры (умных сетей, транспорта, энергоснабжения), управления большими данными (IoT-устройствами, комплексами сенсоров и аналитики), различных цифровых платформ, обеспечивающих участие и связь граждан с бизнесом, городской администрацией, транспортом, энергетикой, здравоохранением, образованием, а также обеспечение экологической безопасности и устойчивости «умного города» через интегрированные и централизованные платформы.

Авторами предложены основы гипотезы семи поколений «умного города», рассматривающей эволюцию «умного города» как последовательную смену технологических поколений «умного города» от 1.0 до 7.0, причем последнее поколение 7.0 будет финальным, после чего сама концепция «умных городов» утратит свою актуальность. При этом «умный город» поколения 1.0 – технологический (technology-driven) комплекс, который сфокусирован на внедрении отдельных технологий для решения конкретных проблем (умного освещения, видеонаблюдения). «Умный город» поколения 2.0 ориентирован на агентства (agency-driven) и осуществляет координацию между отдельными агентствами и ведомствами для улучшения городских сервисов. «Умный город» поколения 3.0 ориентирован на граждан (citizen-driven) и их активное вовлечение в процессы принятия решений и использование данных для улучшения городской жизни. «Умный город» поколения 4.0 ориентирован на бизнес (business-driven) и развитие инноваций за счет использования городских данных и технологий. «Умный город» поколения 5.0 является устойчивым (sustainable) комплексом, направленным на обеспечение устойчивого сбалансированного развития экологических, социальных и экономических процессов для создания комфортной среды для жизни. «Умный город» поколения 6.0 является инклюзивным (inclusive) и обеспечивает равный доступ к городским сервисам и возможностям для всех жителей независимо от их возраста, пола, социального статуса и физических особенностей. «Умный город» поколения 7.0 – сложный саморазвивающийся и самообучающийся (self-learning) урбоэкосоциотехнический комплекс, который постоянно адаптируется и улучшается на основе данных и обратной связи от жителей и включенных систем.

Математические модели «умного города» поколений 1.0 – 7.0. Актуальной проблемой является математическое моделирование рисков развития различных поколений «умного города», для которого необходимо построение математических моделей, характеризующих каждое поколение «умного города» в отдельности.

Модель «умного города» поколения 1.0 включает следующие параметры:

уровень развития «умного города» ($U(t)$) – безразмерная величина, или процент (0–100 %), или балл. Отражает степень интеграции технологий и решений «умного города» в различные сферы, транспорт, энергетику, управление отходами (размерность – баллы, %);

уровень экономического роста ($G(t)$) – валовый городской продукт или его производная (размерность – денежные единицы/год, например, долл. США/год);

уровень инноваций ($I(t)$) – количество патентов, инвестиции в исследования и разработки, доля инновационных компаний (размерность – патенты/год, денежные единицы/год, %);

уровень социальной напряженности ($S(t)$) – количество протестов, уровень преступности, индекс социального неравенства, например коэффициент Джини (размерность – количество/год, %);

уровень устойчивости ($E(t)$) – уровень выбросов CO_2 , потребление ресурсов, индекс экологической устойчивости (размерность – тонны CO_2 /год, объем ресурсов/год, балл);

уровень цифровизации ($D(t)$) – проникновение интернета, использование мобильных устройств, объем данных, генерируемых в городе. Размерность – %, количество устройств/человек, объем данных (Тбайт);

уровень вовлеченности граждан ($C(t)$) – участие в принятии решений, активность в социальных сетях, использование городских сервисов (размерность – количество пользователей, %);

фактор турбулентности ($Turb(t)$) – суммарное влияние социально-экономических и социокультурных потрясений (например, пандемий, экономических кризисов, военных конфликтов). Представляет собой внешнее воздействие, изменяющее параметры системы (размерность – баллы и условные единицы, так как турбулентность обуславливает влияние на коэффициенты уравнений).

Закономерен вопрос о размерности итоговой модели, рассматривающей все переменные. Минимальная размерность модели (Dimension of the Model, DM_{min}) определяется количеством переменных, которые используются для описания состояния системы «умного города». Чем больше переменных, тем выше размерность модели и тем детальнее она описывает систему, но при этом возрастает сложность анализа и интерпретации результатов. В случае модели «умного города» поколения 1.0 (G1.0) ее минимальная размерность 1.0 $DM_{min} = 8$.

Каждая из ключевых переменных может быть описана линейным дифференциальным уравнением:

а) уровень развития «умного города» ($U(t)$):

$$\frac{dU(t)}{dt} = \alpha I(t) + \beta D(t) - \gamma S(t)T(t) - \delta E(t)Turb(t) + \varepsilon C(t), \quad (1)$$

где α – коэффициент, отражающий влияние инноваций на развитие «умного города»;

β – коэффициент, отражающий влияние цифровизации на развитие «умного города»;

γ – коэффициент, отражающий негативное влияние социальной напряженности на развитие «умного города», усиленное турбулентностью;

δ – коэффициент, отражающий негативное влияние экологических проблем на развитие «умного города», усиленное турбулентностью;

ε – коэффициент, отражающий влияние вовлеченности граждан;

б) уровень экономического роста ($G(t)$):

$$\frac{dG(t)}{dt} = \varphi U(t) + \psi I(t) - \omega S(t) - \zeta E(t), \quad (2)$$

где φ – коэффициент, отражающий влияние развития «умного города» на экономический рост;

ψ – коэффициент, отражающий влияние инноваций на экономический рост;

ω – коэффициент, отражающий влияние социальной напряженности на экономический рост;

ζ – коэффициент, отражающий влияние экологических проблем;

в) уровень инноваций ($I(t)$):

$$\frac{dI(t)}{dt} = \eta G(t) + \theta D(t) - \kappa S(t)Turb(t), \quad (3)$$

где η – коэффициент, отражающий влияние экономического роста на уровень инноваций;

θ – коэффициент, отражающий влияние цифровизации на уровень инноваций;

κ – коэффициент, отражающий негативное влияние социальной напряженности и турбулентности на уровень инноваций;

г) уровень социальной напряженности ($S(t)$):

$$\frac{dS(t)}{dt} = \lambda(1 - D(t))Turb(t) - \mu C(t) + \nu E(t)Turb(t), \quad (4)$$

где λ – коэффициент, отражающий влияние турбулентности и низкой цифровизации на социальную напряженность;

μ – коэффициент, отражающий влияние вовлеченности граждан на социальную напряженность;

v – коэффициент, отражающий влияние ухудшения экологической обстановки на социальную напряженность;

д) уровень устойчивости ($E(t)$):

$$\frac{dE(t)}{dt} = \xi G(t) - oU(t)Turb(t) - \pi Pol(t) + \sigma D(t), \quad (5)$$

где ξ – коэффициент, отражающий влияние экономического роста на устойчивость (учитывая, что экономический рост может вести к увеличению загрязнения);

o – коэффициент, отражающий негативное влияние развития «умного города» в условиях турбулентности (например, из-за неэффективного использования ресурсов);

π – коэффициент, отражающий влияние на уровень загрязнения в целом;

σ – коэффициент, отражающий влияние цифровизации на устойчивость (например, за счет повышения энергоэффективности);

е) уровень цифровизации ($D(t)$):

$$\frac{dD(t)}{dt} = \rho G(t) + \tau I(t) - vS(t)Turb(t), \quad (6)$$

где ρ – коэффициент, отражающий влияние экономического роста на цифровизацию;

τ – коэффициент, отражающий влияние инноваций на цифровизацию;

v – коэффициент, отражающий влияние социальной напряженности и турбулентности на цифровизацию;

ж) уровень вовлеченности граждан ($C(t)$):

$$\frac{dC(t)}{dt} = \chi D(t) + \psi_1 Pol(t) - \omega_1 S(t)Turb(t), \quad (7)$$

где χ – коэффициент, отражающий влияние цифровизации;

ψ_1 – коэффициент, отражающий влияние государственной политики;

ω_1 – коэффициент, отражающий влияние социальной напряженности и турбулентности;

$Pol(t)$ – комплексная характеристика внутренней социально-политической среды «умного города», которая может быть представлена как функция многих переменных;

з) фактор турбулентности $Turb(t)$:

$$\frac{dTurb(t)}{dt} = \alpha_1 S(t) - \beta_1 U(t) + \gamma_1 Pol(t), \quad (8)$$

где α_1 – коэффициент, отражающий влияние социальной напряженности на турбулентность;

β_1 – коэффициент, отражающий влияние уровня развития «умного города» на устойчивость к турбулентности;

γ_1 – коэффициент, отражающий влияние государственной политики.

При этом интегральный риск нарушений структурно-функциональной организации и процессов жизнедеятельности «умного города» поколения 1.0 можно охарактеризовать уравнением

$$Risk_{(1.0)} = C_{U(t)} \int U(t) dt + C_{G(t)} \int G(t) dt + C_{I(t)} \int I(t) dt + C_{S(t)} \int S(t) dt + C_{E(t)} \int E(t) dt + C_{D(t)} \int D(t) dt + C_{C(t)} \int C(t) dt + C_{Turb(t)} \int Turb(t) dt, \quad (9)$$

где $C_{U(t)}$, $C_{G(t)}$, $C_{I(t)}$, $C_{S(t)}$, $C_{E(t)}$, $C_{D(t)}$, $C_{C(t)}$, $C_{Turb(t)}$ – показатели потенциального ущерба соответственно от нарушения развития «умного города», стагнации экономического роста, дезорганизации инновационной деятельности и устойчивого развития, социальной напряженности, процессов цифровизации, недостаточной вовлеченности граждан, влияния фактора турбулентности.

Моделирование «умного города» поколения 2.0 рассматривает не только технологическую инфраструктуру, но и активное участие граждан в управлении городом, а также адаптацию к изменяющимся социально-экономическим и социокультурным условиям. Турбулентность связана с неопределенностью, кризисами, быстрыми изменениями в технологиях, социальных нормах и ценностях. Модель должна учитывать следующие ключевые аспекты «умного города» поколения 2.0:

- инфраструктуру – транспорт, энергетику, водоснабжение, связь;
- экономику – инновации, занятость, предпринимательство, туризм;
- социальную сферу – образование, здравоохранение, безопасность, культуру, социальную сплоченность;
- управление – участие граждан, открытые данные, электронное правительство, реагирование на кризисы;
- технологии – IoT, искусственный интеллект (artificial intelligence, AI), большие данные, облачные вычисления, блокчейн;
- окружающую среду – качество воздуха и воды, управление отходами, энергоэффективность, адаптацию к климатическим изменениям;
- человеческий капитал – уровень образования, компетенции, здоровье, креативность, социально-экономическую активность населения.

Минимальная размерность модели «умного города» поколения 2.0 $DM_{min} = 15$ ($8 + 7 = 15$).

Инфраструктуру «умного города» характеризуют следующие переменные:

$I_t(t)$ – индекс эффективности транспортной системы (среднее время в пути, загруженность дорог, доступность общественного транспорта);

$I_e(t)$ – индекс энергоэффективности (потребление энергии на душу населения, доля возобновляемых источников энергии);

$I_w(t)$ – индекс качества водоснабжения (доступность чистой воды, потери воды в сетях);

$I_c(t)$ – индекс развития телекоммуникационной инфраструктуры (покрытие интернетом, скорость передачи данных).

Экономику:

$E_i(t)$ – индекс инновационной активности (количество стартапов, патентов, инвестиции в исследования и разработки);

$E_e(t)$ – уровень занятости населения;

$E_t(t)$ – доходы от туризма.

Социальную сферу:

$S_e(t)$ – уровень образования населения;

$S_h(t)$ – индекс здоровья населения (средняя продолжительность жизни, уровень заболеваемости);

$S_s(t)$ – индекс безопасности (уровень преступности, количество дорожно-транспортных происшествий);

$S_c(t)$ – индекс социальной сплоченности (уровень доверия между людьми, участие в общественных мероприятиях).

Управление:

$G_p(t)$ – уровень участия граждан в управлении городом (использование онлайн-платформ для голосования, участие в общественных обсуждениях);

$G_o(t)$ – уровень открытости данных (доступность городских данных в открытом формате);

$G_e(t)$ – эффективность электронного правительства (удобство получения государственных услуг онлайн);

$G_r(t)$ – эффективность реагирования на кризисы (скорость и адекватность реагирования на стихийные бедствия, эпидемии).

Технологии:

$T_{IoT}(t)$ – уровень внедрения IoT-технологий (количество подключенных устройств, сбор данных);

$T_{AI}(t)$ – уровень использования AI в управлении городом (анализ данных, автоматизация процессов).

Окружающую среду:

$Env_{air}(t)$ – индекс качества воздуха;

$Env_{waste}(t)$ – эффективность системы управления отходами (доля переработанных отходов).

Человеческий капитал:

$HC_{skills}(t)$ – уровень развития навыков, необходимых для работы в условиях общества поколения 5.0 (например, цифровых навыков, критического мышления, креативности);

$HC_{adapt}(t)$ – способность населения адаптироваться к изменениям.

Ключевые переменные, в частности индекс эффективности транспортной системы, уровень образования населения, индекс инновационной активности, эффективность реагирования на кризисы, можно описать следующими линейными дифференциальными уравнениями:

индекс эффективности транспортной системы ($I_t(t)$):

$$\frac{dI_t(t)}{dt} = \alpha (T_{IoT}(t)G_p(t) - \beta(I_t(t) - I_t^*) + \gamma G_p(t) Turb(t)), \quad (10)$$

где α – влияние IoT на участие граждан в транспортных коммуникациях;

β – тенденция к возвращению к некоторому равновесному состоянию (I_t^* – целевое значение);

γ – влияние эффективного реагирования на кризисы (например, пробки) и турбулентности ($Turb(t)$);

уровень образования населения ($S_e(t)$):

$$\frac{dS_e(t)}{dt} = \delta(G_p(t)T_{AI}(t)) - \lambda(S_e(t) - S_e^*) + \mu HC_{skills}(t), \quad (11)$$

где δ – положительное влияние участия граждан и AI на образование (например, онлайн-курсы, персонализированное обучение),

λ – тенденция к возвращению к равновесному состоянию ((S_e^*) – целевой уровень образования),

μ – влияние человеческого капитала;

индекс инновационной активности ($E_i(t)$):

$$\frac{dE_i(t)}{dt} = v(G_p(t)S_e(t)) + \xi T_{eAI}(t) - \pi Turb(t), \quad (12)$$

где v – положительное влияние участия граждан и образования на инновации,

ξ – влияние AI,

π – негативное влияние социально-экономической и социокультурной турбулентности;

эффективность реагирования на кризисы ($G_r(t)$):

$$\frac{dG_r(t)}{dt} = \rho(T_{IoT}(t)T_{AI}(t)) - \sigma(G_r(t) - G_r^*), \quad (13)$$

где ρ – положительное влияние IoT и AI на реагирование на кризисы,

σ – тенденция к возвращению к равновесному состоянию ((G_r^*) – целевой уровень).

При этом интегральный риск нарушений структурно-функциональной организации и процессов жизнедеятельности «умного города» поколения 2.0 можно представить аддитивной моделью

$$Risk_{(2.0)} = Risk_{(1.0)} + C_{I(t)} \int dI_t(t) dt + C_{S(t)} \int dS_e(t) dt + C_{E(t)} \int E_i(t) dt + C_{G_r(t)} \int dG_r(t) dt, \quad (14)$$

где $C_{I(t)}$, $C_{S(t)}$, $C_{E(t)}$, $C_{G_r(t)}$ – показатели потенциального ущерба от нарушений соответственно транспортной, образовательной, инновационной и антикризисной деятельности.

Для **моделирования эволюции «умного города» поколения 3.0** можно выделить следующие ключевые переменные:

экономику «умного города» ($Ec(t)$) – ВВП города, инвестиции в инновации, количество высокотехнологичных предприятий и т. д.;

социальную сферу ($Soc(t)$) – уровень образования, здравоохранения, социальной защиты, индекс счастья населения, уровень преступности и т. д.;

инфраструктуру ($Inf(t)$) – уровень развития транспортной сети, энергетической сети, сети связи, жилищного фонда, коммунального хозяйства и т. д.;

экологию ($Env(t)$) – уровень загрязнения воздуха, воды, уровень шумового загрязнения, объем отходов, площадь зеленых насаждений и т. д.;

технологии ($Tech(t)$) – уровень развития информационных технологий, искусственного интеллекта, интернета вещей, робототехники, блокчейна и т. д.;

управление ($Gov(t)$) – эффективность городского управления, уровень коррупции, участие граждан в управлении, прозрачность принятия решений и т. д.;

безопасность ($Sec(t)$) – уровень кибербезопасности, общественной безопасности, устойчивость к чрезвычайным ситуациям;

устойчивость ($Rec(t)$) – способность города адаптироваться к изменениям и восстанавливаться после шоков (экономических, социальных, экологических);

культуру и самобытность ($Cul(t)$) – поддержка местных традиций, развитие искусства, сохранение исторического наследия, создание новых культурных пространств.

Каждую из этих переменных можно детализировать, увеличив размерность модели. Минимальная размерность модели «умного города» поколения 3.0 $Dm_{min} = 24$ ($15 + 9 = 24$).

Необходимо сформулировать систему дифференциальных уравнений, описывающих динамику изменения каждой переменной во времени. Эти уравнения должны отражать взаимосвязи между переменными и влияние внешних факторов, включая социально-экономическую и социокультурную турбулентность.

Упрощая и агрегируя многие процессы, предположим, что социально-экономическое хозяйство «умного города» можно представить четырехкомпонентной моделью, включающей инфраструктуры ($Inf(t)$), технологии ($Tech(t)$), городское управление ($Gov(t)$) и факторы социально-экономической, социотехнической и социокультурной турбулентности ($Turb(t)$), и описать линейными дифференциальными уравнениями.

В таком случае экономика «умного города» поколения 3.0 в общем виде может быть описана линейным дифференциальным уравнением:

$$\frac{dEc(t)}{dt} = \alpha Inf(t) + \beta Tech(t) + \gamma Gov(t) - \delta Turb(t), \quad (15)$$

где $\alpha, \beta, \gamma, \delta$ – параметры, отражающие влияние инфраструктуры, технологий, управления и турбулентности на экономику города;

– социальный компонент $Soc(t)$:

$$\frac{d Soc(t)}{dt} = \varepsilon Ec(t) + \zeta Gov(t) + \eta Cul(t) - \theta Turb(t), \quad (16)$$

где $\varepsilon, \zeta, \eta, \theta$ – параметры, отражающие соответственно характер влияния экономики, управления, культуры, социально-экономической и социотехнической турбулентности на социальную сферу умного города;

– инфраструктурный компонент $Inf(t)$:

$$\frac{d Inf(t)}{dt} = \lambda Ec(t) + \mu Gov(t) - \nu Env(t), \quad (17)$$

где λ, μ, v – параметры, отражающие влияние экономики, управления и экологии на инфраструктуру;

– экологический компонент $Env(t)$:

$$\frac{d Env(t)}{dt} = \rho Ec(t) + \sigma Tech(t) + \tau Man(t) Gov(t) - \varphi Turb(t), \quad (18)$$

где $\rho, \sigma, \tau, \varphi$ – параметры, отражающие влияние экономики, технологий, управления и турбулентности на окружающую среду (экологию) умного города,

$Man(t)$ – комплекс внешних информационных и иных манипулятивных влияний на управление «умным городом», который может быть представлена как функция многих переменных;

– технологический компонент $Tech(t)$:

$$\frac{dTech(t)}{dt} = \chi Ec(t) + \psi Gov(t) + \omega Soc(t) - \varsigma Turb(t), \quad (19)$$

где $\chi, \psi, \omega, \varsigma$ – параметры, отражающие влияние экономики, управления, социальной сферы и турбулентности на технологии;

– управляемый компонент $Gov(t)$:

$$\frac{dGov(t)}{dt} = ATech(t) + BSoc(t) - CTurb(t), \quad (20)$$

где A, B, C – параметры, отражающие влияние технологий, социальной сферы и турбулентности на управление;

– компонент безопасности $Scrt(t)$:

$$\frac{dScrt(t)}{dt} = DTech(t) + EGov(t) - FTurb(t), \quad (21)$$

где D, E, F – параметры, отражающие влияние технологий, управления и турбулентности на безопасность;

– компонент устойчивости ($Res(t)$):

$$\frac{dRes(t)}{dt} = GEc(t) + HSoc(t) + IEnv(t) + JSec(t) - KTurb(t), \quad (22)$$

где G, H, I, J, K – параметры, отражающие влияние экономики, социальной сферы, экологии, безопасности и турбулентности на устойчивость;

– культура и самобытность ($Cul(t)$):

$$\frac{dCul(t)}{dt} = LGov(t) + MSoc(t) - NTech(t) + OTurb(t), \quad (23)$$

где L, M, N, O – параметры, отражающие влияние управления, социальной сферы, технологий и турбулентности на культуру и социокультурную сферу.

При этом интегральный риск нарушений структурно-функциональной организации и процессов жизнедеятельности «умного города» поколения 3.0 можно охарактеризовать аддитивной моделью

$$\begin{aligned} Risk_{(3.0)} = & Risk_{(2.0)} + C_{Ec(t)} Ec(t) dt + C_{Soc(t)} \int Soc(t) dt + C_{Inf(t)} \int Inf(t) dt + \\ & + C_{Env(t)} \int Env(t) dt + C_{Tech(t)} \int Tech(t) dt + C_{Gov(t)} \int Gov(t) dt + \\ & + C_{Src(t)} \int Src(t) dt + C_{Res(t)} \int Res(t) dt + C_{Cul(t)} \int Cul(t) dt, \end{aligned} \quad (24)$$

где $C_{Ec(t)}$, $C_{Soc(t)}$, $C_{Inf(t)}$, $C_{Env(t)}$, $C_{Tech(t)}$, $C_{Gov(t)}$, $C_{Scrt(t)}$, $C_{Res(t)}$, $C_{Cul(t)}$ – показатели потенциального ущерба от нарушений соответственно развития «умного города», стагнации экономического роста, дезорганизации инновационной деятельности, устойчивого развития, процессов цифровизации, недостаточной вовлеченности граждан, влияния фактора турбулентности.

Моделирование эволюции «умного города» поколения 4.0 требует более комплексного подхода, чем в случае поколения 3.0. «Умный город» поколения 4.0 характеризуется как широким спектром применения цифровых технологий, так и глубокой интеграцией физического и киберпространства, ориентированностью на человека, устойчивостью развития социально-экономических процессов и высокой степенью адаптивности информационно-технологической платформы «умного города». Размерность модели «умного города» поколения 4.0 определяется количеством переменных, ключевыми среди которых являются (кроме описанных для модели «умного города» поколения 3.0 – экономика, социальная сфера, инфраструктура, экология, технологии, управление, безопасность, устойчивость, культура):

вовлеченность граждан ($Cit(t)$) – уровень участия граждан в принятии решений, использовании цифровых сервисов, предоставлении обратной связи;

персонализация сервисов ($Per(t)$) – степень адаптации городских сервисов к индивидуальным потребностям и предпочтениям граждан;

цифровая идентичность и доверие ($Id(t)$) – уровень безопасности и конфиденциальности цифровой информации граждан, уровень доверия к цифровым технологиям и городским властям;

устойчивость к киберугрозам ($Cyb(t)$) – защищенность городской инфраструктуры и сервисов от кибератак и сбоев;

инклюзивность и доступность ($Inc(t)$) – обеспечение равного доступа к городским ресурсам и сервисам для всех групп населения, включая людей с ограниченными возможностями, пожилых людей и представителей различных культур;

креативность и инновации ($Cre(t)$) – создание благоприятной среды для развития креативных индустрий, поддержки инновационных проектов и привлечения талантов;

энергетическая эффективность и декарбонизация ($En(t)$) – сокращение потребления энергии, использование возобновляемых источников энергии, снижение выбросов парниковых газов.

Минимальная размерность модели «умного города» поколения 4.0 $DM_{min} = 31$ ($24 + 7 = 31$).

Система дифференциальных уравнений для «умного города» поколения 4.0 будет более сложной, чем для «умного города» поколения 3.0, из-за большего количества переменных и более сложных взаимосвязей между ними, включая:

– вовлеченность граждан ($Cit(t)$):

$$\frac{dcit(t)}{dt} = aGov(t) + bTech(t) + cSoc(t) - dTurb(t), \quad (25)$$

где a, b, c, d – параметры, отражающие влияние управления, технологий, социальной сферы и турбулентности на вовлеченность граждан;

– персонализацию сервисов ($Per(t)$):

$$\frac{dPer(t)}{dt} = eTech(t) + fCit(t) - gID(t), \quad (26)$$

где e, f, g – параметры, отражающие влияние технологий, вовлеченности граждан и цифровой идентичности на персонализацию сервисов;

– цифровую идентичность и доверие ($ID(t)$):

$$\frac{dId(t)}{dt} = hGov(t) + iCyb(t) + jEth(t) - kTurb(t), \quad (27)$$

где h, i, j, k – параметры, отражающие влияние управления, кибербезопасности, этики и турбулентности на цифровую идентичность и доверие,

$Eth(t)$ – уровень этики использования технологий, который также может быть описан дифференциальным уравнением;

- устойчивость к киберугрозам ($Cyb(t)$):

$$\frac{dCyb(t)}{dt} = lTech(t) + m Gov(t) - nAtk(t), \quad (28)$$

где l, m, n – параметры, отражающие влияние технологий, управления и уровня киберугроз ($Atk(t)$);

- инклузивность и доступность ($Inc(t)$):

$$\frac{dInc(t)}{dt} = oGov(t) + pSoc(t) + qTech(t) - rTurb(t), \quad (29)$$

где o, p, q, r – параметры, отражающие влияние управления, социальной сферы, технологий и турбулентности на инклузивность и доступность;

- креативность и инновации ($Cre(t)$):

$$\frac{dCre(t)}{dt} = sTech(t) + tSoc(t) + uGov(t) - vTurb(t), \quad (30)$$

где s, t, u, v – параметры, отражающие влияние технологий, социальной сферы, управления и турбулентности на креативность и инновации;

- энергетическую эффективность и декарбонизацию ($En(t)$):

$$\frac{dEn(t)}{dt} = wTech(t) + xGov(t) - yEc(t), \quad (31)$$

где w, x, y – параметры, отражающие влияние технологий, управления и экономики на энергетическую эффективность и декарбонизацию.

При этом интегральный риск нарушений структурно-функциональной организации и процессов жизнедеятельности модели «умного города» поколения 4.0 можно описать аддитивной моделью

$$Risk_{(4.0)} = Risk_{(3.0)} + C_{Cit(t)} \int Cit(t) dt + C_{Per(t)} \int Per(t) dt + C_{Id(t)} \int Id(t) dt + \\ + C_{Cyb(t)} + C_{Cyb(t)} \int Cyb(t) dt + C_{Inc(t)} \int Inc(t) dt + C_{Cre(t)} \int Cre(t) dt + C_{En(t)} \int En(t) dt, \quad (32)$$

где $C_{Cit(t)}, C_{Per(t)}, C_{Id(t)}, C_{Cyb(t)}, C_{Inc(t)}, C_{Cre(t)}, C_{En(t)}$ – показатели потенциального ущерба от снижения соответственно вовлеченности граждан, персонализации сервисов, цифровой идентичности и доверия, устойчивости к киберугрозам, инклузивности и доступности, креативности и инновационного потенциала, энергетической эффективности и декарбонизации.

Моделирование эволюции «умного города» поколения 5.0 представляет собой вершину в развитии концепции «умных городов», интегрируя в себе все лучшее от предыдущих поколений и фокусируясь на создании устойчивой, человеко-ориентированной и самоорганизующейся системы. Минимальный набор переменных, помимо тех, что были указаны для «умного города» поколения 4.0 (экономика, социальная сфера, инфраструктура, экология, технологии, управление, безопасность, устойчивость, культура и самобытность, вовлеченность граждан, персонализация сервисов, цифровая идентичность и доверие, устойчивость к киберугрозам, инклузивность и доступность, креативность и инновации, энергетическая эффективность и декарбонизация), должен включать:

когнитивные способности системы ($Cog(t)$) – способность города к самообучению, анализу данных, прогнозированию, принятию решений в реальном времени;

самоорганизацию ($Org(t)$) – степень автономности различных городских подсистем и их способность к координации без централизованного управления;

адаптивность ($Adp(t)$) – способность города адаптироваться к изменяющимся условиям, новым технологиям и вызовам;

устойчивость к сложным системам ($Sys(t)$) – показатель, характеризующий способность города поддерживать стабильность и функциональность в условиях сложных взаимосвязанных систем;

этические рамки AI ($Eth(t)$) – четко определенные этические принципы и механизмы контроля за использованием искусственного интеллекта в городе;

биоцентричность ($Bio(t)$) – оценку соответствия городской среды потребностям биологических систем, включая человека, животных и растений;

циркулярную экономику ($Cir(t)$) – уровень внедрения принципов циркулярной экономики в различные сферы городской жизни.

Минимальная размерность модели «умного города» поколения 5.0 $DM_{min} = 38$ ($31 + 7 = 38$).

Система дифференциальных уравнений (33)–(39) будет еще более сложной и должна отражать взаимосвязи между новыми переменными и остальными показателями. При этом важно учитывать эффекты обратной связи и нелинейные зависимости.

В частности, когнитивные способности населения «умного города» поколения 5.0 как СТС ($Cog(t)$) могут быть описаны уравнением

$$\frac{dCog(t)}{dt} = aTech(t) + bOrg(t) + cEth(t) - dTurb(t), \quad (33)$$

где a, b, c, d – параметры, отражающие влияние технологий, самоорганизации, этических рамок и турбулентности на когнитивные способности системы;

– самоорганизация ($Org(t)$):

$$\frac{dOrg(t)}{dt} = eCog(t) + fInf(t) - g Gov(t), \quad (34)$$

где e, f, g – параметры, отражающие влияние когнитивных способностей, инфраструктуры и управления на самоорганизацию;

– адаптивность ($Adp(t)$):

$$\frac{dAdp(t)}{dt} = k Cog(t) + q Sys(t) - p Turb(t), \quad (35)$$

где k, q, p – параметры, отражающие влияние когнитивных способностей, устойчивости к сложным системам и турбулентности на адаптивность;

– устойчивость к сложным системам ($Sys(t)$):

$$\frac{dSys(t)}{dt} = l Adp(t) + m Tech(t) + n Inc(t), \quad (36)$$

где l, m, n – параметры, отражающие влияние адаптивности, технологий и инклузивности на устойчивость к сложным системам;

– этические рамки AI ($Eth(t)$):

$$\frac{dEth(t)}{dt} = o Gov(t) + p Soc(t) + r Cog(t) - s Turb(t), \quad (37)$$

где l, o, r, s – параметры, отражающие влияние управления, социальной сферы, когнитивных способностей и турбулентности на этические рамки AI;

– биоцентричность ($Bio(t)$):

$$\frac{dBio(t)}{dt} = tEnv(t) + u Inf(t) + wCit(t) + zEth(t), \quad (38)$$

где t, u, w, z – параметры, отражающие влияние экологии, инфраструктуры, вовлеченности граждан и этических рамок АІ на биоцентричность;

– циркулярная экономика ($Cir(t)$):

$$\frac{dCir(t)}{dt} = a_c Env(t) + e_c Tech(t) + i_c Gov(t) + j_c Turb(t), \quad (39)$$

где a_c, e_c, i_c, j_c – параметры, отражающие влияние экологии, технологий, управления и турбулентности на циркулярную экономику.

Моделирование эволюции «умного города» поколения 5.0 в условиях социально-экономической и социокультурной турбулентности при переходе к обществу 5.0 представляет собой сложную, но необходимую задачу. Создание такой модели требует интеграции передовых знаний и технологий из различных областей науки и техники, а также широкого вовлечения заинтересованных сторон. Успешное моделирование позволит городским властям принимать более обоснованные решения, строить устойчивые и процветающие города будущего и создавать комфортную и безопасную среду для жизни граждан. При этом интегральный риск нарушений структурно-функциональной организации и процессов жизнедеятельности модели «умного города» поколения 5.0 можно охарактеризовать аддитивной моделью

$$Risk_{(5.0)} = Risk_{(4.0)} + C_{Cog(t)} \int Cog(t) dt + C_{Org(t)} \int Org(t) dt + C_{Adp(t)} \int Adp(t) dt + \\ + C_{Sys(t)} \int Sys(t) dt + C_{Eth_AI(t)} \int Eth_AI(t) dt + C_{Bio(t)} \int Bio(t) dt + C_{Cir(t)} \int Cir(t) dt, \quad (40)$$

где $C_{Cog(t)}, C_{Org(t)}, C_{Adp(t)}, C_{Sys(t)}, C_{Eth_AI(t)}, C_{Bio(t)}, C_{Cir(t)}$ – потенциальный ущерб соответственно от снижения когнитивного потенциала населения, самоорганизации и адаптивности, устойчивости к сложным системам, этических рамок использования технологий АІ, биоцентричности и качественного состояния циркулярной экономики (экономики замкнутого цикла).

Моделирование эволюции «умного города» поколения 6.0 – это концептуально продвинутый этап, представляющий собой самовосстанавливающуюся, саморазвивающуюся и сознательную экосистему, ориентированную на максимальное благополучие каждого жителя. В условиях социально-экономической и социокультурной турбулентности при переходе к обществу 5.0 такая модель должна обладать не только адаптивностью, но и способностью к предвидению и активному формированию будущего.

Сложность модели «умного города» поколения 6.0 возрастает экспоненциально по сравнению с предыдущими поколениями. Это связано с необходимостью учета не только физических, цифровых и когнитивных аспектов, но и факторов, определяющих самосознание, этическую зрелость и способность к коллективному разуму.

К перечню переменных, определенных для «умного города» поколения 5.0, добавляются следующие:

коллективный разум (collective mind, $CM(t)$) – способность городской системы к коллективному решению проблем, объединению интеллекта и опыта жителей, организаций и АІ для достижения общих целей;

самосознание ($Slf(t)$) – степень осознания городом своих сильных и слабых сторон, ресурсов, ограничений и целей. Включает в себя способность к самооценке и самокритике;

предвидение ($Pre(t)$) – способность системы к долгосрочному прогнозированию, выявлению будущих трендов и рисков, разработке адаптивных стратегий;

этическая зрелость ($Eth(t)$) – уровень развития этических норм и ценностей в системе управления городом и в сознании жителей. Включает принципы справедливости, равенства, устойчивости и ответственности;

креативное самовыражение ($Cre(t)$) – возможности для жителей реализовывать свой творческий потенциал, создавать новые продукты, сервисы и культурные ценности;

психическое здоровье ($Men(t)$) – уровень психического благополучия жителей города, снижение уровня стресса, тревожности и депрессии;

духовное развитие ($Spi(t)$) – поддержка религиозных, философских и духовных потребностей жителей, создание условий для личностного роста и самосовершенствования;

управление смыслами ($Mng(t)$) – способность городской системы создавать и поддерживать общие ценности, смыслы и цели, которые объединяют жителей и направляют развитие города;

устойчивость к экзистенциальным рискам ($Exi(t)$) – готовность города к противодействию глобальным угрозам, таким как изменение климата, пандемии, ядерная война и другие катастрофы.

Минимальная размерность модели «умного города» поколения 5.0 $DM_{min} = 47$ ($38 + 9 = 47$). Между тем реальная размерность может быть на порядок выше, если детализировать каждую из этих переменных.

Система дифференциальных уравнений для «умного города» поколения 6.0 становится чрезвычайно сложной и требует применения продвинутых методов моделирования, таких как мультиагентное моделирование, системная динамика с когнитивными агентами, нейросетевое моделирование и эволюционные алгоритмы.

Дифференциальные уравнения (41)–(48) модели «умного города» поколения 6.0 включают следующие компоненты:

- коллективный разум $CM(t)$:

$$\frac{dCM(t)}{dt} = \alpha Cog(t) + \beta Soc(t) + \gamma Cit(t) + \delta Eth_m(t) - \varepsilon Turb(t), \quad (41)$$

где $\alpha, \beta, \gamma, \delta, \varepsilon$ – параметры, отражающие влияние когнитивных способностей, социальной сферы, вовлеченности граждан, этической зрелости и турбулентности на коллективный разум;

- самосознание ($Slf(t)$):

$$\frac{dSlf(t)}{dt} = \zeta CM(t) + \eta Pre(t) + \theta Gov(t), \quad (42)$$

где ζ, η, θ – параметры, отражающие влияние коллективного разума, предвидения и управления на самосознание;

- этическую зрелость ($dEth_m(t)$):

$$\frac{dEth_m(t)}{dt} = mGov(t) + nSoc(t) + oCM(t) + pSpi(t) - qTurb(t), \quad (43)$$

где m, n, o, p, q – параметры, отражающие влияние управления, социальной сферы, коллективного разума, духовного развития и турбулентности на этическую зрелость;

- креативное самовыражение ($Cre(t)$):

$$\frac{dCre(t)}{dt} = rSoc(t) + sTech(t) + tCul(t) + uMen(t), \quad (44)$$

где r, s, t, u – параметры, отражающие влияние социальной сферы, технологий, культуры и психического здоровья на креативное самовыражение;

- психическое здоровье ($Men(t)$):

$$\frac{dMen(t)}{dt} = vSoc(t) + wCre(t) + xSpi(t) - yTurb(t), \quad (45)$$

где v, w, x, y – параметры, отражающие влияние социальной сферы, креативного самовыражения, духовного развития и турбулентности на психическое здоровье;

- духовное развитие ($Spi(t)$):

$$\frac{dSpi(t)}{dt} = zSoc(t) + aaEth_m(t) + bbMen(t), \quad (46)$$

где z, aa, bb – параметры, отражающие влияние социальной сферы, этической зрелости и психического здоровья на духовное развитие;

- управление смыслами ($Mng(t)$):

$$\frac{dMng(t)}{dt} = ccGov(t) + ddEth_m(t) + eeCM(t) + ffCre(t), \quad (47)$$

где cc, dd, ee, ff – параметры, отражающие влияние управления, этической зрелости, коллективного разума и креативного самовыражения на управление смыслами;

- устойчивость к рискам социально-экономической турбулентности ($Exit(t)$):

$$\frac{dExit(t)}{dt} = ggPre(t) + hhTech(t) + iiInf(t) + jjCM(t) - kkTurb(t), \quad (48)$$

где gg, hh, ii, jj, kk – параметры, отражающие влияние предвидения, технологий, инфраструктуры, коллективного разума и турбулентности на устойчивость к экзистенциальным рискам.

Моделирование эволюции «умного города» поколения 6.0 потребует значительных усилий со стороны мирового научного сообщества и тесного сотрудничества между учеными, инженерами, политиками и гражданами. «Умный город» поколения 6.0 должен стать не только эффективным и удобным местом для жизни, но и платформой для развития человеческого потенциала, создания новых знаний и решения глобальных проблем.

При этом интегральный риск нарушений структурно-функциональной организации и процессов жизнедеятельности «умного города» поколения 6.0 можно описать аддитивной моделью

$$Risk_{(6.0)} = Risk_{(5.0)} + C_{CM(t)} \int CM(t) dt + C_{Slf(t)} \int Slf(t) dt + C_{Eth(t)} \int Eth(t) dt + \\ + C_{Cre(t)} \int Cre(t) dt + C_{Men(t)} \int Men(t) dt + C_{Spi(t)} \int Spi(t) dt + \\ + C_{Mng(t)} \int Mng(t) dt + C_{Exit(t)} \int Exit(t) dt, \quad (49)$$

где $C_{CM(t)}, C_{Slf(t)}, C_{Eth(t)}, C_{Cre(t)}, C_{Men(t)}, C_{Spi(t)}, C_{Mng(t)}, C_{Exit(t)}$ – показатели потенциального ущерба от нарушения соответственно функционирования коллективного разума «умного города», самосознания жителей, предвидения, этической зрелости, креативного самовыражения, психического здоровья, духовного развития, управления смыслами, устойчивости к рискам устойчивого существования «умного города» поколения 6.0 как СТС.

Моделирование эволюции «умного города» поколения 7.0 представляет собой теоретическую вершину, где город достигает симбиотических отношений с окружающей средой, его жителями и даже с другими городами, создавая глобальную сеть устойчивых и саморегулируемых систем. Размерность модели «умного города» поколения 7.0 является наиболее высокой из всех поколений, так как включает в себя все предыдущие факторы, а также новые измерения, связанные с межвидовым взаимодействием, глобальной связностью, технологиями будущего и метафизическими аспектами.

К переменным, определенным для «умного города» поколения 6.0, добавляются следующие: межвидовые взаимоотношения ($Int(t)$) – степень взаимодействия с другими живыми существами (животными, растениями, экосистемами), включая биоразнообразие и экологическое равновесие;

глобальная связность ($Gbl(t)$) – степень интеграции с глобальными сетями знаний, ресурсов, инноваций и управления;

синхронизация с циклами природы ($Nat(t)$) – соответствие ритмам природы (временам года, солнечной активности, биоритмам) и использование природных ресурсов максимально эффективно;

трансцендентность ($Tra(t)$) – способность города к изменению своих фундаментальных принципов и целей, включая выход за рамки привычных ограничений и создание новых способов существования;

квантовая коммуникация ($Qc(t)$) – возможность использования квантовых технологий для коммуникации, обработки данных и управления;

управление сознанием ($Con(t)$) – способность системы понимать и управлять сознанием людей (в этических рамках) для повышения благополучия и решения проблем;

экологическое сознание ($Eco(t)$) – уровень осознания жителями города своей роли в сохранении окружающей среды и бережного отношения к природным ресурсам;

управление временем ($Tim(t)$) – способность города оптимизировать использование времени, учитывая биоритмы, личные предпочтения и приоритеты;

космические связи ($Cos(t)$) – готовность и способность города к взаимодействию с внеземными цивилизациями.

Минимальная размерность модели «умного города» поколения 5.0 $DM_{min} = 56$ ($47 + 9 = 56$).

Система дифференциальных уравнений (50)–(58) модели «умного города» поколения 7.0 включает следующие компоненты:

- межвидовые взаимоотношения ($Int(t)$):

$$\frac{dInt(t)}{dt} = aEnv(t) + bBio(t) + cEco(t) - dTurb(t), \quad (50)$$

где a, b, c, d – параметры, отражающие влияние экологии, биоцентрического, экологического сознания и турбулентности на межвидовые взаимоотношения;

- глобальную связность ($Gbl(t)$):

$$\frac{dGbl(t)}{dt} = eCol(t) + fTech(t) + gCos(t) - hTurb(t), \quad (51)$$

где e, f, g, h – параметры, отражающие влияние коллективного разума, технологий, космических связей и турбулентности на глобальную связность;

- синхронизацию с циклами природы ($Nat(t)$):

$$\frac{dNat(t)}{dt} = iEnv(t) + jTim(t) + kEco(t), \quad (52)$$

где i, j, k – параметры, отражающие влияние экологии, управления временем и экологического сознания на синхронизацию с циклами природы;

- трансцендентность ($Tra(t)$):

$$\frac{dTra(t)}{dt} = l Slf(t) + m Col(t) + n Spi(t), \quad (53)$$

где l, m, n – параметры, отражающие влияние самосознания, коллективного разума и духовного развития на трансцендентность;

- квантовую коммуникацию ($Oc(t)$):

$$\frac{dOc(t)}{dt} = o Tech(t) + p Gbl(t), \quad (54)$$

где o, p – параметры, отражающие влияние технологий и глобальной связности на квантовую коммуникацию;

- управление сознанием ($Con(t)$):

$$\frac{dCon(t)}{dt} = qMen(t) + rEth(t), \quad (55)$$

где q, r – параметры, отражающие влияние психического здоровья и этической зрелости на управление сознанием (в этических рамках);

- экологическое сознание ($Eco(t)$):

$$\frac{dEco(t)}{dt} = sEnv(t) + tSoc(t) + uInt(t), \quad (56)$$

где s, t, u – параметры, отражающие влияние экологии, социальной сферы и межвидовых взаимоотношений на экологическое сознание;

- управление временем ($Tim(t)$):

$$\frac{dTIm(t)}{dt} = vCog(t) + wMng(t) + xPre(t), \quad (57)$$

где v, w, x – параметры, отражающие влияние когнитивных способностей, управления смыслами и предвидения на управление временем;

- космические связи ($Cos(t)$):

$$\frac{dCos(t)}{dt} = yGbl(t) + zTech(t), \quad (58)$$

где y, z – параметры, отражающие влияние глобальной связности и технологий связи с удаленными космическими объектами.

При этом интегральный риск нарушений структурно-функциональной организации и процессов жизнедеятельности «умного города» поколения 7.0 можно представить аддитивной моделью

$$Risk_{(7.0)} = Risk_{(6.0)} + C_{Int(t)} \int Int(t) dt + C_{Gbl(t)} \int Gbl(t) dt + C_{Nat(t)} \int Nat(t) dt + \\ + C_{Tra(t)} \int Tra(t) dt + C_{Jc(t)} \int Jc(t) dt + C_{Con(t)} \int Con(t) dt + C_{Eco(t)} \int Eco(t) dt + \\ + C_{Tim(t)} \int Tim(t) dt + C_{Cos(t)} \int Cos(t) dt, \quad (59)$$

где $C_{Int(t)}, C_{Gbl(t)}, C_{Nat(t)}, C_{Tra(t)}, C_{Jc(t)}, C_{Con(t)}, C_{Eco(t)}, C_{Tim(t)}, C_{Cos(t)}$ – потенциальный ущерб от нарушений соответственно межвидовых взаимоотношений ($Int(t)$), глобальной связности ($Gbl(t)$), синхронизации с природными и биологическими циклами ($Nat(t)$), технологической трансцендентности ($Tra(t)$), квантовой коммуникации ($Jc(t)$), управления индивидуальным и коллективным сознанием граждан ($Con(t)$), экологическими системами ($Eco(t)$), управления временем ($Tim(t)$), управления космическими связями с «умными сообществами», находящимися на околоземных орbitах, за их пределами и на других небесных телах ($Tra(t)$).

Система дифференциальных уравнений для «умного города» поколения 7.0 будет невероятно сложной, требующей использования самых передовых методов математического моделирования, включая мультиагентное моделирование с квантовыми вычислениями и системную динамику с когнитивными агентами, наделенными самосознанием и способностью к трансцендентному мышлению на основе механизмов самообучения и саморазвития; нейросетевое моделирование, учитывающее сознательные процессы, эволюционные алгоритмы для поиска оптимальных решений в динамичных условиях, а также метафизическое моделирование (в рамках этических ограничений) – учет влияния духовных факторов и высших целей на развитие города.

Успешное создание такой модели может не только привести к созданию городов будущего, но и оказать влияние на развитие всего человечества и гармоничного сосуществования с окружающей средой и космосом.

При анализе моделей различных поколений «умного города» целесообразен сценарный подход. Для обсуждения могут быть предложены четыре потенциально возможных сценария:

– устойчивого роста, требующий высокие инвестиции в «зеленую экономику», ресурсосберегающие технологии, эффективную социальную политику, низкую интенсивность факторов социальной напряженности;

– цифровой трансформации, предполагающий быстрое развитие цифровых технологий, но, возможно, рост социального неравенства;

– кризиса, допускающий высокую вероятность турбулентности социально-экономических процессов и дезорганизации социотехнических систем с последующим замедлением экономического развития и ростом социальной напряженности;

– регресса, возможный при условии недостаточных инвестиций, неэффективной политики, быстрого роста социальной напряженности, снижения уровня развития человеческого потенциала населения.

Кроме того, предложены пять сценариев эволюционного развития «умного города» – ноосферный, сбалансированный, устойчивого развития, технократический и постгуманистический. Очевидно, что на практике реальный сценарий развития «умного города» будет включать признаки каждого из вышеперечисленных сценариев.

Моделирование интегрального риска «умного города». Важную роль в моделировании рисков развития «умного города» как СТС играют анализ и моделирование турбулентности социально-экономической среды ($Turb(t)$) «умного города». В целом турбулентность сложно формализовать, но можно ввести функцию $T(t)$, которая отражает уровень неопределенности и резких изменений. Это может быть случайный процесс (например, белый шум) или функция, зависящая от внешних факторов (например, экономических кризисов, социальных волнений, технологических прорывов). Турбулентность «умного города» как развивающейся СТС можно моделировать различными способами в зависимости от конкретных сценариев с помощью стохастической, периодической и импульсной моделей. Авторы предлагают комбинированную модель турбулентности социально-экономической среды «умного города»:

$$Turb(t) = \sum_{i=1}^n A_i \delta(t - t_i) + \sum_{i=1}^n A_i \sin(\omega_i t + \varphi_i) + \varepsilon(t), \quad (60)$$

где A_i – амплитуда импульса событий;

t_i – момент времени, когда происходит импульс событий;

$\delta(t)$ – дельта-функция Дирака;

$\omega_i(t)$ – частота колебаний i -й социально-экономической характеристики;

$\varphi_i(t)$ – фаза колебаний i -й социально-экономической характеристики;

$\varepsilon(t)$ – случайный шум.

Необходимо отметить, что каждое поколение «умного города» характеризуется собственными рисками, обусловленными как используемыми технологиями, уровнем развития и организации социально-экономических, технологических и социокультурных процессов, так и турбулентностью социально-экономической среды.

Здесь уместен как качественный, так и количественный анализ рисков «умного города». С качественной точки зрения эволюционное развитие концепции «умного города» как СТС неизбежно будет сталкиваться с вызовами и рисками киберугроз, связанных с уязвимостью информационно-технологических инфраструктур и платформ, технологической зависимостью, этическим дисбалансом и социальным расслоением.

С количественной точки зрения риски эволюционного развития «умного города» как СТС можно представить в виде модели семи поколений, которая предполагает, что в «умном городе», к какому бы поколению он ни относился, будут потенциально присутствовать риски всех остальных шести поколений. Хотя, безусловно, определяющими будут риски каждого поколе-

ния «умного города» в отдельности ($Risk_{Gi}(t)$). Тогда обобщенную модель интегрального риска «умного города» как СТС, находящейся на уровне поколения G_i , можно представить нелинейным дифференциальным уравнением:

$$\frac{d Risk_{Gi}(t)}{dt} = K_i \left(Risk_{Gi}(t) + \sum_{j \in [1,7]}^n \omega_j (Risk_{Gj}(t)) + Turb_{Gi}(t) \right) Risk_{Gi}(t), \quad (61)$$

где $Risk_{Gi}(t)$ – риски «умного города» поколения G_i ;

$Risk_{Gj}(t)$ – риски «умного города» поколения G_j (причем $i \neq j$);

$Turb_{Gi}(t)$ – турбулентность социально-экономической среды «умного города» поколения G_i ;

K_i – постоянная, характеризующая прирост рисков в «умном городе» поколения G_i .

В общем виде интегральный риск «умного города» ($Risk_{SC}$) любого поколения обусловлен присутствием технологий и факторов риска, относящихся к его другим поколениям:

$$Risk_{SC} = \omega_1 Risk_{(1.0)} + \omega_2 Risk_{(2.0)} + \omega_3 Risk_{(3.0)} + \omega_4 Risk_{(4.0)} + \omega_5 Risk_{(5.0)} + \\ + \omega_6 Risk_{(6.0)} + \omega_7 Risk_{(7.0)}, \quad (62)$$

где ω_{1-7} – удельно-весовые коэффициенты рисков, присущих каждому поколению «умного города».

Численное моделирование рисков поколений «умного города» как СТС может быть осуществлено *in silico* в программной среде MATLAB, Simulink, а также с помощью языка высокого уровня Python с использованием библиотек SciPy, NumPy, PySD. Кроме того, моделирование причинно-следственных логических связей между факторами, событиями и процессами, приводящими к формированию рисков в СТС «умный город», может быть успешно осуществлено с помощью Vensim – специализированного программного обеспечения для моделирования системной динамики сложных СТС.

Заключение. Многомерная интеграция людей, производственных сообществ и технологий в рамках «умного города» как сложной СТС, которая формирует единую интеллектуальную, технологическую и социокультурную среду, будет неизбежно порождать принципиально новые социотехнические, технологические и биосоциальные риски. Такие риски, обусловленные внедрением конвергентных технологий в условиях четвертой промышленной революции, могут иметь многомерные и сложные биосоциальные и социогуманитарные последствия, анализ и оценка которых выходят за пределы настоящей работы и затрагивают проблемы глубокой трансформации человека в перспективных урбанизированных сообществах, таких как «умный город», с учетом смены его поколений.

Феномен «умного города» является продуктом определенной социально-экономической формации и результатом определенного достигнутого уровня технологических возможностей, социальных отношений и социокультурного развития. «Умный город» представляет собой большую и сложную социотехническую систему, включающую разнообразные технологические сервисы, интегрированные друг с другом и обслуживающие различные категории населения. В свою очередь, «умный город» становится инструментом технологических и социотехнических преобразования сформировавшей его социально-экономической формации.

Таким образом, «умный город» можно рассматривать как социотехническую технологию преобразования социально-экономической формации, реализовавшей его как проект на определенном уровне технологического, экономического и социального развития. При этом переход от одного поколения «умного города» к другому может оказывать эффект драйвера, преобразующего реальную социально-экономическую формацию. При этом важнейшее значение приобретает управление рисками «умного города».

На основе гипотезы семи поколений «умного города», рассматривающей эволюцию «умного города» как последовательную смену его технологических поколений от 1.0 до 7.0, предложены математические модели каждого из семи поколений на основе линейных дифференциаль-

ных уравнений. Представлены математические модели рисков нарушения структурно-функциональной организации и процессов жизнедеятельности «умного города» для поколений 1.0 – 7.0 (формулы (9), (14), (24), (32), (40), (49), (59)).

Вместе с тем предложенные модели рисков нарушения структурно-функциональной организации и процессов жизнедеятельности «умного города» в различных поколениях могут являться одной из основ для методологии эволюционного проектирования экосистем «умных сообществ и регионов» не только как киберфизических, но и как больших и сложных социотехнических систем, обладающих потенциалом драйверов развития человеческой цивилизации.

Вклад авторов. А. Г. Давыдовский предложил провести моделирование рисков развития «умного города» на основе сценарного подхода. С. В. Кругликов сформулировал цель и задачи исследования. Оба автора принимали участие в подготовке текста статьи, анализе и интерпретации результатов исследования.

Список использованных источников

1. Одинцов, А. В. Основные риски реализации концепции «умного города» / А. В. Одинцов // Социодинамика. – 2019. – № 10. – С. 1–8. – DOI: 10.25136/2409-7144.2019.10.30636.
2. Беларусь в цифрах. Статистический справочник 2025. – Минск : НСК РБ, 2025. – 62 с.
3. Костина, Е. А. Риски умного города / Е. А. Костина // Мир экономики и управления. – 2023. – Т. 23, № 2. – С. 108–125. – DOI: 10.25205/2542-0429-2023-23-2-108-125.
4. Шваб, К. Технологии четвертой промышленной революции / К. Шваб, Н. Дэвис ; пер. с англ. – М., 2018. – 317 с.
5. Умная инфраструктура, физические и информационные активы, Smart Cities, BIM, GIS и IoT / В. П. Куприяновский, В. В. Алењков, И. А. Соколов [и др.] // International Journal of Open Information Technologies. – 2017. – Vol. 5, № 10. – С. 55–86.
6. The theoretical foundations of sociotechnical systems change for sustainability: A systematic literature review / P. Savaget, M. Geissdoerfer, A. Kharrazi, S. Evans // Journal of Cleaner Production. – 2019. – Vol. 206. – P. 878–892. – DOI: 10.1016/j.jclepro.2018.09.208.
7. Саак, А. Э. Модели информационно-технологической структуры умного города / А. Э. Саак, В. Н. Тюшняков, Е. В. Пахомов // Фундаментальные исследования. – 2017. – № 10-2. – С. 387–391.
8. Kitchin, R. Data-driven, networked urbanism / R. Kitchin // Steps. – 2017. – Vol. 3, no. 2. – P. 98–116.
9. Smart home energy management systems in internet of things networks for green cities demands and services / M. S. Aliero, K. N. Qureshi, M. F. Pasha, G. Jeon // Environmental Technology & Innovation. – 2021. – Vol. 22. – URL: <https://www.sci-hub.ru/10.1016/j.eti.2021.101443> (date of access: 17.06.2024).
10. Frank, E. Smart City = Smart Citizen = Smart Economy? An Economic Perspective of Smart Cities / E. Frank, G. Aznar Fernández-Montesinos // Social, Legal, and Ethical Implications of IoT, Cloud, and Edge Computing Technologies, IGI Global, Hershey / ed.: G. Cornetta, A. Touhafi, G.M. Muntean. – 2020. – Chap. 7. – P. 161–180.
11. Smart cities: advances in research – an information systems perspective / E. Ismagilova, L. Hughes, Y. K. Dwivedi, K. R. Raman // International Journal of Information Management. – 2019. – Vol. 47. – P. 88–100.
12. Суховская, Д. Н. Анализ структуры концепции «умного города» / Д. Н. Суховская, Н. А. Шульгин // E-Scio. – 2019. – № 9(36). – С. 737–746.
13. Камолов, С. Г. Технологии будущего для «умных городов» / С. Г. Камолов, А. М. Корнеева // Вестник Московского государственного областного университета. Серия: Экономика. – 2018. – № 2. – С. 100–114.
14. Yang, R. Smart city development Models: A cross-cultural regional analysis from theory to practice / R. Yang, F. Zhen // Research in Globalization. – 2024. – Vol. 8. – P. 100221 – DOI: 10.1016/j.resglo.2024.100221.
15. Bibri, S. E. A novel model for data-driven smart sustainable cities of the future: A strategic roadmap to transformational change in the era of big data / S. E. Bibri, J. Krogstie // Future Cities and Environment. – 2021. – Vol. 7, no. 1. – P. 1–25.
16. Костина, Е. А. Барьеры использования сервисов обратной связи на муниципальном уровне / Е. А. Костина, А. В. Костин // Журнал Сибирского федерального университета. Гуманитарные науки. – 2023. – Т. 16, № 3. – С. 468–480.

-
17. Пироцкая, А. В. Концепция «умного города» в представлении экспертов и жителей: проблемы и противоречия в реализации концепции / А. В. Пироцкая // Мир экономики и управления. – 2020. – Т. 20, № 2. – С. 178–196.
18. Kitchin, R. The opportunities, challenges and risks of big data for official statistics / R. Kitchin // Statistical Journal of the International Association of Official Statistics. – 2015. – Vol. 31, no. 3. – P. 471–481.
19. A survey of blockchain technology applied to smart cities: Research issues and challenges / X. J. Junfeng, H. Tang, T. Huang [et al.] // IEEE Communications Surveys & Tutorials. – 2019. – Vol. 21, iss. 3. – P. 2794–2830.
-

References

1. Odintsov A. V. *The main risks of implementing the concept of a "smart city"*. Sociodinamika [Sociodynamics], 2019, no. 10, pp. 1–8 (In Russ.). DOI: 10.25136/2409-7144.2019.10.30636.
2. Belarus' v cifrah. Statisticheskij spravochnik 2025. *Belarus in Numbers. Statistical Handbook 2025*. Minsk, Nacional'nyj statisticheskij komitet Respubliki Belarus', 2025, 62 p. (In Russ.).
3. Kostina E. A. *The risks of a smart city*. Mir jekonomiki i upravlenija [The World of Economics and Management], 2023, vol. 23, no. 2, pp. 108–125 (In Russ.). DOI: 10.25205/2542-0429-2023-23-2-108-125.
4. Schwab K., Davis N. *Shaping the Fourth Industrial Revolution*. World Economic Forum, 2018, 287 p.
5. Kupriyanovsky V. P., Alenkov V. V., Sokolov I. A., Zazhigalkin A. V., Klimov A. A., ..., Namot D. E. Smart infrastructure, physical and information assets, Smart Cities, BIM, GIS and IoT. *International Journal of Open Information Technologies*, 2017, vol. 5, no. 10, pp. 55–86 (In Russ.).
6. Savaget P., Geissdoerfer M., Kharrazi A., Evans S. The theoretical foundations of sociotechnical systems change for sustainability: A systematic literature review. *Journal of Cleaner Production*, 2019, vol. 206, pp. 878–892. DOI: 10.1016/j.jclepro.2018.09.208.
7. Saak A. E., Tyushnyakov V. N., Pakhomov E. V. *Models of the information technology structure of a smart city*. Fundamental'nye issledovaniya [Fundamental Research], 2017, no. 10-2, pp. 387–391 (In Russ.).
8. Kitchin R. Data-driven, networked urbanism. *Steps*, 2017, vol. 3, no. 2, pp. 98–116.
9. Aliero M. S., Qureshi K. N., Pasha M. F., Jeon G. Smart home energy management systems in internet of things networks for green cities demands and services. *Environmental Technology & Innovation*, 2021, vol. 22. Available at: <https://www.sci-hub.ru/10.1016/j.eti.2021.101443> (accessed 06.17.2024).
10. Frank E., Aznar Fernández-Montesinos G. Smart City = Smart Citizen = Smart Economy? An Economic Perspective of Smart Cities. *Social, Legal, and Ethical Implications of IoT, Cloud, and Edge Computing Technologies*, IGI Global, Hershey. In G. Cornetta, A. Touhafi, G. M. Muntean (eds.), 2020, chap. 7, pp. 161–180.
11. Ismagilova E., Hughes L., Dwivedi Y. K., Raman K. R. Smart cities: advances in research – an information systems perspective. *International Journal of Information Management*, 2019, vol. 47, pp. 88–100.
12. Sukhovskaya D. N., Shulgin N. A. Analysis of the structure of the "smart city" concept. *E-Scio*, 2019, no. 9(36), pp. 737–746 (In Russ.).
13. Kamolov S. G., Korneeva A. M. *Technologies of the future for "smart cities"*. Vestnik Moskovskogo gosudarstvennogo oblastnogo universiteta. Serija: Jekonomika [Bulletin of the Moscow State Regional university. Series: Economics], 2018, no. 2, pp. 100–114 (In Russ.).
14. Yang R., Zhen F. Smart city development Models: A cross-cultural regional analysis from theory to practice. *Research in Globalization*, 2024, vol. 8, p. 100221. DOI: 10.1016/j.resglo.2024.100221.
15. Bibri S. E., Krogstie J. A novel model for data-driven smart sustainable cities of the future: A strategic roadmap to transformational change in the era of big data. *Future Cities and Environment*, 2021, vol. 7, no. 1, pp. 1–25.
16. Kostina E. A., Kostin A. V. *Barriers to using feedback services communications at the municipal level*. Zhurnal Sibirskogo federal'nogo universiteta. Gumanitarnye nauki [Journal of the Siberian Federal University. Humanities], 2023, vol. 16, no. 3, pp. 468–480 (In Russ.).
17. Pirotskaya A. V. *The concept of a "smart city" in the view of experts and residents: problems and contradictions in the implementation of the concept*. Mir jekonomiki i upravlenija [The world of Economics and Management], 2020, vol. 20, no. 2, pp. 178–196 (In Russ.).
18. Kitchin R. The opportunities, challenges and risks of big data for official statistics. *Statistical Journal of the International Association of Official Statistics*, 2015, vol. 31, no. 3, pp. 471–481.
19. Junfeng X. J., Tang H., Huang T., Yu F. R., Xie R., ..., Liu Y. A survey of blockchain technology applied to smart cities: Research issues and challenges. *IEEE Communications Surveys & Tutorials*, 2019, vol. 21, iss. 3, pp. 2794–2830.

Информация об авторах

Кругликов Сергей Владимирович, доктор военных наук, кандидат технических наук, доцент, генеральный директор, Объединенный институт проблем информатики НАН Беларуси.

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

Давыдовский Анатолий Григорьевич, кандидат биологических наук, доцент, доцент кафедры микропроцессорных систем и сетей Института информационных технологий БГУИР.

E-mail: agd2011@list.ru

Information about the authors

Sergey V. Kruglikov, D. Sc. (Milit.), Ph. D. (Eng.), Assoc. Prof., Director General, The United Institute of Informatics Problems of the National Academy of Sciences of Belarus.

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

Anatoly G. Davydovskii, Ph. D. (Biol.), Assoc. Prof., Assoc. Prof. of the Department of Microprocessor Systems and Networks of the Institute of Information Technologies of BSUIR.

E-mail: agd2011@list.ru

ОБРАБОТКА СИГНАЛОВ, ИЗОБРАЖЕНИЙ, РЕЧИ, ТЕКСТА И РАСПОЗНАВАНИЕ ОБРАЗОВ

SIGNAL, IMAGE, SPEECH, TEXT PROCESSING AND PATTERN RECOGNITION



UDC 004.93
DOI: 10.37661/1816-0301-2025-22-2-33-47

Original Article
Оригинальная статья

Efficient detection of building in remote sensing images using an improved YOLOv10 network

Xiangyi Wu^{1✉}, Sergey V. Ablameyko^{1,2}

¹*Belarusian State University,
av. Nezavisimosti, 4, Minsk, 220030, Belarus*
[✉]*E-mail: tigerv5872@gmail.com*

²*The United Institute of Informatics Problems
of the National Academy of Sciences of Belarus,
st. Surganova, 6, Minsk, 220012, Belarus*

Abstract

Objectives. At present, rapid detection of the location and size of building objects from remote sensing images is important for scientific research value and has practical significance for urban planning, environmental monitoring and disaster management.

Methods. This paper proposes an object detection method based on improved YOLOv10 network, which incorporates Super Token Attention, RepConv and Normalized Weighted Distance to more precisely detect buildings in remote sensing images. This method improves the detection accuracy and efficiency especially for small objects. The LEVIR-CD dataset is used for model training and testing.

Results. The experimental results show that the method demonstrates better accuracy on the building detection task than the traditional YOLOv10 and other methods.

Conclusion. The proposed method significantly enhances the accuracy and efficiency of building detection in remote sensing images.

Keywords: YOLOv10, remote sensing images, attention mechanism, building detection, RepConv, Super Token Attention

For citation. Wu X., Ablameyko S. V. *Efficient detection of building in remote sensing images using an improved YOLOv10 network*. Informatika [Informatics], 2025, vol. 22, no. 2, pp. 33–47. DOI: 10.37661/1816-0301-2025-22-2-33-47.

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

Received | Поступила в редакцию 29.04.2025
Accepted | Подписана в печать 13.05.2025
Published | Опубликована 30.06.2025

Эффективное обнаружение зданий на изображениях дистанционного зондирования на основе улучшенной сети YOLOv10

С. Ву^{1✉}, С. В. Абламейко^{1, 2}

¹Белорусский государственный университет,
пр. Независимости, 4, Минск, 220030, Беларусь
✉E-mail: tigerv5872@gmail.com

²Объединенный институт проблем информатики
Национальной академии наук Беларуси,
ул. Сурганова, 6, Минск, 220012, Беларусь

Аннотация

Цели. В настоящее время быстрое определение местоположения и размера объектов зданий с помощью изображений дистанционного зондирования имеет важное научно-исследовательское и практическое значение для городского планирования, мониторинга окружающей среды и управления стихийными бедствиями.

Методы. Предлагается метод обнаружения объектов на основе улучшенной сети YOLOv10, которая включает в себя механизм внимания Супертокен, модель RepConv (повторно параметризованная свертка) и нормализованное взвешенное расстояние для более точного обнаружения зданий на изображениях дистанционного зондирования. Метод повышает точность и эффективность обнаружения, особенно для небольших объектов. Набор данных LEVIR-CD используется для обучения и тестирования модели.

Результаты. Экспериментальные результаты показывают, что предлагаемый метод демонстрирует лучшую точность при решении задачи обнаружения зданий, чем традиционный YOLOv10 и другие методы.

Заключение. Предлагаемый метод эффективно повышает точность и эффективность обнаружения зданий на изображениях дистанционного зондирования.

Ключевые слова: YOLOv10, изображения дистанционного зондирования, механизм внимания, обнаружение зданий, RepConv, Super Token Attention

Для цитирования. Ву, С. Эффективное обнаружение зданий на изображениях дистанционного зондирования на основе улучшенной сети YOLOv10 / С. Ву, С. В. Абламейко // Информатика. – 2025. – Т. 22, № 2. – С. 33–47. – DOI: 10.37661/1816-0301-2025-22-2-33-47.

Конфликт интересов. Авторы заявляют об отсутствии конфликта интересов.

Introduction. With the development of aerospace technology, remote sensing image resources play an increasingly important role in many fields such as urban planning, military reconnaissance, and land resource management [1]. In these applications, the detection and identification of buildings, as an important landmark feature on the ground in urban or suburban areas, has become a key topic.

Traditional object detection methods rely on empirically based feature hand design, and these methods are often difficult to mine feature information in higher dimensions of the image and have high computational costs when dealing with high-resolution remote sensing images [2]. In recent years, the development of deep learning technology has brought revolutionary advances in the field of image processing, especially in object detection, semantic segmentation and image classification, which have achieved remarkable results. However, in building detection in remote sensing images, there are still some problems especially to detect building in very high-resolution remote sensing images:

1) Different sources of images, the original pictures have different sizes of objects in remote sensing images because of different heights and angles of detectors, the influence of weather, the refraction of the atmosphere, the time of imaging, the curvature of the earth, and other factors. This increases the difficulty of object detection [3].

2) The background in remote sensing images is usually complex and varied, including roads, vegetation, lake waves, and so on. These background disturbances increase the difficulty of object detection, and high resolution remote sensing images also imply more complex backgrounds [4].

3) Occlusion problems, serious interference caused by shadows and occlusion due to shooting weather or angle, lighting conditions may change with time and weather. Changes in lighting can lead to changes in the appearance of the object, thus increasing the difficulty of object detection [5].

4) Building textures are relatively homogeneous, usually with only one or two outlines or colour information [6].

5) Remote sensing images usually have large-scale data volume and need to process a large amount of image data. This poses a challenge to the computational efficiency and storage space of object detection algorithms.

Fig. 1 illustrates the problems of building detection in several scenarios.



Fig. 1. Building in different scenarios in remote sensing images

Small object detection is an important research direction and also a difficult point in object detection, there are many researchers who have conducted a lot of research on small object detection and they have proposed many methods from feature fusion, contextual connectivity, and adversarial learning for improving the performance of small object detection [7]. With the development of deep learning techniques, different attention mechanism techniques such as EfficientNetV2, CBAM, RepViTblock, and EMA attention are beginning to be applied to small object detection tasks.

EfficientNetV2 achieves faster training speed and better parameter efficiency by training perceptual neural architecture search and expansion [8]. CBAM (Convolutional Block Attention Module) enhances the sensitivity of the network to small objects through the attentional mechanism, and has shown its effectiveness in several remote sensing image detection tasks its effectiveness [9]. RepViTblock is a lightweight new backbone network [10], which revisits mobile CNNs from the perspective of ViT, showing a superior balance of latency and accuracy. EMA attention, on the other hand, enhances the model's feature extraction capability for small objects through weighted averaging [11].

With the continuous iterative updating of new technologies, object detection techniques are also evolving. Future research directions may include further improving the generalisation ability of the model, increasing technology fusion with multi-task learning, and developing more efficient training algorithms to reduce the demand for computational resources. It will also be a way to improve the performance of small object detection. Of course, small object detection in high-resolution remote sensing images is still challenging.

The YOLO series of networks have better balance detection accuracy and speed by capturing the deep and high-level features of objects. They have been selected for optimization and improvement to complete small object detection tasks in complex scenes such as satellite remote sensing data. The YOLO networks have dominated the field of object detection with its superior performance and efficiency [12]. YOLOv10, the latest in the series, has been designed to achieve state-of-the-art performance through the introduction of a consistent dual allocation and optimised model components, eliminating the need for non-maximum suppression (NMS) and significantly reducing computational overhead while achieving state-of-the-art performance [13]. However, detecting tiny objects is a very challenging problem [14].

In order to further promote the development of small object detection, this paper proposes an algorithm to optimize the backbone network and improve the attention mechanism to solve the problem of poor small object detection. In this paper, we propose to integrate STA (Super Token Attention) [15], RepConv (Re-Parameterizable Convolution) [16] and NWD (Normalized Weighted Distance) [17] techniques into the YOLOv10 model by adjusting the parameters in order to adapt the accuracy of building object detection in remote sensing images. Firstly, RepConv is applied in the convolutional layer of YOLOv10 model to simplify the network structure by reparameterizing the convolutional and batch normalisation layers. And the model parameters are reduced by RepConv to improve the inference speed and accuracy of the model. Then, the backbone network of the improved YOLOv10 model is used to extract features from the input image, and Super Tokens are generated on the feature map through the STA mechanism, and the Super Tokens are used to adjust the model's attention distribution so that the model pays more attention to the regions that contain information about buildings. Finally, the NWD loss function is introduced to calculate the weighted distance between the predicted frame and the real frame to optimise the model's ability to locate small objects and improve the accuracy of object detection for buildings.

Method

Overview of YOLOv10. YOLOv10 is the latest real-time end-to-end target detection model that significantly improves the performance and efficiency of the YOLO family by optimising post-processing and model architecture. Based on the C2f (Cross Stage Partial fusion) structure of YOLOv8, the model introduces CIB (Channel Interaction Block) to strengthen the inter-channel information interaction; and combines the spatial pyramid with the attention mechanism PSA (Pyramid Spatial Attention) to optimise the Multi-scale feature fusion; SCDown (Spatial-Channel Downsampling) module is introduced to jointly optimise the downsampling process in spatial and channel dimensions, the specific structure is shown in fig. 2. Non-maximum suppression free training is achieved and the inference delay is reduced. In addition, YOLOv10 introduces an integrated efficiency-accuracy oriented model design strategy that reduces computational redundancy and improves model capability.

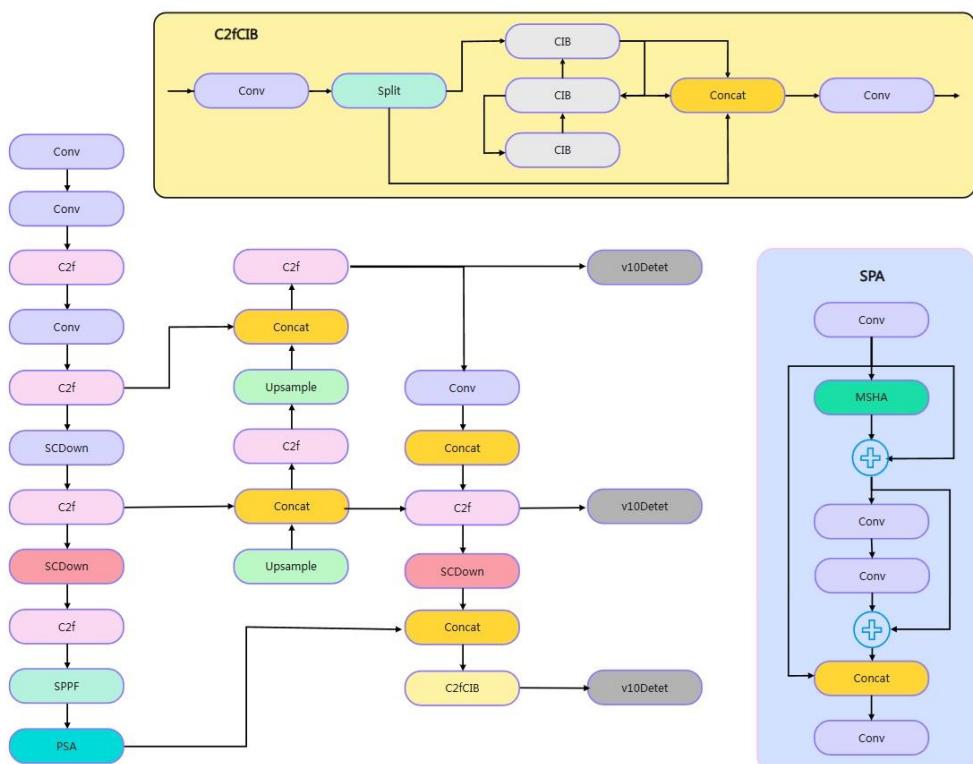


Fig. 2. YOLOv10 structure

In this study, we chose YOLOv10n as the baseline model because it provides a better balance of performance and efficiency while maintaining a smaller model size, and is suitable to be used as a starting point for improving the algorithm for further exploring and optimising the detection of building objects in remote sensing images.

Currently, buildings in remotely sensed imagery face many problems, such as: a large number of small building targets and a mixture of buildings of different sizes; and the background of remotely sensed imagery contains a variety of feature types, such as forests, fields, and roads. Together, these factors make it difficult to automatically extract building information from images. In addition, buildings have different appearances, colours and proportions, and targets such as buildings may be naturally obscured by trees etc., resulting in incomplete or partially visible targets, which is often difficult to cope with by traditional recognition methods, and this puts forward higher requirements for target recognition and localisation. Therefore, this paper proposes an improved method that fuses RepConv, Super Token Attention and NWD techniques into the YOLOv10 model, hereafter referred to as YOLO-RSTA.

RepConv. RepConv (Re-Parameterizable Convolution) is a model reparameterization technique that improves the efficiency and performance of models by optimising the network structure in the field of deep learning. Satellite remote sensing images often have complex backgrounds and varying building shapes and sizes. The standard Conv layer may not capture all the necessary features for accurate building detection. RepConv, with its multi-branch structure, can extract more diverse and comprehensive features, helping the model better distinguish buildings from the background and other objects.

The core idea of RepConv is to use multi-branched convolutional layers in the training phase, which can include convolutional kernels of different sizes (e.g., 1×1 , 3×3 , etc.) and possibly BN layers. In the inference phase, the parameters of these branches are reparameterized onto a master branch, usually an equivalent 3×3 convolutional layer, thus reducing computation and memory consumption and increasing inference speed [18]. This is shown in fig. 3 below.

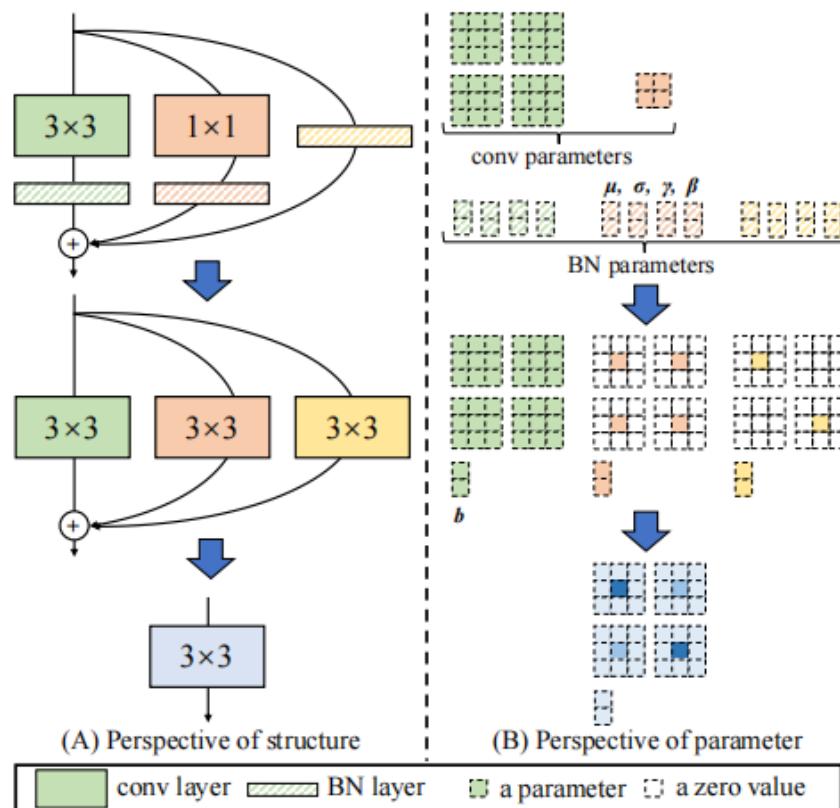


Fig. 3. RepConv Schematic Illustration

RepConv works based on the following formula:

$$W_t = T(W_b, W_r), \quad (1)$$

where W_t are transformed weights, used to manipulate input features, and W_b is (base weights), frozen weights inherited from the pre-trained model. W_r is refocusing weights, additional trainable parameters introduced by the refocusing transform. $T(*)$ is the refocusing transform, a trainable operation applied to the base weights to generate new weights. Operation for generating new weights. RepConv also introduces residual concatenation, which makes the refocusing transform learn an increment of the base weights instead of the original mapping, similar to residual blocks in *ResNet*:

$$W_t = W_b * W_r + W_b, \quad (2)$$

where $*$ denotes a convolutional operation. RepConv enhances the model's use of prior knowledge through the refocusing transform, allowing the model to focus on different feature representations encoded in the pre-trained model, thus improving feature extraction for small targets.

With the reparameterization technique, RepConv is able to simplify the multi-branch structure into a single-branch structure during inference, which significantly reduces the computation and memory consumption and improves the inference speed.

The improved algorithm in this paper replaces the first and third Conv convolution layers with RepConv layers in the YOLOv10 backbone network. The purpose is to use the feature map output by the previous convolution layer as input, perform a reparameterization operation through the RepConv layer, and then perform a convolution operation through the subsequent convolution layer. At this time, the number of channels of the output feature map does not change. This replacement method can reduce the size of the spatial dimension without losing information. It retains more information in the channel than traditional convolution operations, so it can effectively improve the feature extraction ability for small targets.

Super Token Attention (STA). Satellite images usually have buildings of various shapes and scales, and their edges are blurred due to shadows or vegetation. Traditional convolutional layers have limited ability to capture global dependencies, making it difficult to distinguish between buildings and backgrounds and other objects. Therefore, in remote sensing image building target detection, global context information is crucial for accurate recognition. For this reason, the fusion of STA mechanism is proposed to enhance the global perception capability of the YOLOv10 model. STA enhances the detection performance by aggregating global information in the image to form Super Tokens and using these tokens to guide the attention allocation.

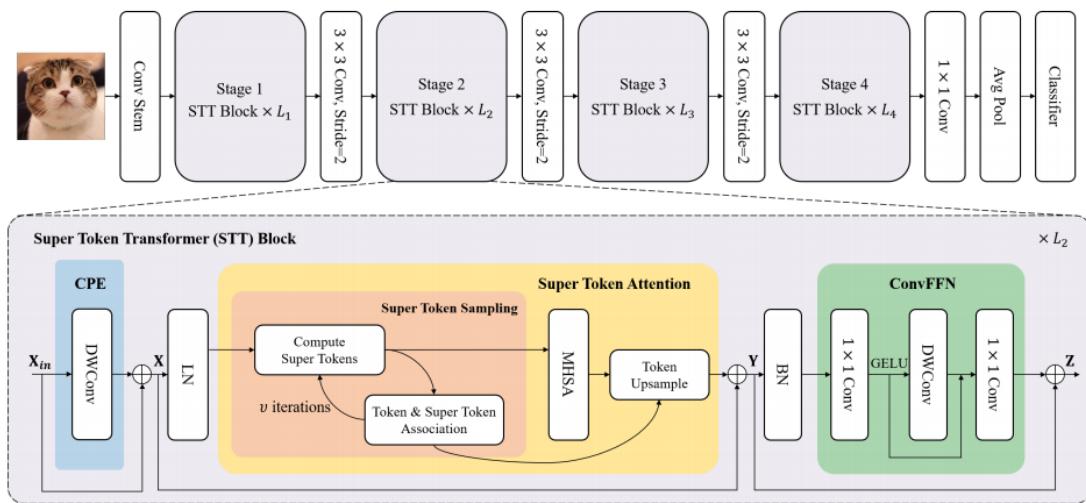


Fig. 4. Super Token Vision transformer architecture diagram

The core of the STA mechanism is to identify and utilise the global features in the image, the structure of which is shown in fig. 4. Firstly, global contextual information is extracted through the self-attention layer, which is then combined with local features to form supertags. These supertags not only contain rich semantic information, but also represent key regions in the image [19]. In the attention allocation stage, the model dynamically adjusts the attention weights according to the importance of the supertags, which enables the model to pay more attention to the regions containing building information.

The main principle of STA is divided into the following 3 steps:

Super Token Sampling (STS): the STS is a key step in the STA mechanism, which aims to aggregate the visual tokens in the input image into super tokens to reduce the number of tokens in the self-attention operation while preserving global contextual information. This process mimics the idea of superpixels, i.e., segmenting an image into regions with similar features to simplify subsequent processing. Specifically, the algorithm first generates an initial set of superlabels through average pooling, and then updates the superlabels through an iterative process. In each iteration, the algorithm calculates the association matrix between each pixel feature and the superlabelled feature, and then uses this association matrix to update the superlabels. This process can be represented by the following equation:

$$Q_t = \text{Softmax}\left(\frac{XS^{t-1^T}}{\sqrt{d}}\right), \quad (3)$$

where Q is the correlogram computed in the t iteration, $\text{Softmax}()$ function that converts a set of values into a probability distribution. Its central role is to make all output values sum to 1 by normalizing them so that each value is between 0 and 1. X is the visual marker, S^{t-1} is the supertag from the previous iteration, and d is the square root of the number of channels C for normalization. STS progressively improves the discriminative properties of Super Tokens by optimising the association matrix over multiple iterations. This process, combined with the self-attention mechanism, allows the model to reduce the computational complexity: reducing the $O(N^2)$ self-attention computation to $O(NK)$ ($K \ll N$).

Self-Attention for Super Tokens: Performs a self-attention operation in the space of super tokens to capture long-term dependencies between super tokens. This step uses the standard self-attention mechanism to compute an attention graph between super tokens, and uses this attention graph to perform a weighted summation of super tokens to obtain a new super token representation. The core formula is:

$$\text{Attn}(S) = \text{Softmax}\left(\frac{q(S)K^T(S)}{\sqrt{d}}\right)v(S), \quad (4)$$

where $q(S)$, $K(S)$ and $v(S)$ are the supertagged query, key and value, respectively, and d is the square root of the number of channels C . The $\text{Softmax}()$ function for normalised correlation matrix.

Token Upsampling (TU): Finally, the learned association map is used to map the hyperlabelling back into the visual marker space. This step fuses the information from the hyperlabelling back into the original pixel features so that the model can use the global contextual information for more accurate detection. The formula is as follows:

$$TU(\text{Attn}(S)) = Q\text{Attn}(S), \quad (5)$$

where Q is the association map for mapping the hyperlabelling back into the visual labelling space.

In this way, STA is able to reduce the computational complexity in self-attention while enhancing the model's ability to capture global context by aggregating local features. In remote sensing image detection, this means that the model is able to handle large-scale information in the image more effectively, improving the detection of small targets and complex scenes.

Normalized Weighted Distance (NWD). NWD is a method for measuring the similarity between predicted and real bounding boxes in object detection tasks. It evaluates the proximity of two bounding boxes by calculating a weighted distance between them. Compared to the traditional IoU (intersection and integration ratio) metric, NWD is able to capture the positional relationship between bounding boxes more accurately, especially when dealing with small objects or low overlap. The core idea of NWD is to represent the bounding box as a Gaussian distribution and then calculate the Wasserstein distance between these two distributions. This distance metric is more sensitive to changes in the position and scale of the bounding box, thus providing a more accurate similarity assessment in object detection [20].

The Wasserstein distance is used to calculate the distance between two probability distributions, i.e. to assess the minimum cost (the minimum of the mean distance of a move) required to convert one distribution into the other. For two two-dimensional Gaussian distributions $N(\mu_1, \Sigma_1)$ and $N(\mu_2, \Sigma_2)$ the Wasserstein distance between them is:

$$W_2(N(\mu_1, \Sigma_1), N(\mu_2, \Sigma_2)) = \sqrt{\|\mu_1 - \mu_2\|^2 + \text{tr}(\Sigma_1 + \Sigma_2 - 2\sqrt{\Sigma_1 \Sigma_2})}, \quad (6)$$

where μ_1 and μ_2 are mean vectors, Σ_1 and Σ_2 are covariance matrices, and tr denotes the sum of the diagonal elements of the matrix. In NWD, this concept is used to calculate the distance between two bounding boxes corresponding to a Gaussian distribution.

Since the Wasserstein distance itself is an unbounded distance metric, NWD converts it to a similarity metric between 0 and 1 by normalizing it in exponential form, where 0 means exactly the same and 1 means completely different. To convert the Wasserstein distance into a similarity metric, NWD uses the exponential form of normalization:

$$\text{NWD} = \exp\left(-\frac{W_2}{C}\right), \quad (7)$$

where C is a constant closely related to the dataset that is used to adjust the scale of normalization. In practice, C is usually set to the average absolute size of the targets in the dataset for optimal performance.

NWD provides better scale invariance, smoothing of positional deviations, and the ability to measure similarity between non-overlapping or mutually inclusive bounding boxes than traditional IoU metrics. In addition, NWD is scale-invariant, allowing it to maintain consistent performance in the detection of objects of different sizes.

The proposed method. Our proposed method YOLO-RSTA is a fusion of STA, RepConv and NWD techniques into the model of YOLOv10, and the detailed structure is shown in fig. 5.

Since the original YOLOv10 improves the model's detection ability for targets at different scales, but the detection accuracy in the background complexity and occlusion problem is not satisfactory enough. Therefore, we replace the convolution (Conv) of P1/2 and P2/4 layers with RepConv to enhance the feature extraction capability. The structure of RepConv can capture richer feature information during training, and feature extraction at different scales is more flexible and effective, and in the process of extracting feature maps at different scales, the key features of the target can be captured more effectively. Capture the key features of the target in the process of extracting feature maps at different scales.

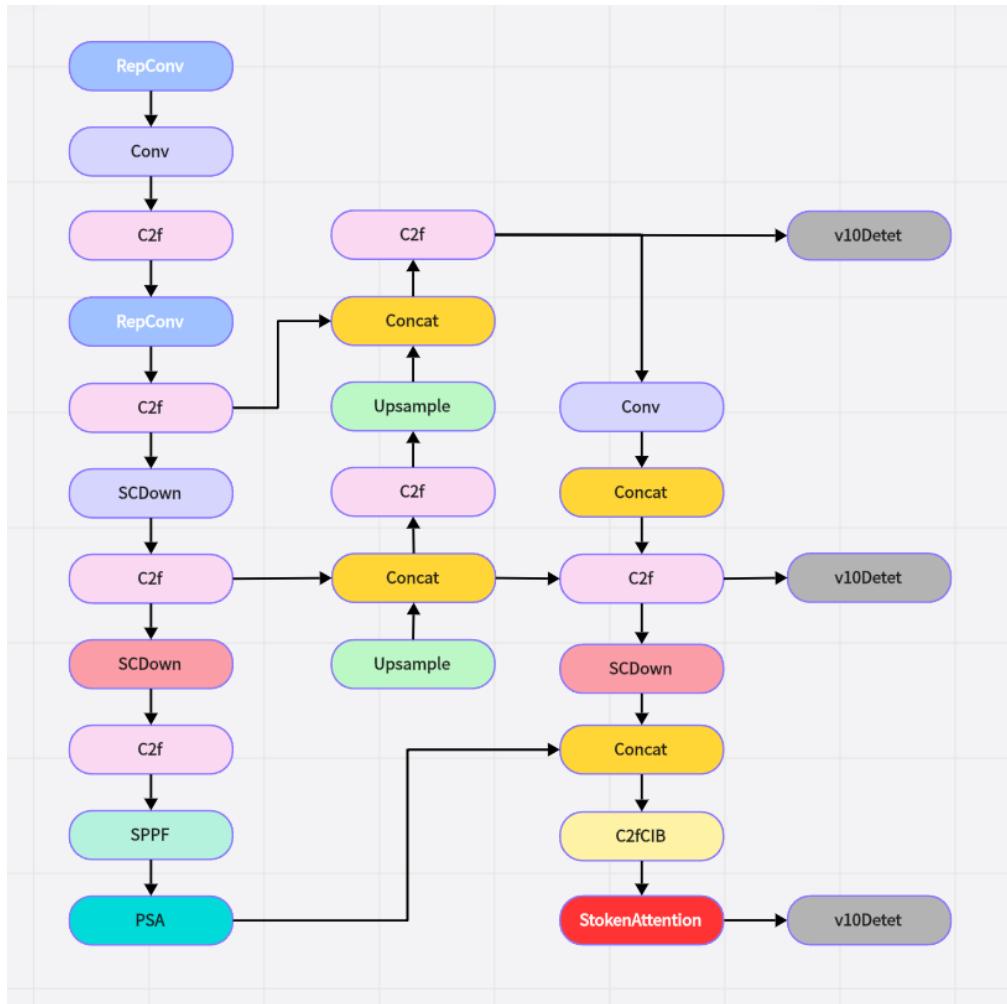


Fig. 5. Structural diagram of the improved YOLO-RSTA network

Meanwhile, we enable the NWD loss function, which will be used by the model to measure the similarity between the predicted bounding box and the real bounding box. However, the introduction of NWD leads to changes in accuracy, so the IoU ratio parameter needs to be adjusted to find the optimal balance of the loss function, which sacrifices some of the accuracy but improves on both Recall and mAP.

Finally, we added the STA attention mechanism after the P5/32 layer, which aims to improve the accuracy and efficiency in detecting small targets and processing complex scenes. In complex scenes, background interference is a major factor affecting the detection of small targets. STA reduces the interference of background noise on the model's decision making by focusing on key regions in the image. STA enhances the feature representation by applying self-attention to the super tokens and STA helps the model to capture these features at multiple scales, which improves the detection of small targets in complex backgrounds.

Experimental results

Experimental Environment Setting. In this experiment, the experiment is mainly conducted on the LEVIR-CD dataset, which contains large-scale remote sensing data. In this paper, PyTorch version 2.4.0 and the corresponding CUDA version 2.4.1 are used as the experimental environment. According to the above method, we use NVIDIA GeForce RTX 4060 as GPU for training, 445 images in the training set, 64 images in the validation set, and 128 images in the test set. The image resolution is 1024×1024 . The objects in the dataset are mostly small object buildings. The specific settings are shown in fig. 6.

Python	CUDA	PyTorch	GPU	CPU
Python 3.9	12.4.1_551.78	2.4.0	NVIDIA GeForce RTX 4060	Intel Core i5-12400F

Fig. 6. Experimental environment

Data sets. Group B of the LEVIR-CD [21] dataset, a large-scale remotely sensed building change detection dataset provided by LEVIR Labs, was selected for the experiments in this study. The LEVIR-CD contains 637 pairs of very high-resolution (0.5 metres/pixel) Google Earth image blocks, with the size of each pair of images being 1024×1024 pixels, as shown in fig. 7. The LEVIR-CD covers a wide range of building types, including cottage homes, high-rise flats, small garages, and large warehouses. The LEVIR-CD dataset contains a total of 31,333 individual instances of building change. The geographic distribution of the LEVIR-CD dataset covers several cities in the state of Texas in the United States, including Austin, Lakeway, Bee Cave, Buda, Kyle, Manor, Pflugerville, Dripping Springs, etc., with image data captured from 2002 to 2018.



Fig. 7. LEVIR-CD dataset

Evaluation indicators. The experiments were conducted to evaluate the algorithm's target detection performance using mean accuracy (mAP), precision (P), recall (R), and parametric counts (params). The formulas for precision, recall and mAP are shown in eqs. (8) to (10).

$$P = \frac{TP}{TP + FP}, \quad (8)$$

$$R = \frac{TP}{TP + FN}, \quad (9)$$

$$\text{mAP} = \frac{1}{n} \sum_{i=1}^n AP_i \quad (10)$$

In the formula: P is precision; R is recall; TP is the number of samples predicted to be positive and actually positive; FP is the number of samples predicted to be positive but actually negative; FN is the number of samples predicted to be negative but actually positive; AP (Average Precision) is a

numerical representation of the area of a certain type of object under the Precision-Recall Curve. mAP (Mean Average Precision) is the average value obtained by averaging the average precision (AP) of different types of object detection. In formula (10), it is necessary to first calculate the current AP_i of different types of objects based on P and R , and take the average value of $n AP_i$ as mAP.

Parameters is the size of the model's trainable parameters. The mAP50 used in the experiment represents the average detection accuracy when the IoU threshold is 0.5, while mAP50-95 represents the average accuracy of multiple thresholds in the range of IoU thresholds from 0.5 to 0.95 (step size 0.05), which can more comprehensively evaluate the performance of the model under different positioning accuracies.

Experimental results. We choose YOLOv10n as the baseline for experiments. We train for 300 epochs in the whole network structure. We need to evaluate the impact of different improvement strategies on the performance of the YOLOv10 model for building object detection in remote sensing images.

The main evaluation metrics we focus on include the model's precision, recall, and mean average precision (mAP) under different IoU thresholds. With these metrics, we can comprehensively evaluate the accuracy of the model in detecting buildings.

Experiment 1 was conducted on the LEVIR-CD dataset using YOLOv5n, YOLOv8n, YOLOv10n, and the improved method in this paper. The experimental metrics are shown in table 1.

Table 1
 Results of LEVIR-CD experiment 1

Model-name	Parameters	Precision	Recall	mAP50	mAP50-95
YOLOv5n	2 503 139	0.884	0.817	0.9	0.577
YOLOv8n	3 005 843	0.883	0.829	0.909	0.582
YOLOv10n	2 694 806	0.858	0.827	0.898	0.572
YOLO-RSTA	2 957 368	0.891	0.845	0.914	0.571

In terms of precision, YOLO-RSTA leads with the highest value of 0.891, proving its advantage in reducing false alarms and predicting positive samples more accurately. In terms of recall, YOLO-RSTA also performs the best at 0.845, implying that it is better at identifying actual positive samples and covering the real change scenarios. mAP50 metrics, YOLO-RSTA and YOLOv8n occupy the top two positions respectively, indicating that the two models are better at detecting at the IoU threshold of 0.5, with YOLO-RSTA having a slight edge over the others. However, in the mAP50-95 metrics, YOLOv8n is outstanding with the highest value of 0.582. In summary, YOLO-RSTA performs well in the key indicators of precision, recall and mAP50.

Experiment 2 compares the proposed method YOLO-RSTA with some common attention mechanisms (EfficientNetV2, CBAM, RepViT block and EMA attention) and the results are shown in table 2 below.

Table 2
 Results of building detection for LEVIR-CD dataset

Model-name	Parameters	Precision	Recall	mAP50	mAP50-95
YOLOv10n	2 694 806	0.858	0.827	0.898	0.572
YOLOv10n-EMA_attention	3 027 350	0.864	0.797	0.891	0.571
YOLOv10n-RepViTblock	3 027 350	0.862	0.82	0.897	0.57
YOLOv10n-CBAM	2 703 369	0.872	0.784	0.869	0.515
YOLOv10n-SE	2 695 318	0.853	0.779	0.881	0.569
YOLOv10n-EfficientNetv2	2 544 510	0.853	0.784	0.869	0.515
YOLO-RSTA	2 957 368	0.891	0.845	0.914	0.571

In this building detection experiments, our modified model demonstrates significant performance improvement compared to other models. In terms of precision, YOLO-RSTA reaches 0.891, which is about 0.8 % improvement compared to YOLOv5n and YOLOv8n, and about 3.8 % improvement compared to YOLOv10n. In terms of recall, YOLO-RSTA is a distant second with 0.845, an improvement of about 3.4 % over YOLOv5n's 0.817, and an improvement of about 2.3 % compared to YOLOv10n's 0.827. In the mAP50 metric, YOLO-RSTA is slightly higher than YOLOv8n, with an improvement of about 1.6 % over YOLOv5n's 0.9 and about 1.8 % over YOLOv10n's 0.898. Despite the slight increase in the number of parameters of YOLO-RSTA, its inference time of 1.946 seconds is still within the acceptable range. Overall, the STA attention mechanism significantly improves the model's detection accuracy and recall, albeit at the slight expense of inference speed, but this trade-off may be worthwhile in real-world applications as it achieves performance gains in key performance metrics.

From the above data, it can be seen that most of the YOLOv10 models with the introduction of different attention mechanisms can lead the original YOLOv10n model in terms of accuracy, and the model proposed in this paper improves by about 2.21 % in terms of accuracy compared to the YOLOv10n model and outperforms the other models with attention mechanisms, which also indicates that the model has a significant effect in terms of reducing false detections. In recall, the model improves by about 2.05 % over the original YOLOv10n and performs the best, indicating that the model also improves in detecting more real buildings. The model also manages to improve the mAP50 metric by about 1.67 % compared to the original YOLOv10n model, suggesting that the model can improve the detection performance overall.

Experiment 3. The proposed YOLO-RSTA method is subjected to an ablation experiment, this ablation experiment aims to explore the effects of RepConv, STA and NWD on the performance of the YOLOv10n model. The results are shown in table 3 below.

Table 3
Results of ablation experiments

YOLOv10n	Repconv	STA	NWD	Parameters	Precision	Recall	mAP50	mAP50-95
✓				2 694 806	0.858	0.827	0.898	0.572
✓	✓			2 694 806	0.871	0.79	0.884	0.57
✓		✓		2 957 368	0.884	0.824	0.901	0.57
✓		✓	✓	2 957 368	0.877	0.844	0.913	0.57
✓	✓	✓		2 957 368	0.861	0.797	0.894	0.571
✓	✓	✓	✓	2 957 368	0.891	0.845	0.914	0.571

When only RepConv is enabled, Precision improves from 0.858 to 0.871, but Recall decreases to 0.79, indicating that RepConv improves detection accuracy but loses some of the true change detection. When only Super Token Attention is enabled, the number of parameters increases to 2 957 368, Precision improves to 0.884, and Recall is 0.824, showing that it enhances model detection. When RepConv and STA are enabled at the same time, Precision is 0.877 and Recall is 0.844, indicating that the combination of the two can effectively improve Recall and maintain a high Precision, and when RepConv and NWD are enabled, Precision is 0.861 and Recall is 0.797, which is a slight decrease in Recall compared to enabling RepConv only. When all three modules are enabled, the optimal performance is achieved with Precision of 0.891 and Recall of 0.845, but the number of parameters increases to 2 957 368. The experiments show that all three modules have positive effects on the performance improvement of the YOLOv10n model, and that all of them achieve the best results on the key metrics when they are all enabled, although they increase the number of parameters.

In order to comprehensively analyze the model efficiency, this paper compares the computational complexity indicators of the baseline model (YOLOv10n), other YOLO series models (YOLOv5n, YOLOv6n, YOLOv8n) and the improved model (YOLO-RSTA) (as shown in table 4). Specifically including:

GLFOPs (Giga Floating-Point Operations): represents the billion floating-point operations required for the model forward reasoning, used to measure the computational complexity;

Inference Time: The total processing time (unit: milliseconds) of a single image from input to output prediction box consists of the time of four stages: pre-processing time, inference time, loss time, and post-processing time for each image.

Table 4
 Computational complexity experiments

Model-name	Parameters	Preprocess	Inference	Loss	Postprocess	GLFOPs
YOLO-10n	2 694 806	0.4	3.5	0	1.8	8.2
YOLO-5n	2 503 139	0.4	2.6	0	10	7.1
YOLO-6n	4 233 843	0.3	2.6	0	10.3	11.8
YOLO-8n	3 005 843	0.4	2.7	0	9.5	8.1
YOLO-RSTA	2 957 368	0.3	3.6	0	0.7	8.2

Experimental results show that the number of parameters of YOLO-RSTA increases by about 9.7 % compared with the baseline, but the GLFOPs remains unchanged at 8.2, which is close to YOLOv8n, but significantly lower than YOLOv6n, indicating that it maintains a lightweight design. The GLFOPs of YOLO-RSTA is 8.2, which is the same as YOLOv10n and YOLOv8n, and better than YOLOv6n's 11.8, proving that the improved module does not introduce additional computational burden. The inference time of YOLO-RSTA is 3.6 ms, slightly higher than YOLOv8n's 2.7 ms, but significantly better than YOLOv5n's (2.6 ms) post-processing efficiency. The post-processing optimization of YOLO-RSTA benefits from the suppression of redundant prediction boxes by the NWD loss function. The post-processing time of YOLO-RSTA is 0.7 ms, which is 61 % lower than the 1.8 ms of the baseline YOLOv10n model, and a significant improvement over YOLOv5n. YOLO-RSTA significantly improves the accuracy (mAP50: 0.914) and recall rate (0.845) while maintaining reasonable computational efficiency, making it suitable for actual remote sensing image processing scenarios.

Conclusion. This study proposes an enhanced remote sensing image building detection framework YOLO-RSTA, which innovatively integrates RepConv, super-labeled attention mechanism (STA) and optimized NWD loss function into YOLOv10. Experimental validation on the LEVIR-CD dataset demonstrates the superiority of the method, achieving the best available performance of 0.891 accuracy, 0.845 recall, and 0.914 mAP50 while maintaining an efficient parameter utilisation (2.95 M). In terms of computational complexity, YOLO-RSTA maintains the same amount of computation as the baseline model with only a 9.7 % increase in parameter size, and optimises the post-processing process with NWD, reducing the post-processing time by 61 %. This demonstrates that the proposed method improves the accuracy without obviously sacrificing the computational efficiency, and is suitable for real remote sensing image processing scenarios.

The performance improvement mainly comes from three key improvements: (1) The RepConv module enhances the feature representation ability, and the mAP50 is improved by 3.0 % compared with the baseline; (2) The STA mechanism effectively models the global contextual relationship in complex remote sensing scenes, and outperforms traditional attention modules (such as CBAM, SE, etc.) in terms of the balance between accuracy and recall; (3) The NWD loss function significantly improves the localization accuracy of small buildings, and the recall rate is improved by 2.1 %. Ablation studies confirm the synergistic effect of these components, where the full model achieves the best indicators through staged optimization.

Compared with mainstream lightweight models (YOLOv5n, v8n), YOLO-RSTA achieves higher detection accuracy with moderate parameter growth (mAP50 is 1.6 % higher than YOLOv8n). Compared with variant models using other attention mechanisms or backbone replacement, this method also shows stronger generalization ability, especially in maintaining mAP50-95 stability.

Experimental results show that YOLO-RSTA can effectively improve the model's ability to recognize building features, especially in complex backgrounds and building detection of different scales. In the future, how to further improve the model's ability to detect small objects, it is expected that this method can provide valuable references for researchers in related fields and inspire more innovative research directions.

Authors' contributions. X. Wu proposed an object detection method based on an improved YOLOv10 network, and conducted experimental studies including training and testing the models on the LEVIR-CD dataset. S. V. Ablameyko participated in summarizing, analyzing, and presenting the obtained results, which demonstrate the superiority of the proposed method over the traditional YOLOv10 and other comparative approaches.

References

1. Li S. T., Li C. Y., Kang X. D. Current status and future prospects of multi-source remote sensing image fusion. *National Remote Sensing Bulletin*, 2021, vol. 25, no. 1, pp. 148–166. DOI: 10.11834/jrs.20210259.
2. Luo H. L., Wang W. X., Ye X. Y., Zhu S. X., Bai Y. Q. Research progress on directed object detection based on deep learning. *Image and Signal Processing*, 2024, vol. 13, no. 3, pp. 258–270. DOI: 10.12677/jisp.2024.133022.
3. Abdikan S., Bilgin G., Sanli F. B., Ushu E., Ustuner M. Enhancing land use classification with fusing dual-polarized terrasar-x and multispectral rapideye data. *Journal of Applied Remote Sensing*, 2015, vol. 9, no. 1, p. 096054. DOI: 10.1117/1.JRS.9.096054.
4. Liu F. F., Zhu C. M., Zhao N. N., Wu J. H. Remote sensing small target detection based on multimodal fusion. *Laser & Optoelectronics Progress*, 2024, vol. 61, no. 24, p. 2428010. DOI: 10.3788/LOP241203.
5. Li J., Wei X. M. Research on efficient detection network method for remote sensing images based on self-attention mechanism. *Image and Vision Computing*, 2024, vol. 142, p. 104884. DOI: 10.1016/j.imavis.2023.104884.
6. Liu D., Zhong L., Wu H., Li S., Li Y. Remote sensing image super-resolution reconstruction by fusing multi-scale receptive fields and hybrid transformer. *Scientific Reports*, 2025, vol. 15, p. 2140. DOI: 10.1038/s41598-025-86446-5.
7. Li Z., Wang H., Ma G., Yang W., Ablameyko S. Effective small object detection in remote sensing images based on improved YOLOv8 network. *Nonlinear Phenomena in Complex Systems*, 2024, vol. 27, no. 3, pp. 278–291. DOI: 10.5281/zenodo.13960639.
8. Tan M., Le Q. V. *EfficientNetV2: Smaller models and faster training*, 2021. Available at: <https://arxiv.org/abs/2104.00298> (accessed 13.02.2025). DOI: 10.48550/arXiv.2104.00298. (Preprint).
9. Woo S., Park J., Lee J. Y., Kweon I. S. CBAM: Convolutional Block Attention Module. *Proceedings of the European Conference on Computer Vision (ECCV), Munich, Germany, 8–14 September 2018*. Springer, Cham, 2018, pp. 3–19. DOI: 10.1007/978-3-030-01234-2_1.
10. Wang A., Chen H., Lin Z., Han J., Ding G. *RepViT: Revisiting mobile CNN from ViT perspective*, 2023. Available at: <https://arxiv.org/abs/2307.09283> (accessed 13.02.2025). DOI: 10.48550/arXiv.2307.09283. (Preprint).
11. Dosovitskiy A., Beyer L., Kolesnikov A., Weissenborn D., Zhai X., ..., Houlsby N. *An image is worth 16×16 words: Transformers for image recognition at scale*, 2020. Available at: <https://arxiv.org/abs/2010.11929> (accessed 13.02.2025). DOI: 10.48550/arXiv.2010.11929. (Preprint).
12. Wang X., Zhu D., Yan Y. Towards efficient detection for small objects via attention-guided detection network and data augmentation. *Sensors*, 2022, vol. 22, no. 19, p. 7663. DOI: 10.3390/s22197663.
13. Wang A., Chen H., Liu L. H., Chen K., Lin Z. J., ..., Ding G. G. *YOLOv10: Real-time end-to-end object detection*, 2024. Available at: <https://doi.org/10.48550/arXiv.2405.14458> (accessed 13.02.2025). DOI: 10.48550/arXiv.2405.14458. (Preprint).
14. Wang J., Xu C., Yang W., Yu L. *A normalized Gaussian Wasserstein distance for tiny object detection*, 2021. Available at: <https://doi.org/10.48550/arXiv.2110.13389> (accessed 13.02.2025). DOI: 10.48550/arXiv.2110.13389. (Preprint).
15. Huang H. B., Zhou X. Q., Cao J., He R., Tan T. N. *Vision transformer with super token sampling*, 2022. Available at: <https://doi.org/10.48550/arXiv.2211.11167> (accessed 13.02.2025). DOI: 10.48550/arXiv.2211.11167. (Preprint).
16. Wan D. H., Lu R., Tian S., Xu T., Lang X., Ren Z. Mixed local channel attention for object detection. *Engineering Applications of Artificial Intelligence*, 2023, vol. 123, p. 106442. DOI: 10.1016/j.engappai.2023.106442.

17. Wang H., Ablameyko S. Enhancing small object detection in remote sensing images using mixed local channel attention with YOLOv8. *Journal of Computer Technology and Applied Mathematics*, 2024, vol. 1, no. 1, pp. 40–45. DOI: 10.5281/zenodo.10986298.
18. Ding X., Zhang X., Ma N., Han J., Ding G., Sun J. RepVGG: Making VGG-style ConvNets great again. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Nashville, TN, USA, 20–25 June 2021*, pp. 13 733–13 742.
19. Vasa V. K., Zhu W., Chen X., Qiu P., Dong X., Wang Y. *STA-Unet: Rethink the semantic redundant for medical imaging segmentation*, 2024. Available at: <https://arxiv.org/pdf/2410.11578> (accessed 13.02.2025). DOI: 10.48550/arXiv.2410.11578. (Preprint).
20. Yu Z., Huang H., Chen W., Su Y., Liu Y., Wang X. *YOLO-FaceV2: A scale and occlusion aware face detector*, 2022. Available at: <https://arxiv.org/pdf/2208.02019v2> (accessed 13.02.2025). DOI: 10.48550/arXiv.2208.02019. (Preprint).
21. Chen H., Shi Z. A spatial-temporal attention-based method and a new dataset for remote sensing image change detection. *Remote Sensing*, 2020, vol. 12, no. 10, p. 1662. DOI: 10.3390/rs12101662.

Information about the authors

Xianyi Wu, Postgraduate Student of the Faculty of Mechanics and Mathematics of the Belarusian State University.
E-mail: tigerv5872@gmail.com
<https://orcid.org/0009-0003-6976-5386>

Sergey V. Ablameyko, Acad. of the National Academy of Sciences of Belarus, D. Sc. (Eng.), Prof. of the Faculty of Mechanics and Mathematics of the Belarusian State University.
E-mail: ablameyko@bsu.by
<https://orcid.org/0000-0001-9404-1206>

Информация об авторах

By Сянъи, аспирант механико-математического факультета Белорусского государственного университета.
E-mail: tigerv5872@gmail.com
<https://orcid.org/0009-0003-6976-5386>

Абламейко Сергей Владимирович, академик НАН Беларуси, доктор технических наук, профессор механико-математического факультета Белорусского государственного университета.
E-mail: ablameyko@bsu.by
<https://orcid.org/0009-0003-6976-5386>

МАТЕМАТИЧЕСКОЕ МОДЕЛИРОВАНИЕ MATHEMATICAL MODELING



УДК 519.872
DOI: 10.37661/1816-0301-2025-22-2-48-62

Оригинальная статья
Original Article

Анализ полуоткрытой сети массового обслуживания с адаптацией скоростей обслуживания к скоростям поступления запросов

А. Н. Дудин, О. С. Дудина, С. А. Дудин[✉]

Белорусский государственный университет,
пр. Независимости, 4, Минск, 220030, Беларусь
[✉]E-mail: dudins@bsu.by

Аннотация

Цели. Решается задача нахождения основных вероятностных характеристик функционирования полуоткрытой сети массового обслуживания, в которой интенсивность входного потока в узлы имеет несколько возможных уровней. При изменении уровня интенсивности потока возможно изменение скорости обслуживания запросов в узлах сети с целью оптимизации ее функционирования.

Методы. Для решения используется аппарат многомерных цепей Маркова с непрерывным временем и специальной блочной структурой инфинитезимального генератора.

Результаты. Получены алгоритмы и формулы для вычисления блоков генератора этой цепи, что позволяет вычислять ее инвариантное распределение вероятностей. Найдены формулы для вычисления основных характеристик производительности сети по известному инвариантному распределению вероятностей состояний цепи Маркова. Приведен численный пример, иллюстрирующий зависимость основных характеристик производительности сети от пропускной способности ее узлов. При выбранном экономическом критерии качества функционирования сети продемонстрирована возможность оптимизации перераспределения ресурсов сети между ее узлами при изменении уровня входного потока.

Заключение. Полученные результаты могут быть использованы для оптимизации функционирования различных реальных объектов, описываемых полуоткрытыми сетями массового обслуживания, например телекоммуникационных и логистических систем, мобильных роботизированных систем хранения, за счет адаптации распределения ресурсов сети между ее узлами к изменяющейся скорости поступления запросов.

Ключевые слова: полуоткрытая сеть массового обслуживания, изменение скорости поступления запросов, анализ производительности, цепь Маркова, вероятность отказа

Для цитирования. Дудин, А. Н. Анализ полуоткрытой сети массового обслуживания с адаптацией скоростей обслуживания к скоростям поступления запросов / А. Н. Дудин, О. С. Дудина, С. А. Дудин // Информатика. – 2025. – Т. 22, № 2. – С. 48–62. – DOI: 10.37661/1816-0301-2025-22-2-48-62.

Конфликт интересов. Авторы заявляют об отсутствии конфликта интересов.

Поступила в редакцию | Received 17.04.2025
Подписана в печать | Accepted 02.05.2025
Опубликована | Published 30.06.2025

Analysis of a semi-open queueing network with adaptation of service rates to the rates of arriving requests

Alexander N. Dudin, Olga S. Dudina, Sergei A. Dudin[✉]

Belarusian State University,
av. Nezavisimosti, 4, Minsk, 220030, Belarus
✉E-mail: dudins@bsu.by

Abstract

Objectives. The problem of computation of the main probabilistic characteristics of operation of a semi-open queueing network, in which the intensity of the input flow to the nodes has several possible levels, is considered. When the changing the level of flow intensity occurs, it is possible to change the rate of requests service in the nodes in order to optimize the network functioning.

Methods. The solution is based on the apparatus of multidimensional Markov chains with continuous time and a special block structure of the infinitesimal generator.

Results. The generator blocks of this chain are calculated using algorithms and formulas, which allows the invariant probability distribution to be determined. Formulas for calculating the main characteristics of network performance using the known invariant probability distribution of the states of the Markov chain are derived. A numerical example is provided to illustrate how the dependence of the main characteristics of network performance depend on the throughput of its nodes. Using the selected economic criterion for network performance quality, it is demonstrated that the redistribution of network resources between its nodes can be optimised with the change of the arrival flow level.

Conclusion. The obtained results can be used to optimize the functioning of various real objects described by semi-open queueing networks, for example, telecommunication and logistic systems, mobile robotic fulfillment systems, by adapting the distribution of network resources between its nodes to the changing rate of incoming requests.

Keywords: semi-open queueing network, request arrival rate variation, performance analysis, Markov chain, loss probability

For citation. Dudin A. N., Dudina O. S., Dudin S. A. *Analysis of a semi-open queueing network with adaptation of service rates to the rates of arriving requests*. Informatika [Informatics], 2025, vol. 22, no. 2, pp. 48–62 (In Russ.). DOI: 10.37661/1816-0301-2025-22-2-48-62.

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

Введение. Сеть массового обслуживания представляет собой совокупность конечного числа взаимосвязанных узлов, работа каждого из которых описывается системой массового обслуживания. По сети циркулируют запросы, которые после обслуживания в одном из узлов переходят в другой узел сети в соответствии с маршрутной матрицей или покидают сеть. Такие сети являются адекватными математическими моделями, например, телекоммуникационных, транспортных, производственных, логистических и других реальных систем. Поэтому существует очень обширная литература, посвященная их исследованию. В большинстве работ исследованы так называемые открытые сети, в которых неограниченное число запросов прибывают извне и после получения обслуживания покидают сеть навсегда, и замкнутые сети, в которых непрерывно циркулирует конечное фиксированное число запросов. Относительно недавно (в начале 2000-х гг.) начато исследование так называемых полуоткрытых сетей, в которых запросы так же, как и в открытых сетях, прибывают извне и уходят после получения обслуживания. Однако число запросов, которые могут одновременно присутствовать в системе, ограничено некоторой константой, она иногда называется уровнем мультипрограммирования. Популярность исследования полуоткрытых сетей объясняется их применимостью для анализа и оптимизации многих реальных систем, например для описания работы телекоммуникационных сетей, комплексов для проведения погрузочно-разгрузочных работ, гибких роботизированных производств и т. д. Для подробного ознакомления с тематикой исследования таких сетей можно рекомендовать, например, работы [1–8].

Во всех работах по исследованию полуоткрытых сетей массового обслуживания за исключением работы [8], в которой допускалось изменение параметров сети в зависимости от числа находящихся в ней запросов, предполагается, что параметры сети не изменяются в процессе ее функционирования. Однако хорошо известно, что данные параметры могут изменяться. В частности, может существенно изменяться интенсивность входного потока запросов в различные узлы сети (например, в зависимости от времени суток, погодных условий, возникновения чрезвычайных ситуаций и т. д.). Для обеспечения хорошего качества обслуживания запросов имеет смысл реагировать на изменения интенсивности потока соответствующим изменением интенсивности обслуживания запросов. Например, применительно к телекоммуникационным сетям можно запрашивать большую пропускную способность каналов или оборудования, предоставляемых для обеспечения устойчивой работы узлов. Если это невозможно, необходимо производить перераспределение имеющегося ресурса сети между ее узлами в зависимости от их загруженности.

В настоящей статье проанализирована модель полуоткрытой сети массового обслуживания со случайно меняющейся скоростью поступления запросов в узлы, которая может быть полезна для решения задачи оптимизации стратегии требования необходимого ресурса или перераспределения ресурса сети между ее узлами.

Отметим, что большинство существующих результатов в теории сетей массового обслуживания в силу сложности их анализа относятся к анализу tandemных сетей, имеющих линейную топологию; получению условий мультипликативности стационарного распределения сети, т. е. возможности получения стационарного распределения вероятностей состояний сети как произведения вероятностей состояний узлов сети, либо к получению различного рода приближенных решений.

В данной работе приведены точные алгоритмические результаты для нахождения стационарного распределения вероятностей состояний полуоткрытой сети с произвольной топологией и маршрутизацией запросов.

Математическая модель. Рассмотрим полуоткрытую сеть массового обслуживания, состоящую из K , $K \geq 2$, узлов (рис. 1).

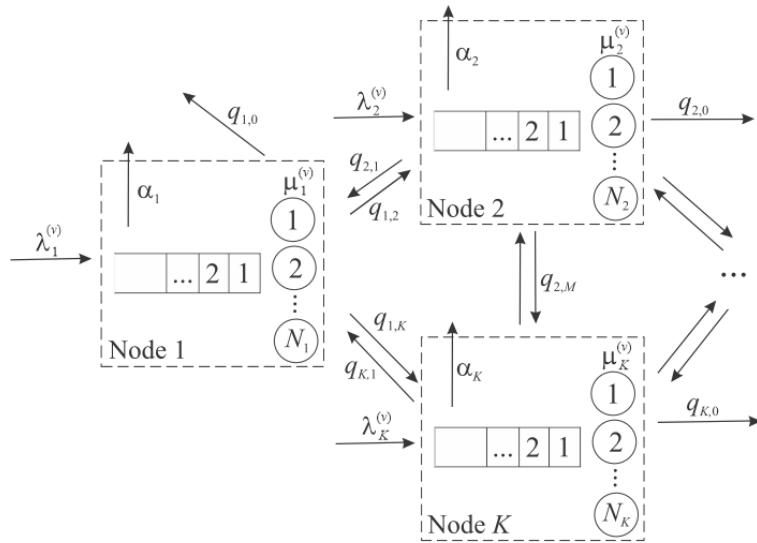


Рис. 1. Структура сети

Fig. 1. The structure of the network

Прибытие запросов в сеть определяется следующим образом. Предполагается, что имеется $V, V \geq 2$, уровней интенсивности входного потока. Переходы между уровнями регулируются цепью Маркова с непрерывным временем $v_t, t \geq 0$, и конечным пространством состояний

$\{1, 2, \dots, V\}$. Интенсивности переходов цепи v_t определяются неприводимым генератором Λ . Инвариантный вероятностный вектор v цепи v_t вычисляется как решение системы уравнений $v\Lambda = \mathbf{0}, v\mathbf{e} = 1$. Здесь $\mathbf{0}$ – это вектор-строка, состоящий из нулей, а \mathbf{e} – вектор-столбец, состоящий из единиц.

Предположим, что при уровне v процесса $v_t, t \geq 0$, в узел $k, k = \overline{1, K}$, поступает стационарный пуассоновский поток с интенсивностью $\lambda_k^{(v)}$. Общая интенсивность поступления в сеть при фиксированном уровне v входного процесса определяется как $\lambda^{(v)} = \sum_{k=1}^K \lambda_k^{(v)}$. Общая средняя интенсивность поступления запросов в сеть составляет $\lambda = \sum_{v=1}^V (v)_v \lambda^{(v)}$. Средняя интенсивность поступления в k -й узел $\lambda_k = \sum_{v=1}^V (v)_v \lambda_k^{(v)}, k = \overline{1, K}$.

В сети могут одновременно обрабатываться не более $N, N \geq 1$, запросов. Запрос, прибывающий, когда количество запросов в сети равно N , не принимается на обслуживание и считается потерянным. Работа каждого узла сети описывается многолинейной системой массового обслуживания с буфером емкости, гарантирующей невозможность его переполнения. Предполагается, что k -й узел имеет $N_k, k = \overline{1, K}$, идентичных приборов. Время обслуживания прибором в k -м узле при состоянии v процесса поступления запросов распределено экспоненциально с параметром $\mu_k^{(v)}$.

После получения обслуживания в k -м узле, $k = \overline{1, K}$, с вероятностью $q_{k,k'}$ запрос переходит на обслуживание в k' -й узел, $k' = \overline{1, M}$, или с вероятностью $q_{k,0}$ завершает обслуживание в сети, $\sum_{k'=0}^M q_{k,k'} = 1$.

Предполагается, что запросы, ожидающие в буферах, являются нетерпеливыми. Запрос, находящийся в k -м буфере, навсегда покидает сеть после времени, распределенного экспоненциально с параметром $\alpha_k, k = \overline{1, K}$, независимо от других ожидающих запросов.

Целями исследования являются анализ стационарного распределения состояний описанной сети массового обслуживания, вывод выражений для вычисления ее основных характеристик производительности, а также численная иллюстрация поведения этих характеристик.

Случайный процесс, описывающий динамику сети. Для описания поведения исследуемой сети может быть использована следующая цепь Маркова с непрерывным временем:

$$\Psi_t = \{n_t, v_t, n_t^{(1)}, \dots, n_t^{(K)}\}, t \geq 0,$$

где n_t – общее количество запросов в сети, $n_t = \overline{0, N}$;

v_t – уровень входного процесса, $v_t = \overline{1, V}$;

$n_t^{(k)}$ – количество запросов в k -м узле, $n_t^{(k)} = \overline{0, n_k}$, $\sum_{k=1}^K n_t^{(k)} = n_t, k = \overline{1, K}$,

в произвольный момент времени t .

Цепь Маркова $\Psi_t, t \geq 0$, является регулярной, неприводимой и имеет конечное пространство состояний. Поэтому стационарные (инвариантные) вероятности состояний цепи

$$\pi(n, v, n^{(1)}, \dots, n^{(K)}) = \lim_{t \rightarrow \infty} P\{n_t = n, v_t = v, n_t^{(1)} = n^{(1)}, \dots, n_t^{(K)} = n^{(K)}\},$$

$$n = \overline{0, N}, v = \overline{1, V}, n^{(k)} = \overline{0, n}, \sum_{k=1}^K n^{(k)} = n, k = \overline{1, K},$$

существуют для любых значений параметров рассматриваемой сети массового обслуживания.

Пусть $\boldsymbol{\pi}_n = (\pi(n, 1), \pi(n, 2), \dots, \pi(n, V))$, $n = \overline{0, N}$, есть векторы-строки, состоящие из вероятностей состояний со значением n первой компоненты, перенумерованных в прямом лексикографическом порядке значений компоненты v_t и в обратном лексикографическом порядке значений компоненты $n_t^{(1)}, \dots, n_t^{(K)}$.

Известно, что векторы $\boldsymbol{\pi}_n$, $n = \overline{0, N}$, удовлетворяют системе линейных алгебраических уравнений

$$(\boldsymbol{\pi}_0, \boldsymbol{\pi}_1, \dots, \boldsymbol{\pi}_N)G = \mathbf{0}, \quad (\boldsymbol{\pi}_0, \boldsymbol{\pi}_1, \dots, \boldsymbol{\pi}_N)\mathbf{e} = 1, \quad (1)$$

где G – инфинитезимальный генератор цепи Маркова ψ_t , $t \geq 0$. Для решения системы (1) нужно определить генератор G .

Теорема. Инфинитезимальный генератор G цепи Маркова ψ_t , $t \geq 0$, имеет блочную трехдиагональную структуру $G = (G_{n,n'})$, $|n - n'| \leq 1$, $n, n' = \overline{0, N}$, с ненулевыми блоками:

$$\begin{aligned} G_{0,0} &= \Lambda - \text{diag}\{\lambda^{(v)}, v = \overline{1, V}\}, \\ G_{n,n} &= \Lambda \otimes I_{J_n} + \text{diag}\{-\lambda^{(v)} I_{J_n} + (\Psi_n^{(v)}(Q^{(v)}) + \Delta_n^{(v)}), v = \overline{1, V}\}, n = \overline{1, N-1}, \\ G_{N,N} &= \Lambda \otimes I_{J_N} + \text{diag}\{(\Psi_N^{(v)}(Q^{(v)}) + \Delta_N^{(v)}), v = \overline{1, V}\}, \\ G_{n,n-1} &= \text{diag}\{S_n^{(v)}(\mathbf{q}^{(v)}) + A_n(\boldsymbol{\alpha}), v = \overline{1, V}\}, n = \overline{1, N}, \\ G_{n,n+1} &= \text{diag}\left\{\sum_{k=1}^K \lambda_k^{(v)} P_n(\boldsymbol{\beta}_k), v = \overline{1, V}\right\}, n = \overline{0, N-1}. \end{aligned}$$

Здесь и далее:

- $\text{diag}\{A_v, v = \overline{1, V}\}$ – диагональная матрица с диагональными элементами, указанными в скобках;
- $\text{diag}^+\{A_v, v = \overline{1, V}\}$ – наддиагональная матрица с наддиагональными элементами, указанными в скобках;
- $\text{diag}^-\{A_v, v = \overline{1, V}\}$ – поддиагональная матрица с поддиагональными элементами, указанными в скобках;
- \otimes – символ Кронекерова произведения матриц [9];
- $J_n = C_{n+K-1}^{K-1} = \frac{(n+K-1)!}{n!(K-1)!}$;
- $\boldsymbol{\alpha}$ – вектор-столбец, определенный как

$$\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \dots, \alpha_K)^T;$$

- $\boldsymbol{\beta}_k$, $k = \overline{1, K}$, – векторы размера K со всеми нулевыми элементами, кроме k -го элемента, который равен единице;
- $P_n(\boldsymbol{\beta}_k)$, $n = \overline{0, N-1}$, $k = \overline{1, K}$, – матрица, задающая вероятности переходов процесса $\mathbf{n}_t = \{n_t^{(1)}, \dots, n_t^{(K)}\}$, $t \geq 0$, в момент принятия запроса в k -й узел сети, при условии, что в сети

находится n запросов. Данные вероятности не зависят от состояния процесса поступления. Таким образом, для вычисления $P_n(\beta_k)$ можно применить алгоритм из работы [6, лемма 4];

- $A_n(\alpha), n = \overline{1, N}$, – матрицы, содержащие интенсивности переходов процесса $\mathbf{n}_t, t \geq 0$, которые вызваны потерей запроса из-за нетерпеливости в случае нахождения n запросов в сети. Они определяются как $A_n(\alpha) = A_n^{(K-1)}(\alpha)$, где матрицы $A_n^{(K-1)}(\alpha)$ вычисляются из рекурсий:

$$A_n^{(0)}(\alpha) = \max\{0, n - N_K\} \alpha_K,$$

$$A_n^{(k)}(\alpha) = \begin{pmatrix} \max\{0, n - N_{K-k}\} \alpha_{K-k} I & O & \cdots & O \\ A_1^{(k-1)}(\alpha) & \max\{0, n - 1 - N_{K-k}\} \alpha_{K-k} I & \cdots & O \\ O & A_2^{(k-1)}(\alpha) & \cdots & O \\ \vdots & \vdots & \ddots & \vdots \\ O & O & \cdots & \max\{0, 1 - N_{K-k}\} \alpha_{K-k} I \\ O & O & \cdots & A_n^{(k-1)}(\alpha) \end{pmatrix},$$

$$k = \overline{1, K-1}, n = \overline{2, N};$$

- $\mathbf{q}^{(v)}, v = \overline{1, V}$, – вектор-столбец, заданный как

$$\mathbf{q}^{(v)} = (q_1^{(v)}, q_2^{(v)}, \dots, q_K^{(v)})^T = (q_{1,0} \mu_1^{(v)}, q_{2,0} \mu_2^{(v)}, \dots, q_{K,0} \mu_K^{(v)})^T;$$

- $S_n^{(v)}(\mathbf{q}^{(v)}), n = \overline{1, N}, v = \overline{1, V}$, – матрицы, определяющие интенсивности переходов процесса $\mathbf{n}_t, t \geq 0$, вызванных завершением обслуживания в сети запроса в случае наличия n запросов в сети при уровне v процесса поступления. Они вычисляются следующим образом:

$$S_n^{(v)}(\mathbf{q}^{(v)}) = S_n^{(K-1,v)}(\mathbf{q}^{(v)}),$$

где матрицы $S_n^{(K-1,v)}(\mathbf{q}^{(v)})$ находятся с использованием рекурсий

$$S_n^{(0,v)}(\mathbf{q}^{(v)}) = \min\{n, N_K\} q_K^{(v)},$$

$$S_n^{(k,v)}(\mathbf{q}^{(v)}) = \begin{pmatrix} \min\{n, N_{K-k}\} q_{K-k}^{(v)} I & O & \cdots & O \\ S_1^{(k-1,v)}(\mathbf{q}^{(v)}) & \min\{n - 1, N_{K-k}\} q_{K-k}^{(v)} I & \cdots & O \\ O & S_2^{(k-1,v)}(\mathbf{q}^{(v)}) & \cdots & O \\ \vdots & \vdots & \ddots & \vdots \\ O & O & \cdots & \min\{1, N_{K-k}\} q_{K-k}^{(v)} I \\ O & O & \cdots & S_n^{(k-1,v)}(\mathbf{q}^{(v)}) \end{pmatrix},$$

$$k = \overline{1, K-1}, n = \overline{1, N};$$

- $Q^{(v)}, v = \overline{1, V}$, – квадратные матрицы размера K со следующими элементами:

$$Q_{k,k}^{(v)} = -\mu_k^{(v)}, k = \overline{1, K}, \quad Q_{k,k'}^{(v)} = q_{k,k'} \mu_k^{(v)}, k, k' = \overline{1, K}, k' \neq k;$$

- $Q_j^{(v)}, j = \overline{1, K-2}, v = \overline{1, V}$, – матрицы, полученные путем удаления первых $K-2-j$ строк и столбцов из матрицы $Q^{(v)}$;

- $\Psi_n^{(v)}(Q^{(v)})$, $n = \overline{1, N}$, $v = \overline{1, V}$, – матрицы, задающие интенсивности переходов запросов между узлами сети при n запросах в ней и уровне v и определяемые как

$$\Psi_n^{(v)}(Q^{(v)}) = \Psi_n^{(K-2,v)},$$

где матрицы $\Psi_n^{(k,v)} = \Psi_n^{(k,v)}(Q^{(v)})$ вычисляются с использованием рекурсивных формул

$$\Psi_n^{(0,v)} = \text{diag}^+ \{ \min\{i, N_{K-1}\} Q_{K-1,K}^{(v)}, i = \overline{n, 1} \} +$$

$$+ \text{diag}^- \{ \min\{i, N_K\} Q_{K,K-1}^{(v)}, i = \overline{1, n}, n = \overline{1, N},$$

$$\Psi_n^{(k,v)} = \text{diag}\{O, \Psi_i^{(k-1,v)}, i = \overline{1, n}\} + \text{diag}^+ \{ \min\{n-i, N_{K-1-k}\} \Gamma_{i,m}^{(v)}, i = \overline{0, n-1} \} +$$

$$+ \text{diag}^- \{ \Omega_{i,k}^{(v)}, i = \overline{1, n}, n = \overline{1, N}, k = \overline{1, K-2} \}.$$

Здесь матрицы $\Omega_{n,j}^{(v)} = \Omega_{n,j}^{(v_j-2,v)}$, $n = \overline{1, N}$, и $\Gamma_{n,j}^{(v)} = \Gamma_{n,j}^{(v_j-2,v)}$, $n = \overline{0, N-1}$, $j = \overline{1, K-2}$, вычисляются следующим образом:

$$\Omega_{n,j}^{(0,v)} = \min\{n, N_K\} q_{v_j,1}^{(j,v)}, n = \overline{1, N}, j = \overline{1, K-2},$$

$$\Omega_{n,j}^{(k,v)} = \begin{pmatrix} \min\{n, N_{K-k}\} q_{v_j-k,1}^{(j,v)} I & O & \cdots & O \\ \Omega_{1,j}^{(k-1,v)} & \min\{n-1, N_{K-k}\} q_{v_j-k,1}^{(j,v)} I & \cdots & O \\ O & \Omega_{2,j}^{(k-1,v)} & \cdots & O \\ \vdots & \vdots & \ddots & \vdots \\ O & O & \cdots & \min\{1, N_{K-k}\} q_{v_j-k,1}^{(j,v)} I \\ O & O & \cdots & \Omega_{n,j}^{(k-1,v)} \end{pmatrix},$$

$$k = \overline{1, v_j - 2}, n = \overline{1, N}, j = \overline{1, K-2},$$

$$\Gamma_{n,j}^{(0,v)} = q_{1,v_j}^{(j,v)}, n = \overline{0, N-1}, j = \overline{1, K-2},$$

$$\Gamma_{n,j}^{(k,v)} = \begin{pmatrix} q_{1,v_j-k}^{(j,v)} I & \Gamma_{0,j}^{(k-1,v)} & O & \cdots & O & O \\ O & q_{1,v_j-k}^{(j,v)} I & \Gamma_{1,j}^{(k-1,v)} & \cdots & O & O \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ O & O & O & \cdots & q_{1,v_j-k}^{(j,v)} I & \Gamma_{n,j}^{(k-1,v)} \end{pmatrix},$$

$$k = \overline{1, v_j - 2}, n = \overline{0, N-1}, j = \overline{1, K-2},$$

где $q_{k,k'}^{(j,v)} = (k, k')$ -й элемент матрицы $Q_j^{(v)}$, а v_j – количество строк матрицы $Q_j^{(v)}$;

- $\Delta_n^{(v)}$, $n = \overline{1, N}$, $v = \overline{1, V}$, – диагональные матрицы, диагональные элементы которых с точностью до знака определяют суммарные интенсивности выхода из соответствующего состояния в случае n запросов в сети при уровне v процесса прибытия. Очевидно, что эти матрицы можно найти с помощью выражения

$$\Delta_n^{(v)} = -\text{diag}\{\Psi_n^{(v)}(Q^{(v)})e + S_n^{(v)}(q^{(v)})e + A_n(\alpha)e\}, n = \overline{1, N}.$$

Доказательство осуществляется путем анализа всех возможных переходов цепи Маркова $\psi_t, t \geq 0$, с последующей записью соответствующих интенсивностей в виде блоков генератора. Основная трудность здесь – определение матриц, задающих интенсивности переходов процесса $n_t, t \geq 0$. Алгоритмы для вычисления данных матриц основаны на результатах работы [6]. В этой работе интенсивности в узлах были фиксированы и не могли изменяться, что существенно облегчало исследование.

Определив вид генератора G , можно подсчитать векторы $\pi_n, n = \overline{0, N}$, стационарных вероятностей состояний цепи Маркова $\psi_t, t \geq 0$, как решение конечной системы линейных алгебраических уравнений (1). Количество уравнений этой системы может быть большим. Для ее решения следует использовать алгоритмы, учитывающие блочную трехдиагональную структуру генератора, например алгоритм, предложенный в работе [10].

Формулы для вычисления характеристик производительности. Вычислив векторы $\pi_n, n = \overline{0, N}$, можно подсчитать значения различных характеристик производительности сети и ее узлов.

Интенсивность выходного потока успешно обслуженных в сети запросов из k -го узла вычисляется как

$$\lambda_k^{succ} = \sum_{n=1}^N \sum_{v=1}^V \pi(n, v) (S_n^{(v)}(\mathbf{q}^{(v,k)})) e, k = \overline{1, K},$$

где $\mathbf{q}^{(v,k)}$ – вектор-столбец размером K со всеми нулевыми элементами, кроме k -го элемента $(\mathbf{q}^{(v,k)})_k$, который равен $q_k^{(v)}$.

Среднее количество запросов, находящихся в k -м узле, вычисляется как

$$L_k = \sum_{n=1}^N \sum_{v=1}^V \pi(n, v) \Phi_n(\beta_k) e, k = \overline{1, K},$$

где матрицы $\Phi_n(\beta) = \Phi_n^{(K-1)}(\beta), \beta \in \{\beta_1, \beta_2, \dots, \beta_K\}$, определяются с использованием рекурсии

$$\Phi_n^{(0)}(\beta) = n\beta_K, \\ \Phi_n^{(k)}(\beta) = \begin{pmatrix} n\beta_{K-k}I & O & \cdots & O \\ \Phi_1^{(k-1)}(\beta) & (n-1)\beta_{K-k}I & \cdots & O \\ O & \Phi_2^{(k-1)}(\beta) & \cdots & O \\ \vdots & \vdots & \ddots & \vdots \\ O & O & \cdots & \beta_{K-k}I \\ O & O & \cdots & \Phi_n^{(k-1)}(\beta) \end{pmatrix}, k = \overline{1, K-1}, n = \overline{1, N}.$$

Кроме того, вычисляются:

– среднее количество занятых приборов в k -м узле, $k = \overline{1, K}$,

$$N_k^{\text{serv}} = \sum_{n=1}^N \sum_{v=1}^V \pi(n, v) S_n^{(v)}(\beta_k) e;$$

– среднее количество запросов в буфере k -го узла, $k = \overline{1, K}$,

$$L_k^{\text{buf}} = \sum_{n=1}^N \sum_{v=1}^V \pi(n, v) A_n(\beta_k) e = L_k - N_k^{\text{serv}};$$

– вероятность потери произвольного запроса при прибытии в k -й узел из-за наличия N запросов в сети

$$P_k^{\text{arr-loss}} = \lambda^{-1} \sum_{v=1}^V \lambda_k^{(v)} \boldsymbol{\pi}(N, v) \mathbf{e}, k = \overline{1, K};$$

– вероятность потери произвольного запроса из-за нетерпеливости в k -м узле

$$P_k^{\text{imp-loss}} = \lambda^{-1} \sum_{n=1}^N \sum_{v=1}^V \boldsymbol{\pi}(n, v) A_n(\boldsymbol{\alpha}^{(k)}) \mathbf{e}, k = \overline{1, K},$$

где $\boldsymbol{\alpha}^{(k)}$ – векторы-столбцы размером K со всеми нулевыми элементами, кроме k -го элемента $(\boldsymbol{\alpha}^{(k)})_k$, который равен α_k ;

– вероятность потери произвольного запроса в k -м узле

$$P_k^{\text{loss}} = P_k^{\text{arr-loss}} + P_k^{\text{imp-loss}}, k = \overline{1, K};$$

– среднее количество запросов в сети

$$L = \sum_{n=1}^N n \boldsymbol{\pi}_n \mathbf{e};$$

– интенсивность выходного потока успешно обслуженных запросов в сети

$$\lambda^{\text{succ}} = \sum_{n=1}^N \sum_{v=1}^V \boldsymbol{\pi}(n, v) S_n^{(v)} (\boldsymbol{q}^{(v)}) \mathbf{e} = \sum_{k=1}^K \lambda_k^{\text{succ}};$$

– среднее количество занятых приборов в сети

$$N^{\text{serv}} = \sum_{k=1}^M N_k^{\text{serv}};$$

– среднее количество запросов во всех буферах сети

$$L^{\text{buf}} = \sum_{k=1}^M L_k^{\text{buf}};$$

– вероятность потери произвольного запроса по прибытии из-за наличия N запросов в сети

$$P^{\text{arr-loss}} = \lambda^{-1} \sum_{v=1}^V \lambda^{(v)} \boldsymbol{\pi}(N, v) \mathbf{e};$$

– вероятность потери произвольного запроса из-за нетерпеливости

$$P^{\text{imp-loss}} = \lambda^{-1} \sum_{n=1}^N \sum_{v=1}^V \boldsymbol{\pi}(n, v) A_n(\boldsymbol{\alpha}) \mathbf{e};$$

– вероятность потери произвольного запроса

$$P^{\text{loss}} = P^{\text{arr-loss}} + P^{\text{imp-loss}} = 1 - \frac{\lambda^{\text{succ}}}{\lambda};$$

– вероятность успешного обслуживания произвольного запроса в сети

$$P^{\text{succ}} = 1 - P^{\text{loss}} = \frac{\lambda^{\text{succ}}}{\lambda}.$$

Численный эксперимент. Приведем численный пример, в котором предполагается, что сеть состоит из $K=3$ узлов и до $N=25$ запросов могут находиться в сети одновременно. Количество приборов в узлах составляет $N_1=5, N_2=4$ и $N_3=3$.

Интенсивность нетерпеливости в буферах узлов задана как $\alpha_1=0,015, \alpha_2=0,008, \alpha_3=0,015$.

Процесс прибытия имеет $V=2$ уровня, переходы между которыми определяются генератором $\Lambda = \begin{pmatrix} -0,002 & 0,002 \\ 0,005 & -0,005 \end{pmatrix}$.

Если процесс прибытия находится на уровне 1, интенсивности прибытия в узлы равны $\lambda_1^{(1)}=1, \lambda_2^{(1)}=2, \lambda_3^{(1)}=0,5$ соответственно. Интенсивности прибытия в узлы сети при уровне 2 процесса прибытия равны $\lambda_1^{(2)}=1,2, \lambda_2^{(2)}=1, \lambda_3^{(2)}=2,5$ соответственно.

Вероятности переходов запросов после завершения обслуживания в соответствующих узлах сети задаются выражениями

$$p_{1,2}=0,2, p_{1,3}=0,3, p_{2,1}=0,4, p_{2,3}=0,1, p_{3,1}=0,2, p_{3,2}=0,1.$$

Предположим, что общая пропускная способность, предоставляемая для передачи информации в сети, составляет $\tilde{B}=120$ Мбит/с, а запросы идентичны и имеют средний размер $\tilde{v}=8$ Мбит. При уровне 1 процесса прибытия пропускная способность делится между узлами следующим образом: пропускная способность узла 1 составляет $B_1^{(1)}=45$ Мбит/с, узла 2 – $B_2^{(1)}=60$ Мбит/с, а узла 3 – $B_3^{(1)}=15$ Мбит/с. Таким образом, интенсивности обслуживания приборами в узлах сети при уровне 1 входного процесса определяются как

$$\mu_k^{(1)} = \frac{B_k^{(1)}}{\frac{N_k}{\tilde{v}}}, k=1,2,3, \text{ и равны } \mu_1^{(1)}=9/8, \mu_2^{(1)}=15/8, \mu_3^{(1)}=5/8.$$

Целью эксперимента является определение того, как пропускная способность должна оптимально перераспределяться между узлами, когда процесс прибытия переходит на уровень 2. Предполагается, что минимальная пропускная способность, предоставляемая каждому узлу, составляет 5 Мбит/с. Для нахождения оптимального перераспределения пропускной способности при переходе на уровень 2 входного процесса будем изменять пропускную способность узла 1 $B_1^{(2)}$ на интервале $[5, \tilde{B}-10]$ и пропускную способность узла 2 $B_2^{(2)}$ на интервале $[5, \tilde{B}-5-B_1^{(2)}]$ с шагом 5 Мбит/с и исследовать зависимость основных показателей производительности сети от этих параметров. Отметим, что при фиксированной пропускной способности узлов 1 и 2 пропускная способность узла 3 рассчитывается как $B_3^{(2)}=\tilde{B}-B_1^{(2)}-B_2^{(2)}$. Интенсивности обслуживания при уровне 2 входного процесса рассчитываются таким же

образом, как и при уровне 1 входного процесса, и равны $\mu_k^{(2)} = \frac{B_k^{(2)}}{\frac{N_k}{\tilde{v}}}, k=1,2,3$.

На рис. 2 показана зависимость среднего количества L запросов в сети от параметров $B_1^{(2)}$ и $B_2^{(2)}$.

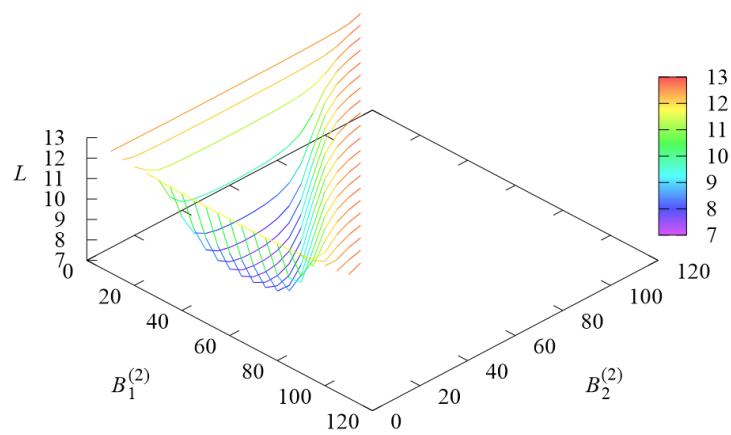


Рис. 2. Зависимость L от $B_1^{(2)}$ и $B_2^{(2)}$
Fig. 2. The dependence of L on $B_1^{(2)}$ and $B_2^{(2)}$

Рис. 3–5 иллюстрируют зависимость вероятностей потерь P_k^{loss} , $k = 1, 2, 3$, произвольного запроса в k -м узле от параметров $B_1^{(2)}$ и $B_2^{(2)}$.

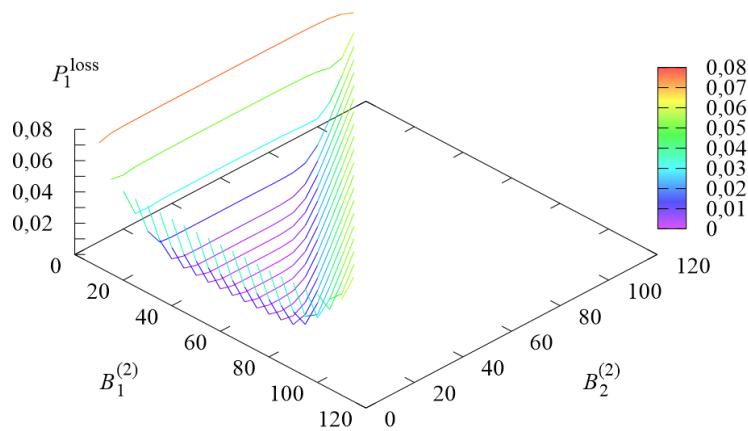


Рис. 3. Зависимость P_1^{loss} от $B_1^{(2)}$ и $B_2^{(2)}$
Fig. 3. The dependence of P_1^{loss} on $B_1^{(2)}$ and $B_2^{(2)}$

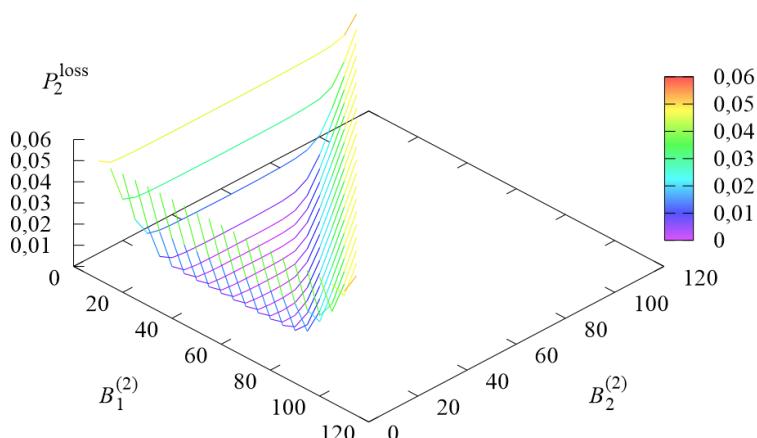


Рис. 4. Зависимость P_2^{loss} от $B_1^{(2)}$ и $B_2^{(2)}$
Fig. 4. The dependence of P_2^{loss} on $B_1^{(2)}$ and $B_2^{(2)}$

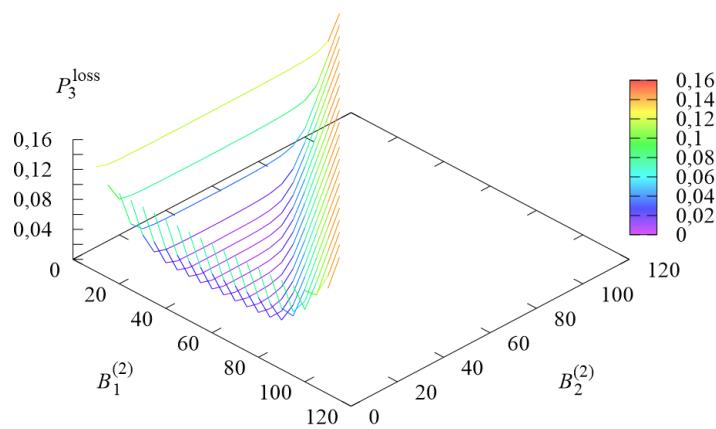


Рис. 5. Зависимость P_3^{loss} от $B_1^{(2)}$ и $B_2^{(2)}$
Fig. 5. The dependence of P_3^{loss} on $B_1^{(2)}$ and $B_2^{(2)}$

На рис. 6–8 изображена зависимость вероятности потери $P^{\text{arr-loss}}$ произвольного запроса по прибытии, вероятности потери $P^{\text{imp-loss}}$ произвольного запроса из-за нетерпеливости и вероятности потери P^{loss} произвольного запроса от параметров $B_1^{(2)}$ и $B_2^{(2)}$.

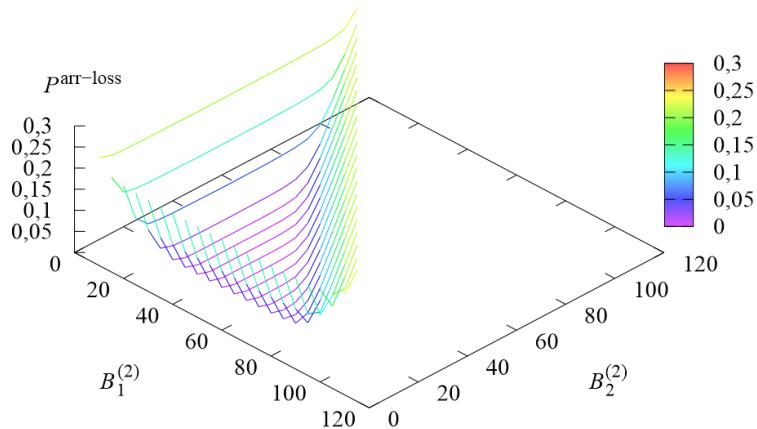


Рис. 6. Зависимость $P^{\text{arr-loss}}$ от $B_1^{(2)}$ и $B_2^{(2)}$
Fig. 6. The dependence of $P^{\text{arr-loss}}$ on $B_1^{(2)}$ and $B_2^{(2)}$

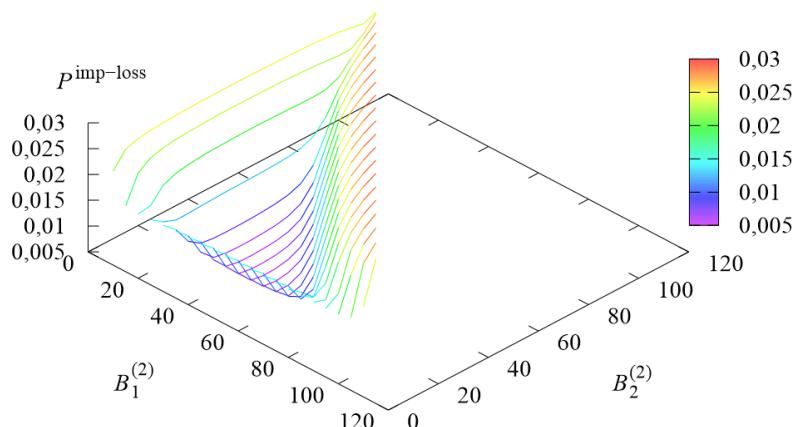


Рис. 7. Зависимость $P^{\text{imp-loss}}$ от $B_1^{(2)}$ и $B_2^{(2)}$
Fig. 7. The dependence of $P^{\text{imp-loss}}$ on $B_1^{(2)}$ and $B_2^{(2)}$

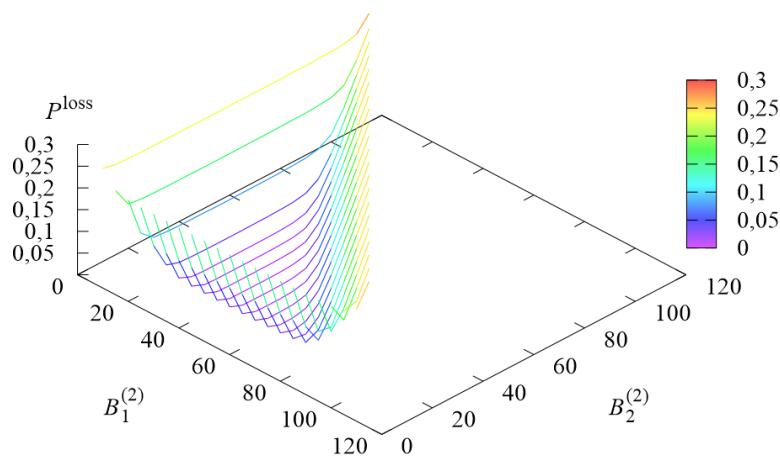


Рис. 8. Зависимость P^{loss} от $B_1^{(2)}$ и $B_2^{(2)}$

Fig. 8. The dependence of P^{loss} on $B_1^{(2)}$ and $B_2^{(2)}$

Как видно из приведенных рисунков, основные характеристики системы существенно зависят от параметров $B_1^{(2)}$ и $B_2^{(2)}$. При слишком больших или слишком маленьких значениях пропускной способности в узлах число запросов в сети, а также вероятности потери запросов принимают большие значения. Поэтому очевидно, что задача перераспределения ресурсов сети при изменении интенсивностей входного потока является важной. Для определения оптимальных значений параметров $B_1^{(2)}$ и $B_2^{(2)}$ необходимо выбрать функционал, описывающий качество функционирования системы.

Предположим, что качество работы сети определяется следующим стоимостным критерием:

$$E = E(B_1^{(2)}, B_2^{(2)}) = a_1 \lambda^{\text{succ}} - b_1 \lambda P^{\text{arr-loss}} - b_2 \lambda P^{\text{imp-loss}},$$

где a_1 – прибыль, получаемая сетью от обслуживания одного запроса; b_1 – штраф за потерю произвольного запроса при поступлении, а b_2 – штраф за потерю произвольного запроса из-за нетерпеливости. Таким образом, стоимостный критерий E определяет прибыль системы в единицу времени. В данном эксперименте зафиксируем следующие значения стоимостных коэффициентов: $a_1 = 2$, $b_1 = 3$, $b_2 = 10$.

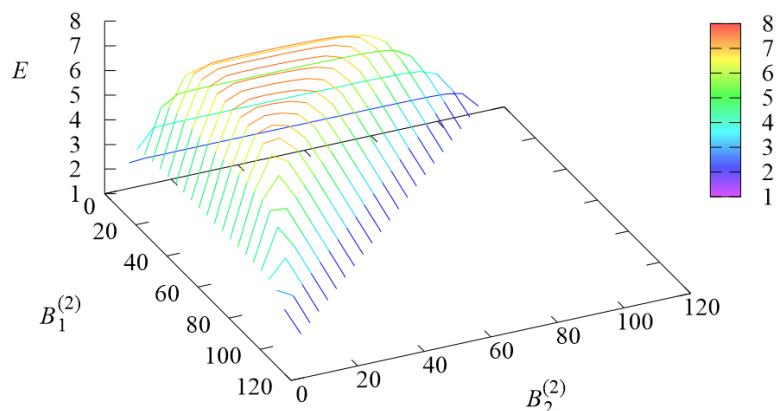


Рис. 9. Зависимость E от $B_1^{(2)}$ и $B_2^{(2)}$

Fig. 9. The dependence of E on $B_1^{(2)}$ and $B_2^{(2)}$

На рис. 9 показана зависимость значений критерия E от параметров $B_1^{(2)}$ и $B_2^{(2)}$. Видно, что при выборе слишком малых и слишком больших значений $B_1^{(2)}$ и $B_2^{(2)}$ прибыль сети за обслуживание запросов существенно уменьшается. Оптимальное значение критерия E равно $E^* = 7,42844$ и достигается при $B_1^{(2)} = 40, B_2^{(2)} = 30$ и, соответственно, $B_3^{(2)} = \tilde{B} - B_1^{(2)} - B_2^{(2)} = 50$. Существенное перераспределение пропускной способности сети в пользу третьего узла легко объясняется пятикратным увеличением интенсивности потока, поступающего в этот узел, при переключении входного потока с уровня 1 на уровень 2. Отметим, что если перераспределение пропускной способности не осуществлять, т. е. оставить значения параметров $B_1^{(2)}$ и $B_2^{(2)}$ равными соответственно $B_1^{(1)} = 45$ и $B_2^{(1)} = 60$, то значение критерия E будет равняться 4,0833, что существенно ниже оптимального значения. Также отметим, что при фиксированных оптимальных значениях $B_1^{(2)} = 40, B_2^{(2)} = 30$ вероятность потери произвольного запроса P^{loss} равна 0,00575, в то время как при отсутствии перераспределения пропускной способности вероятность P^{loss} составляет 0,15118, что более чем в 26 раз больше. Таким образом, перераспределение пропускной способности между узлами имеет смысл.

Вычисления, результаты которых приведены выше, проводились на компьютере с процессором Intel Core i7-8700 3,2 ГГц, 16 ГБ RAM с использованием Wolfram Mathematica 13.2. Время вычисления составило 923 с для расчета 253 точек при максимальном размере блоков генератора, равном 702. В среднем на одну точку тратилось 3,65 с.

Заключение. В статье рассмотрена задача расчета характеристик полуоткрытой сети массового обслуживания, в которой входной поток в узлы имеет несколько возможных уровней, переключение между ними управляет цепью Маркова с конечным пространством состояний. При изменении уровня потока есть возможность изменять и скорости обслуживания запросов в узлах сети. При фиксированных значениях параметров сети ее поведение описано многомерной цепью Маркова. Получены формулы для вычисления генератора этой цепи, имеющего блочную трехдиагональную структуру, и формулы для вычисления основных характеристик производительности сети через векторы стационарных вероятностей состояний цепи Маркова. Приведен численный пример, иллюстрирующий реализуемость предложенных алгоритмов для расчета, а также поведение некоторых характеристик производительности сети. Показана возможность решения задачи оптимизации перераспределения ресурсов сети при изменении уровня входного потока.

В дальнейшем планируется обобщение полученных результатов на случаи более сложного поведения процесса поступления запросов при каждом уровне, рассмотрение зависимости от уровня потока не только скорости обслуживания в узлах сети, но и числа приборов в узлах.

Вклад авторов. А. Н. Дудин – построение модели, аналитическое исследование, подготовка текста статьи. О. С. Дудина и С. А. Дудин – аналитическое исследование, написание текста, программирование алгоритмов и тестирование программ, проведение и оформление численных экспериментов.

References

- Roy D. Semi-open queueing networks: a review of stochastic models, solution methods and new research areas. *International Journal of Production Research*, 2016, vol. 54, no. 6, pp. 1735–1752.
- Otten S., Krenzler R., Xie L., Daduna H., Kruse K. Analysis of semi-open queueing networks using lost customers approximation with an application to robotic mobile fulfilment systems. *OR Spectrum*, 2022, vol. 44, no. 2, pp. 603–648.
- Jia J., Heragu S. S. Solving semi-open queueing networks. *Operations Research*, 2009, vol. 57, no. 2, pp. 391–401.
- Kim J., Dudin A., Dudin S., Kim C. Analysis of a semi-open queueing network with Markovian arrival process. *Performance Evaluation*, 2018, vol. 120, pp. 1–19.

5. Kim C., Dudin S., Dudin A., Samouylov K. Analysis of a semi-open queueing network with a state dependent marked Markovian arrival process, customers retrials and impatience. *Mathematics*, 2019, vol. 7, no. 8, p. 715.
6. Dudin S., Dudin A., Manzo R., Rarità L. Analysis of semi-open queueing network with correlated arrival process and multi-server nodes. *Operations Research Forum*. Cham, Springer International Publishing, 2024, vol. 5, no. 4, p. 99.
7. Amjath M., Kerbache L., Elomri A., MacGregor Smith J. Queueing network models for the analysis and optimisation of material handling systems: a systematic literature review. *Flexible Services and Manufacturing Journal*, 2024, vol. 36, no. 2, pp. 668–709.
8. D'Apice C., Dudin A., Dudin S., Manzo R. Study of a semi-open queueing network with hysteresis control of service regimes. *AIMS Mathematics*, 2025, vol. 10, no. 2, pp. 3095–3123.
9. Graham A. *Kronecker Products and Matrix Calculus with Applications*. Courier Dover Publications, 2018, 144 p.
10. Baumann H., Sandmann W. Numerical solution of level dependent quasi-birth-and-death processes. *Procedia Computer Science*, 2010, vol. 1, no. 1, pp. 1561–1569.

Информация об авторах

Дудин Александр Николаевич, доктор физико-математических наук, профессор, заведующий НИЛ прикладного вероятностного анализа, Белорусский государственный университет.

E-mail: dudin@bsu.by

Дудина Ольга Сергеевна, кандидат физико-математических наук, ведущий научный сотрудник НИЛ прикладного вероятностного анализа, Белорусский государственный университет.

E-mail: dudina@bsu.by

Дудин Сергей Александрович, кандидат физико-математических наук, ведущий научный сотрудник НИЛ прикладного вероятностного анализа, Белорусский государственный университет.

E-mail: dudins@bsu.by

Information about the authors

Alexander N. Dudin, D. Sc. (Phys.-Math.), Prof., Head of the Research Laboratory of Applied Probability Analysis, Belarusian State University.
E-mail: dudin@bsu.by

Olga S. Dudina, Ph. D. (Phys.-Math.), Leading Researcher of the Research Laboratory of Applied Probability Analysis, Belarusian State University.
E-mail: dudina@bsu.by

Sergei A. Dudin, Ph. D. (Phys.-Math.), Leading Researcher of the Research Laboratory of Applied Probability Analysis, Belarusian State University.
E-mail: dudins@bsu.by

ЛОГИЧЕСКОЕ ПРОЕКТИРОВАНИЕ LOGICAL DESIGN



УДК 004.33.054
DOI: 10.37661/1816-0301-2025-22-2-63-80

Оригинальная статья
Original Article

Масштабирование управляемых вероятностных тестов с применением матриц Адамара

В. Н. Ярмолик^{1✉}, Н. А. Шевченко², В. В. Петровская¹

¹Белорусский государственный университет
информатики и радиоэлектроники,
ул. П. Бровки, 6, Минск, 220013, Беларусь
[✉]E-mail: yarmolik10ru@yahoo.com

²Дармштадтский технический университет,
Каролиненплац, 5, Дармштадт, 64289, Германия

Аннотация

Цели. Решается задача двухмерного масштабирования управляемых вероятностных тестов небольшого размера, задающих исходные шаблоны, с помощью матриц Адамара. Показывается ограниченность применения классических подходов генерирования тестов, основанных на перечислении кандидатов в тестовые наборы. С повышением пороговых значений мер различия двоичных тестовых наборов увеличивается вычислительная сложность построения таких тестов. Главной целью настоящей статьи является развитие методов построения тестов на базе исходных шаблонов путем их расширения до требуемой разрядности на основании применения формальных правил.

Методы. Для двухмерного масштабирования исходных шаблонов с заданными пороговыми значениями расстояния Хэмминга применяются матрицы Адамара и рекурсивная процедура Сильвестра для их построения. При проведении экспериментальных исследований использовался 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 искомых тестов используются как шаблоны, характеризующиеся минимальной разрядностью наборов, так и любые управляемые вероятностные тесты. Сама процедура характеризуется как одномерное масштабирование, которое увеличивает разрядность тестовых наборов, но сохраняет их количество. С целью одновременного увеличения разрядности наборов и их количества предлагается метод, основанный на двухмерном масштабировании шаблонов с применением матриц Адамара. Это позволяет строить управляемые вероятностные тесты без трудоемкой процедуры перечисления кандидатов в тестовые наборы и вычисления для них значений меры (мер) различия. Показано, что уникальное свойство ортогональности матриц Адамара с ростом их порядка позволяет достигать отношений среднего значения расстояния Хэмминга между тестовыми наборами к их разрядности, близких к 1/2. Отмечено, что характеристики исходных шаблонов несущественно влияют на характеристики результирующих тестов, построенных с применением матриц Адамара, которые

получены на основании рекурсивной процедуры Сильвестра. Работоспособность и эффективность предложенного подхода к построению управляемых вероятностных тестов оценены для случая тестирования запоминающих устройств. Показано, что управляемые вероятностные тесты, построенные с применением матриц Адамара, имеют заметно большую покрывающую способность по сравнению с вероятностными тестами.

Заключение. Рассмотрен метод генерирования тестовых наборов при формировании управляемых вероятностных тестов с использованием матриц Адамара. Основой предложенного метода является двухмерное масштабирование исходных шаблонов с применением указанных матриц. Показано, что использование различных шаблонов и их двухмерное масштабирование позволяют строить управляемые вероятностные тесты с требуемой разрядностью тестовых наборов и большим их количеством.

Ключевые слова: управляемые вероятностные тесты, двоичный тестовый набор, мера различия символьных наборов, расстояние Хэмминга, матрицы Адамара, процедура Сильвестра

Для цитирования. 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. – DOI: 10.37661/1816-0301-2025-22-2-63-80.

Конфликт интересов. Авторы заявляют об отсутствии конфликта интересов.

Поступила в редакцию | Received 06.03.2025

Подписана в печать | Accepted 24.03.2025

Опубликована | Published 30.06.2025

Scaling controlled random tests based on Hadamard matrices

Vyacheslav N. Yarmolik^{1✉}, Mikalai A. Shauchenka², Vita V. Petrovskaya¹

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

[✉]E-mail: yarmolik10ru@yahoo.com

²Technical University of Darmstadt,
Karolinenplatz, 5, Darmstadt, 64289, Germany

Abstract

Objectives. The problem of constructing controlled random tests is solved by two-dimensional scaling of initial templates using Hadamard matrices. The limitations of classical approaches to generating test patterns based on enumeration of candidates for test patterns are shown. With an increase in the threshold values of the difference measures of binary test patterns, the computational complexity of constructing such tests increases. The main goal of this article is to develop methods for constructing tests based on initial templates and their expansion to the required bit size based on the application of formal rules.

Methods. For two-dimensional scaling of initial templates with specified Hamming distance thresholds, Hadamard matrices and the Sylvester recursive procedure for their construction are applied. The experimental research employed the method of statistical trials.

Results. It is demonstrated that methods for constructing controlled random tests based on templates can be viewed as a procedure for scaling controlled random tests to the required bit size. Both templates characterized by the minimum bit size of patterns and any controlled random tests are used to construct the desired tests. The procedure itself is characterized as one-dimensional scaling, which increases the bit size of patterns while maintaining their quantity. To simultaneously increase the bit size and quantity of test sets, a method based on two-dimensional scaling of templates using Hadamard matrices is proposed. This allows for the construction of controlled random tests without the labor-intensive process of enumerating candidate test patterns and computing their difference measure values. It is shown that the unique orthogonality property of Hadamard matrices, as their order increases, enables achieving ratios of the average Hamming distance between test patterns to their bit

size close to 1/2. It is noted that the characteristics of the initial templates do not significantly affect the characteristics of the resulting tests constructed using Hadamard matrices obtained through the Sylvester recursive procedure. The feasibility and efficiency of the proposed approach to constructing controlled random tests are evaluated for the case of testing memory devices. It is demonstrated that controlled random tests constructed using Hadamard matrices have significantly higher coverage capability compared to random tests.

Conclusion. An approach for generating test patterns in the formation of controlled random tests using Hadamard matrices is considered. The proposed approach is based on two-dimensional scaling of initial templates using these matrices. 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 of test patterns and a larger number of them.

Keywords: technical diagnostics, controlled random tests, binary test pattern, measure of difference of symbol sets, Hamming distance, Hadamard matrices, Sylvester procedure

For citation. 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.). DOI: 10.37661/1816-0301-2025-22-2-63-80.

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

Введение. Для программного обеспечения, аппаратных средств и запоминающих устройств современных вычислительных систем важным является применение систематических методов построения тестовых процедур [1–4]. Большинство подходов автоматизированного создания тестовых процедур и тестовых данных реализуется с помощью различных форм их вероятностного формирования [1, 2, 5, 6]. Вероятностное тестирование (random testing) служит основополагающим подходом для тестирования вычислительных систем и их компонент в силу простоты своей концепции и реализуемости, а также эффективности при выявлении неисправностей [1, 2]. Основным недостатком вероятностного тестирования является его временная сложность, возникающая из-за большого количества тестовых наборов. Временные ограничения на реализацию процедур самотестирования являются критическими для современных встроенных систем (embedded systems), что также влияет на эффективность тестов [7–9]. Поэтому активно развивались и развиваются новые методы построения тестов, в которых случайный фактор используется наряду с детерминированными подходами к формированию тестов [1]. Лидирующее место среди них занимает управляемое (адаптивное) вероятностное тестирование (adaptive random testing) [10–14]. Данный вид тестирования основан на вычислении некоторых характеристик для управляемого генерирования очередного случайного тестового набора. Большинство таких тестов основано на применении расстояния Хэмминга в качестве характеристики, определяющей выбор очередного набора [15–17]. Тестовый набор выбирается из потенциальных кандидатов, сгенерированных случайным образом, по критерию максимума значения расстояния Хэмминга. В работах [1, 2, 10, 11] показано, что чем выше значения критериев выбора, в частности расстояния Хэмминга, тем сложнее процедура выбора тестового набора и заметно меньше длина формируемого управляемого вероятностного теста.

Высокая вычислительная сложность 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 других ограничениях удается избежать рутинной процедуры поиска тестовых наборов среди достаточно большого количества кандидатов в тесты, удовлетворяющих заданным критериям [1, 2, 10]. В качестве одного из многообещающих подходов можно рассматривать использование стандартных шаблонов, представляющих собой управляемые вероятностные тесты с малым числом тестовых наборов и небольшой их размерностью, для синтеза тестов с малыми вычислительными затратами [1, 2, 10, 18, 19]. Таким образом, задача построения управляемых вероятностных тестов с заданным расстоянием Хэмминга как критерием включения кандидата в генерируемый тест и невысокой вычислительной сложностью по-прежнему является актуальной.

1. Анализ управляемых вероятностных тестов. Управляемым вероятностным тестом $CRT = \{T_0, T_1, \dots, T_{q-1}\}$ является тест, состоящий из сгенерированных случайным образом тестовых наборов $T_i = t_{i,0} t_{i,1} \dots t_{i,m-1}$, где $t_{i,l} \in \{0, 1\}$, $l \in \{0, 1, \dots, m - 1\}$, а $i \in \{0, 1, \dots, q - 1\}$, таких, что характеристика (характеристики) набора T_i удовлетворяют некоторому критерию (кри-

териям) по отношению к предыдущим наборам T_0, T_1, \dots, T_{i-1} теста *CRT* [1, 2]. Очередной тестовый набор T_i управляемого вероятностного теста формируется максимально отличающимся от наборов, ранее включенных в тест. Таким образом, принимается гипотеза, что для двух тестовых наборов, имеющих максимальное отличие, количество обнаруживаемых неисправностей (ошибок) вторым набором будет максимальным [1, 2, 10, 11]. В качестве критерия различия очередного тестового набора T_i от предыдущих наборов T_0, T_1, \dots, T_{i-1} часто применяется 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 [20–22]. В качестве критерия выбора кандидата в тестовые наборы и описания формируемого управляемого вероятностного теста часто используется пороговое значение расстояния Хэмминга $minHD(T_i, T_j)$, его среднее значение $aveHD(T_i, T_j)$ и суммарное значение расстояний Хэмминга $totalHD(T_i, T_j)$ [1, 2, 11, 22].

Отметим, что в дальнейшем, не нарушая общности изложения материала, будем рассматривать случай двоичных данных $t_{i,l} \in \{0, 1\}$, $l \in \{0, 1, \dots, m-1\}$, тогда тестовый набор T_i будет представлять собой двоичный вектор $t_{i,0} t_{i,1} \dots t_{i,m-1}$, состоящий из m бит.

Отличие очередного тестового набора по отношению к предыдущим наборам определяется значением минимального порогового расстояния Хэмминга $minHD(T_i, T_j)$, величина которого достаточно часто является критерием выбора. В терминах теории помехоустойчивого кодирования характеристику $minHD(T_i, T_j)$ можно рассматривать как кодовое расстояние d кода T , которое равно наименьшему расстоянию Хэмминга между различными парами кодовых слов T_0, T_1, \dots, T_{q-1} . Соответственно, исходя из основ теории помехоустойчивого кодирования, были сформулированы важные выводы, которые учитывались при генерировании управляемых вероятностных тестов [2, 18, 20, 21, 23]. В указанных работах подчеркивалось, что особенностью управляемых вероятностных тестов является ограниченность их длины, которая определяется значением $d = minHD(T_i, T_j)$, используемым как критерий включения тестового набора в тест. Как следует из предельной оценки Хэмминга (Hamming bound), чем больше $minHD(T_i, T_j)$, тем меньше количество тестовых наборов, удовлетворяющих этому критерию [18, 21].

Как альтернатива построению управляемых вероятностных тестов классическим методом, основанным на переборе кандидатов в тесты, были предложены формальные процедуры генерирования тестов, исключающие трудоемкую процедуру поиска тестовых наборов среди потенциальных кандидатов в тесты [1, 2, 10].

Для построения управляемых вероятностных тестов с малым количеством тестовых наборов q в работе [18] рассмотрены управляемые вероятностные тесты *MMHD*(q) (Maximum Minimum Hamming Distance) с максимальным минимальным расстоянием Хэмминга $max_minHD(T_i, T_j) > m/2$ для произвольных тестовых наборов T_i и T_j теста *MMHD*(q). Данное множество тестов основано на применении теоремы Плоткина, которая позволяет оценить максимально возможное количество q кодовых слов в двоичном коде длиной m для максимально возможного кодового расстояния $d = max_minHD(T_i, T_j)$, а граница Плоткина (Plotkin bound) дает верхний предел этого количества [24]. Отметим, что указанная граница является оценкой величины q , которая достигается не для всех соотношений d и m . Например, для случая $q = 3$ в соответствии с теоремой Плоткина имеем $max_minHD(T_i, T_j) \leq 3m/4$. Однако для тестов *MMHD*(3) и *MMHD*(4) $max_min\{MMHD(3)\} = max_min\{MMHD(4)\} = 2m/3$ [18].

На основании величины расстояния Хэмминга $HD(T_i, T_j)$ для тестовых наборов T_i и T_j совместно с их декартовым расстоянием $CD(T_i, T_j)$ в работах [12, 19] рассмотрен метод синтеза оптимальных управляемых вероятностных тестов (Optimal Controlled Random Tests, *OCRT*). Для этих тестов $HD(T_i, T_j) \geq m/2$, а количество наборов *OCRT* определяется как $q = 2(\lceil \log_2 m \rceil + 1)$. Например, для $m = 4$ количество q тестовых наборов *OCRT* равняется шести, а для $m = 8$, соответственно, восьми.

Приведенные два примера управляемых вероятностных тестов, а также исчерпывающие тесты *ExT*(2^r) для малых разрядностей r , рассмотренные в работах [18, 25], позволили сформулировать понятие шаблона *CRT*(q, d, r) управляемого вероятностного теста.

Под тестовыми шаблонами в дальнейшем будем понимать управляемый вероятностный тест $CRT(q, d, r)$ с фиксированным количеством тестовых наборов q , заданным значением минимального расстояния Хэмминга $d = \min HD(T_i, T_j)$ и разрядностью r тестовых наборов [25].

Показано, что на основании исходного шаблона $CRT(q, d, r)$ оказывается возможным построение их семейства при использовании правил преобразования двоичных кодов, исследуемых и применяемых в теории помехоустойчивого кодирования [18, 21, 24, 25]. Эти правила позволяют модифицировать исходный шаблон, сохраняя при этом значения и соотношения его характеристик q , d и r . Примеры возможных вариантов шаблонов для случая четырех наборов теста $MMHD(4)$ с разрядностью $r = 3$ и теста $ExT(2^2)$ с разрядностью $r = 2$ приведены в табл. 1 [2, 18, 25].

Таблица 1

Примеры шаблонов, построенных на основании тестов $MMHD(4)$ и $ExT(2^2)$

Table 1

Examples of templates built on the basis of tests $MMHD(4)$ and $ExT(2^2)$

MMHD(4)				$ExT(2^2)$		
$CRT_0(4, 2, 3)$	$CRT_1(4, 2, 3)$	$CRT_2(4, 2, 3)$	$CRT_3(4, 2, 3)$	$CRT_0(4, 1, 2)$	$CRT_1(4, 1, 2)$	$CRT_2(4, 1, 2)$
$T_0 = 0\ 0\ 0$	$T_3 = 1\ 0\ 1$	$T_0 = 0\ 0\ 0$	$0\ 0\ 1$	$0\ 0$	$0\ 0$	$1\ 1$
$T_1 = 0\ 1\ 1$	$T_1 = 0\ 1\ 1$	$T_3 = 1\ 0\ 1$	$0\ 1\ 0$	$0\ 1$	$1\ 0$	$0\ 1$
$T_2 = 1\ 1\ 0$	$T_2 = 1\ 1\ 0$	$T_2 = 1\ 1\ 0$	$1\ 1\ 1$	$1\ 0$	$0\ 1$	$1\ 0$
$T_3 = 1\ 0\ 1$	$T_0 = 0\ 0\ 0$	$T_1 = 0\ 1\ 1$	$1\ 0\ 0$	$1\ 1$	$1\ 1$	$0\ 0$

Шаблон $CRT(4, 2, 3)$ и его варианты, представленные в табл. 1, соответствуют тесту $MMHD(4)$. Исчерпывающий тест $ExT(2^2)$, состоящий из всех возможных двухразрядных комбинаций двоичных значений, порождает семейство шаблонов $CRT(4, 1, 2)$. Для получения каждого экземпляра шаблона, приведенного в табл. 1, использовались простейшие правила, заключающиеся в перестановке строк и (или) столбцов [18, 11, 25].

С помощью подобных шаблонов с заданными характеристиками q , d и r строится управляемый вероятностный тест $CRT(q, d_m, m)$ для требуемой разрядности $m > r$ тестовых наборов, в котором сохраняется относительное соотношение расстояния Хэмминга и разрядности тестовых наборов, т. е. выполняется равенство $d/r = d_m/m$. Предположим, необходимо построить управляемый вероятностный тест $CRT(4, 10, 15)$, состоящий из четырех тестовых наборов с разрядностью $m = 15$, для которого $d_m = 10$. Классический подход предполагает генерирование подмножества кандидатов в тестовые наборы, представляющего собой равномерно распределенные двоичные векторы, и выбор одного из них, обеспечивающего характеристики q , d и r формируемого теста. Последовательно набор за набором строится искомый тест. Методика, основанная на применении шаблонов, исключает трудоемкую процедуру случайного перебора кандидатов в тесты [25]. Для построения теста $CRT(4, 10, 15)$ с относительным соотношением $d_m/m = 10/15 = 2/3$ можно применить один из шаблонов либо их подмножество, приведенное в табл. 1, так как для них отношение d/r также равняется $2/3$. Пример одного из возможных управляемых вероятностных тестов $CRT(4, 10, 15)$, построенных с применением шаблонов $CRT(4, 2, 3)$, приведен на рис. 1.

$$\begin{aligned} T_0 &= 0\ 0\ 0 \quad 1\ 0\ 1 \quad 0\ 0\ 0 \quad 1\ 0\ 1 \quad 0\ 0\ 0 \\ T_1 &= 0\ 1\ 1 \quad 0\ 1\ 1 \quad 1\ 0\ 1 \quad 0\ 0\ 0 \quad 0\ 1\ 1 \\ T_2 &= 1\ 1\ 0 \quad 1\ 1\ 0 \quad 1\ 1\ 0 \quad 1\ 1\ 0 \quad 1\ 1\ 0 \\ T_3 &= 1\ 0\ 1 \quad 0\ 0\ 0 \quad 0\ 1\ 1 \quad 0\ 1\ 1 \quad 1\ 0\ 1 \end{aligned}$$

Рис. 1. Управляемый вероятностный тест $CRT(4, 10, 15)$

Fig. 1. Controlled random test $CRT(4, 10, 15)$

При построении теста $CRT(4, 10, 15)$ использовались четыре из $4!$ возможных шаблонов $CRT(4, 2, 3)$, три из которых, а именно $CRT_0(4, 2, 3)$; $CRT_1(4, 2, 3)$; $CRT_2(4, 2, 3)$, в качестве примеров данных шаблонов приведены в табл. 1. Рассмотренные ранее методики построения управляемых вероятностных тестов, основанные на использовании шаблонов, можно представлять как процедуру масштабирования управляемых вероятностных тестов до требуемой раз-

рядности m . Для этой процедуры могут быть использованы как шаблоны, построенные для минимальных разрядностей, так и любые управляемые вероятностные тесты, для которых $r < m$. Сама процедура характеризуется как одномерная процедура масштабирования, которая сохраняет количество тестовых наборов q . Это, например, видно на рис. 1.

Весьма важным представляется построение управляемых вероятностных тестов с фиксированными значениями d и заданным m на основании масштабирования шаблонов, но с большим числом наборов, чем в используемых шаблонах. Другими словами, предлагается рассмотреть подход к двухмерному масштабированию шаблонов, при котором увеличивается как количество разрядов в тестовых наборах, так и число самих тестовых наборов в искомом teste.

2. Матрицы Адамара. Существенную роль в математике и информатике играют матрицы Адамара, которые были введены в математический обиход и интенсивно изучались и использовались в прошлом веке [26, 27]. Было замечено, что эти матрицы могут быть применены для цифровой обработки сигналов и построения кодов с большим кодовым расстоянием d . На практике широко применяются дискретное преобразование Уолша – Адамара и двоичные коды Адамара с кодовым расстоянием $d_m \geq m/2$ для четных m [26].

Квадратная матрица H_n порядка n с элементами $+1$ и -1 называется матрицей Адамара, если выполняется условие $H_n \cdot H_n^T = n \cdot E_n$, где E_n – единичная матрица размером n . На рис. 2 приведено несколько примеров матриц Адамара.

$$\begin{array}{c} \left| \begin{array}{cccc} 1 & 1 & 1 & 1 \\ 1 & -1 & -1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{array} \right| \\ \left| \begin{array}{cccc} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{array} \right| \\ \left| \begin{array}{cccc} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{array} \right| \end{array} \quad \begin{array}{c} n=2 \\ n=4 \\ n=8 \end{array}$$

Рис. 2. Примеры матриц Адамара для различных значений их порядка n

Fig. 2. Examples of Hadamard matrices for different values of their order n

Матрицы Адамара порядков $n = 2^k$ строятся на основании рекурсивной процедуры Сильвестра, которая на основании матрицы Адамара порядка $n = 2^k$ позволяет получить матрицу Адамара, порядок n которой равняется 2^{k+1} согласно соотношению [28]

$$H_{2^{k+1}} = \begin{vmatrix} H_{2^k} & H_{2^k} \\ H_{2^k} & -H_{2^k} \end{vmatrix}. \quad (1)$$

На рис. 2 видно, что матрица H_4 получена на основании H_2 , а матрица H_8 – на основании H_4 в соответствии с процедурой (1).

В общем случае для произвольной матрицы Адамара с элементами $h_{ij} \in \{+1, -1\}$, где $i, j \in \{1, 2, \dots, n\}$, для любой пары строк с индексами $i \neq j$ верно равенство

$$h_{i,1}h_{j,1} + h_{i,2}h_{j,2} + \dots + h_{i,n}h_{j,n} = 0. \quad (2)$$

Последнее равенство доказывает, что различные строки матрицы Адамара попарно ортогональны. Это следует из того, что число слагаемых в выражении (2), равных $+1$, совпадает с числом слагаемых, равных -1 . Соответственно, при n четном любые две строки совпадают ровно в $n/2$ позициях и различаются в остальных $n/2$ позициях.

Уникальное свойство ортогональности матриц Адамара предопределило их широкое применение в различных областях, а процедура Сильвестра позволяет решать задачу масштабирования полученных решений на их основе. Рассмотрим применение матриц Адамара для задачи

синтеза управляемых вероятностных тестов на базе шаблонов $CRT(q, d, r)$ путем их масштабирования до необходимой разрядности тестовых наборов $m > r$ и одновременного увеличения их количества. В отличие от ранее рассмотренных методов синтеза управляемых вероятностных тестов $CRT(q, d_m, m)$ применение матриц Адамара позволяет реализовать двухмерное масштабирование исходных шаблонов.

3. Применение матриц Адамара для построения управляемых вероятностных тестов. Свойство ортогональности матриц Адамара легко обобщается на случай матриц Адамара, представленных элементами, которые описываются двоичными символами 0 и 1. Для перехода к двоичному представлению матриц Адамара используется преобразование $t = (1 - b)/2$, где $b \in \{+1, -1\}$ и $t \in \{0, 1\}$. При этом элемент +1 оригинальной матрицы Адамара может принимать как двоичное значение t , так и значения набора, состоящего из $r > 1$ двоичных переменных $T = t_1 t_2 \dots t_r$. В обоих случаях элемент -1 заменяется на инверсные представления, соответственно \bar{t} и $\bar{T} = \bar{t}_1 \bar{t}_2 \dots \bar{t}_r$. Исходя из свойства ортогональности матриц Адамара (2), можно заключить, что для случая, когда элементами h_{ij} матрицы являются двоичные значения t , величина расстояния Хэмминга между двумя произвольными строками матрицы Адамара порядка n равняется $n/2$. Для матриц Адамара, использующих в качестве элементов h_{ij} двоичные векторы $T = t_1 t_2 \dots t_r$, величина расстояния Хэмминга определяется как $(n \cdot r)/2$. Это следует как из свойства ортогональности матриц Адамара, так и из свойства расстояния Хэмминга, задаваемого равенством $HD(T, \bar{T}) = r$. Например, для двух элементов матрицы Адамара, представленных двоичным набором $T = 0 0 1$ ($r = 3$) и его инверсией $\bar{T} = 1 1 0$, соответствующая матрица Адамара порядка $n = 4$ (см. рис. 2) выглядит следующим образом:

$$H_4 = \begin{vmatrix} 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 & 1 & 0 & 1 & 1 & 0 \end{vmatrix}. \quad (3)$$

Видно, что расстояние Хэмминга d между строками приведенной матрицы равняется $(n \cdot r)/2 = (4 \cdot 3)/2 = 6$, а их совокупность можно рассматривать как управляемый вероятностный тест $CRT(q, d, m) = CRT(4, 6, 12)$, состоящий из четырех тестовых наборов по 12 двоичных значений. Следует отметить, что аналогичный тест может быть построен для любого двоичного набора $T = t_1 t_2 \dots t_r$, состоящего из r двоичных значений. Такой тест $CRT(n, (n \cdot r)/2, n \cdot r)$ будет состоять из q тестовых наборов, число которых равняется порядку матрицы Адамара n . Каждый из тестовых наборов состоит из $m = n \cdot r$ двоичных значений, а расстояние Хэмминга между наборами определяется значением $d = (n \cdot r)/2$.

Основываясь на семействе управляемых вероятностных тестов $CRT(n, (n \cdot r)/2, n \cdot r)$, порождаемых матрицей Адамара H_n порядка n , как семейство шаблонов, легко синтезировать аналогичные тесты вида $CRT(2 \cdot n, (n \cdot r)/2, n \cdot r)$ с удвоенным количеством наборов. Данная процедура основана на использовании исходной матрицы H_n и матрицы $-H_n$, которая строится на основании H_n путем инвертирования всех ее двоичных элементов [21, 26]. Для $n = 4$ и теста $CRT(4, 6, 12)$, описываемого матрицей (3), соответствующий управляемый вероятностный тест $CRT(8, 6, 12)$ примет вид рис. 3.

$$\begin{aligned} T_0 &= 0 0 1 \ 0 0 1 \ 0 0 1 \ 0 0 1 \\ T_1 &= 0 0 1 \ 1 1 0 \ 0 0 1 \ 1 1 0 \\ T_2 &= 0 0 1 \ 0 0 1 \ 1 1 0 \ 1 1 0 \\ T_3 &= 0 0 1 \ 1 1 0 \ 1 1 0 \ 0 0 1 \\ T_4 &= 1 1 0 \ 1 1 0 \ 1 1 0 \ 1 1 0 \\ T_5 &= 1 1 0 \ 0 0 1 \ 1 1 0 \ 0 0 1 \\ T_6 &= 1 1 0 \ 1 1 0 \ 0 0 1 \ 0 0 1 \\ T_7 &= 1 1 0 \ 0 0 1 \ 0 0 1 \ 1 1 0 \end{aligned}$$

Рис. 3. Управляемый вероятностный тест $CRT(8, 6, 12)$

Fig. 3. Controlled random test $CRT(8, 6, 12)$

Двоичный шаблон $CRT(q, d, r)$, используемый для построения управляемых вероятностных тестов, представляет собой матрицу, состоящую из q строк и r столбцов, которая также может быть использована в качестве элемента $+1$ матрицы Адамара. Тогда элементу -1 будет соответствовать инверсная матрица $-CRT(q, d, r)$. С применением процедуры Сильвестра для случая матрицы Адамара, использующей в качестве элемента $+1$ шаблон $CRT_0(4, 1, 2)$, который приведен в табл. 1, строится результирующая матрица для $n = 4$:

$$\begin{array}{cccc|cccc} CRT(4,1,2) & CRT(4,1,2) & CRT(4,1,2) & CRT(4,1,2) & & 00 & 00 & 00 & 00 \\ & & & & & 01 & 01 & 01 & 01 \\ & & & & & 10 & 10 & 10 & 10 \\ & & & & & 11 & 11 & 11 & 11 \\ \\ CRT(4,1,2) & -CRT(4,1,2) & CRT(4,1,2) & -CRT(4,1,2) & = & 00 & 11 & 00 & 11 \\ & & & & & 01 & 10 & 01 & 10 \\ & & & & & 10 & 01 & 10 & 01 \\ & & & & & 11 & 00 & 11 & 00 \\ \\ CRT(4,1,2) & CRT(4,1,2) & -CRT(4,1,2) & -CRT(4,1,2) & & 00 & 00 & 11 & 11 \\ & & & & & 01 & 01 & 10 & 10 \\ & & & & & 10 & 10 & 01 & 01 \\ & & & & & 11 & 11 & 00 & 00 \\ \\ CRT(4,1,2) & -CRT(4,1,2) & -CRT(4,1,2) & CRT(4,1,2) & & 00 & 11 & 11 & 00 \\ & & & & & 01 & 10 & 10 & 01 \\ & & & & & 10 & 01 & 01 & 10 \\ & & & & & 11 & 00 & 00 & 11 \end{array} . \quad (4)$$

Двоичные значения строк полученной матрицы (4) можно рассматривать как тестовые наборы управляемого вероятностного теста $CRT(n \cdot q, n \cdot d, m) = CRT(16, 4, 8)$. Таким образом, в результате применения шаблона $CRT(4, 1, 2)$ в качестве элемента матрицы Адамара был сформирован управляемый вероятностный тест $CRT(16, 4, 8)$ как результат его двухмерного масштабирования. Соответственно, в $n = 4$ раза было увеличено количество двоичных разрядов тестовых наборов и во столько же раз число самих наборов. Нетрудно убедиться, что для теста $CRT(16, 4, 8)$ величина $HD(T_i, T_j) = 4$. Соответственно, в нем сохраняется относительное соотношение 1/2 между величиной расстояния Хэмминга и разрядностью наборов, такое же, как и для исходного шаблона.

Матрицы Адамара позволяют выполнять процедуру масштабирования в $n \in \{2, 4, 8, \dots\}$ раз в зависимости от порядка n используемой матрицы. Само масштабирование заключается в использовании в качестве элемента $+1$ оригинальной матрицы Адамара масштабируемый тест $CRT(q, d, r)$, а вместо элемента -1 – его инверсное представление $-CRT(q, d, r)$. Для общего случая двухмерного масштабирования шаблонов либо произвольных управляемых вероятностных тестов с использованием матриц Адамара справедливо следующее утверждение.

Утверждение. Результатом масштабирования управляемого вероятностного теста $CRT(q, d, r)$ при $d \leq r/2$ с помощью матрицы Адамара n -го порядка является тест $CRT(n \cdot q, n \cdot d, n \cdot r)$, а при $d > r/2$ – тест $CRT(n \cdot q, (n \cdot r)/2, n \cdot r)$.

Доказательство. Результатом масштабирования теста $CRT(q, d, r)$ будет являться тест, состоящий из $n \cdot q$ тестовых наборов, каждый из которых включает $n \cdot r$ двоичных данных. В зависимости от значения d его $\min HD(T_i, T_j)$ принимает величину $n \cdot d$ или $(n \cdot r)/2$. В матричном представлении результатом масштабирования является матрица размерностью $(n \cdot q) \times (n \cdot r)$, в которой строки группируются в n блоков, состоящих из q строк. Каждый блок определяется строкой матрицы Адамара, используемой для масштабирования. Например, первый блок матрицы (4) определяется как $CRT(4, 1, 2) \ CRT(4, 1, 2) \ CRT(4, 1, 2) \ CRT(4, 1, 2)$ и включает $q = 4$ строки результирующей двоичной матрицы.

Рассмотрим значение величины расстояния Хэмминга $HD(T_i, T_j)$ между двумя произвольными строками матрицы, полученной в результате масштабирования исходного теста. Указанные строки матрицы представляют собой два тестовых набора T_i и T_j , каждый из которых состоит из $n \cdot r$ двоичных значений. Первоначально отметим, что для расстояния Хэмминга между двоичными наборами T_i и T_j произвольной разрядности m справедливо соотношение $HD(T_i, T_j) + HD(T_i, \bar{T}_j) = m$.

Рассмотрим тестовые наборы T_i и T_j , принадлежащие различным блокам строк. В силу свойства ортогональности матриц Адамара и отмеченной зависимости для расстояния Хэмминга $HD(T_i, T_j) = (n \cdot r)/2$. Если же тестовые наборы принадлежат одному из n блоков строк, то для них не выполняется свойство ортогональности. В этом случае определяющим будет свойство расстояния Хэмминга, задаваемое равенством $HD(T_i, T_j) = HD(\bar{T}_i, \bar{T}_j)$, а структура каждого блока будет повторять результат применения процедуры одномерного масштабирования, для которой сохраняется отношение расстояния Хэмминга к разрядности наборов. Таким образом, для двух тестовых наборов T_i и T_j , принадлежащих одному блоку строк, расстояние Хэмминга $HD(T_i, T_j)$ равняется $n \cdot d$.

Значением $\min HD(T_i, T_j)$ для результата масштабирования управляемого вероятностного теста $CRT(q, d, r)$ будет минимальное значение из двух ранее полученных $HD(T_i, T_j) = (n \cdot r)/2$ и $HD(T_i, T_j) = n \cdot d$. Соответственно, если $d \leq r/2$, то $n \cdot d \leq (n \cdot r)/2$, и результатом масштабирования будет управляемый вероятностный тест $CRT(n \cdot q, n \cdot d, n \cdot r)$. В противном случае, когда $d > r/2$, будет получен тест $CRT(n \cdot q, (n \cdot r)/2, n \cdot r)$, что и требовалось доказать.

В качестве иллюстрации приведенного утверждения рассмотрим три примера двухмерного масштабирования шаблонов, один из которых обозначен в табл. 1 как $CRT_0(4, 2, 3)$ и состоит из четырех трехразрядных тестовых наборов. В качестве второго примера используем шаблон $CRT(4, 1, 3)$, состоящий также из четырех трехразрядных наборов $T_0 = 0\ 0\ 0$, $T_1 = 0\ 0\ 1$, $T_2 = 0\ 1\ 0$, и $T_3 = 0\ 1\ 1$. Третьим шаблоном $CRT(4, 0, 3)$ является нулевая двоичная матрица размерностью 4×3 . Отличием приведенных шаблонов является значение $d = \min HD(T_i, T_j)$, которое в первом случае равняется двум и удовлетворяет неравенству $d \leq r/2$, а во втором случае равняется единице, что соответствует соотношению $d \leq r/2$. В третьем случае имеем вырожденный случай шаблона, который мог быть получен в результате его вероятностного формирования и для которого $d = 0$. Результат двухмерного двукратного ($n = 2$) масштабирования всех трех шаблонов показан на рис. 4.

$T_0 =$	$0\ 0\ 0$	$0\ 0\ 0$	$T_0 =$	$0\ 0\ 0$	$0\ 0\ 0$	$T_0 =$	$0\ 0\ 0$	$0\ 0\ 0$
$T_1 =$	$0\ 1\ 1$	$0\ 1\ 1$	$T_1 =$	$0\ 0\ 1$	$0\ 0\ 1$	$T_1 =$	$0\ 0\ 0$	$0\ 0\ 0$
$T_2 =$	$1\ 1\ 0$	$1\ 1\ 0$	$T_2 =$	$0\ 1\ 0$	$0\ 1\ 0$	$T_2 =$	$0\ 0\ 0$	$0\ 0\ 0$
$T_3 =$	$1\ 0\ 1$	$1\ 0\ 1$	$T_3 =$	$0\ 1\ 1$	$0\ 1\ 1$	$T_3 =$	$0\ 0\ 0$	$0\ 0\ 0$
$T_4 =$	$0\ 0\ 0$	$1\ 1\ 1$	$T_4 =$	$0\ 0\ 0$	$1\ 1\ 1$	$T_4 =$	$0\ 0\ 0$	$1\ 1\ 1$
$T_5 =$	$0\ 1\ 1$	$1\ 0\ 0$	$T_5 =$	$0\ 0\ 1$	$1\ 1\ 0$	$T_5 =$	$0\ 0\ 0$	$1\ 1\ 1$
$T_6 =$	$1\ 1\ 0$	$0\ 0\ 1$	$T_6 =$	$0\ 1\ 0$	$1\ 0\ 1$	$T_6 =$	$0\ 0\ 0$	$1\ 1\ 1$
$T_7 =$	$1\ 0\ 1$	$0\ 1\ 0$	$T_7 =$	$0\ 1\ 1$	$1\ 0\ 0$	$T_7 =$	$0\ 0\ 0$	$1\ 1\ 1$

a)

b)

c)

Рис. 4. Результат масштабирования шаблонов: a) $CRT(4, 2, 3)$; b) $CRT(4, 1, 3)$; c) $CRT(4, 0, 3)$

Fig. 4. Result of template scaling: a) $CRT(4, 2, 3)$; b) $CRT(4, 1, 3)$; c) $CRT(4, 0, 3)$

Анализ приведенных результатов свидетельствует о том, что в первом случае на основании масштабирования шаблона $CRT(q, d, r) = CRT(4, 2, 3)$ в $n = 2$ раза был получен управляемый вероятностный тест $CRT(n \cdot q, (n \cdot r)/2, n \cdot r) = CRT(2 \cdot 4, (2 \cdot 3)/2, 2 \cdot 3) = CRT(8, 3, 6)$. Свидетельством этому, например, является значение $HD(T_0, T_4) = 3$ (рис. 4). Во втором случае результатом масштабирования является тест $CRT(n \cdot q, n \cdot d, n \cdot r) = CRT(8, 2, 6)$. Результатом, полученным на основании третьего шаблона, является тест $CRT(n \cdot q, n \cdot d, n \cdot r) = CRT(8, 0, 6)$, так как в нем присутствуют тестовые наборы, для которых расстояние Хэмминга равняется нулю. Все три примера подтверждают справедливость приведенного утверждения.

Важным следствием рассмотренного выше утверждения является усреднение свойств результирующего управляемого вероятностного теста независимо от характеристик исходных шаблонов. В первую очередь это касается расстояния Хэмминга, среднее значение $aveHD(T_i, T_j)$ которого с ростом порядка n используемой матрицы Адамара стремится к $(n \cdot r)/2$. Причиной данного факта является то, что удельный вес пар тестовых наборов T_i и T_j , принадлежащих различным блокам строк, существенно возрастает с ростом n . Для таких тестовых наборов $HD(T_i, T_j) = (n \cdot r)/2$, что следует из свойства ортогональности строк матриц Адамара. В то же время тестовые наборы, принадлежащие одному и тому же блоку строк, имеют различные значения расстояний Хэмминга и могут быть оценены средним расстоянием Хэмминга используемых шаблонов $aveHD_t(T_i, T_j)$. Например, для второго шаблона $CRT(4, 1, 3)$, приведенного на рис. 4, b, это значение будет вычисляться как $aveHD_2(T_i, T_j) = (HD(T_0, T_1) + HD(T_0, T_2) + HD(T_0, T_3) + HD(T_1, T_2) + HD(T_1, T_3) + HD(T_2, T_3))/6 = (1 + 1 + 2 + 2 + 1 + 1)/6 = 1,333$. Для первого шаблона $aveHD_1(T_i, T_j) = 2$, а для третьего – $aveHD_3(T_i, T_j) = 0$. Зависимость значения $aveHD(T_i, T_j)$ управляемого вероятностного теста от аналогичной характеристики $aveHD_t(T_i, T_j)$ используемого для его построения шаблона и порядка матрицы Адамара n определяется соотношением

$$\begin{aligned} aveHD(T_i, T_j) &= \frac{\{aveHD_t(T_i, T_j) \cdot n^2 \cdot \binom{q}{2} + \frac{n \cdot r}{2} \cdot [\binom{n \cdot q}{2} - n \cdot \binom{q}{2}]\}}{\binom{n \cdot q}{2}} = \\ &= \frac{n \cdot r}{2} + \frac{\{n^2 \cdot \binom{q}{2} [aveHD_t(T_i, T_j) - \frac{r}{2}]\}}{\binom{n \cdot q}{2}} = \frac{n \cdot r}{2} + \frac{n \cdot q - n}{n \cdot q - 1} [aveHD_t(T_i, T_j) - \frac{r}{2}]. \end{aligned} \quad (5)$$

Значение первого слагаемого $(n \cdot r)/2$ в выражении (5) с ростом n является определяющим, и оно нивелирует влияние второго слагаемого, в котором учитываются свойства выбранного шаблона в виде его среднего расстояния Хэмминга $aveHD_t(T_i, T_j)$. Приведенное значение среднего расстояния Хэмминга $aveHD(T_i, T_j)$ к одному разряду $(n \cdot r)$ -разрядного тестового набора ($aveHD(T_i, T_j)/(n \cdot r)$), полученного в результате масштабирования с помощью матриц Адамара для трех ранее рассмотренных шаблонов, представлено в табл. 2.

Таблица 2

Приведенное значение среднего расстояния Хэмминга $aveHD(T_i, T_j)/(n \cdot r)$

Table 2

The given value of the average Hamming distance $aveHD(T_i, T_j)/(n \cdot r)$

Шаблон Sample	Значение $aveHD(T_i, T_j)/(n \cdot r)$ Value $aveHD(T_i, T_j)/(n \cdot r)$			
	$n = 2$	$n = 4$	$n = 8$	$n = 16$
$CRT(4, 2, 3)$	0,571	0,533	0,516	0,507
$CRT(4, 1, 3)$	0,476	0,488	0,494	0,497
$CRT(4, 0, 3)$	0,285	0,400	0,451	0,476

Из табл. 2 видно, что численное значение $aveHD(T_i, T_j)/(n \cdot r)$ при увеличении n сходится к величине 0,5. Отметим, что в данном случае в качестве элемента +1 оригинальной матрицы Адамара использовался шаблон $CRT(q, d, r)$, а вместо элемента -1 – его инверсное представление $-CRT(q, d, r)$, для которых $q > 1$. Как следует из соотношения (5), при $q = 1$ значение $aveHD(T_i, T_j)/(n \cdot r)$ строго равняется 0,5, что является следствием уникальных свойств матриц Адамара.

4. Результаты экспериментальных исследований. Для подтверждения полученных авторами результатов был проведен ряд вычислительных и практических экспериментов. Исследовалась эффективность управляемых вероятностных тестов, полученных в результате масштабирования шаблонов с помощью матриц Адамара. В качестве объекта применения управляемых вероятностных тестов рассматривались запоминающие устройства (ЗУ) и их наиболее

сложные для обнаружения кодочувствительные неисправности [2, 25, 29, 30]. В частности, оценивалась эффективность обнаружения пассивных кодочувствительных неисправностей (Passive Neighborhood Pattern Sensitive Faults, PNPSF k), где k обозначает количество произвольных ячеек ЗУ, участвующих в моделируемой неисправности в ЗУ емкостью N бит. Подобные неисправности являются объектом активных исследований с целью нахождения эффективных тестовых процедур для их обнаружения [2, 25, 30–33]. Отметим, что результаты, полученные для PNPSF k , легко обобщаются и на другие классы кодочувствительных и иных неисправностей ЗУ в силу того, что PNPSF k является наиболее трудно обнаруживаемой разновидностью неисправностей ЗУ [2, 31–33].

В качестве теста ЗУ при проведении экспериментальных исследований использовался широко применяемый на практике маршевый тест MATS++ [2, 25, 30–33]. Его неразрушающая версия $\{\uparrow(Rt, W\bar{t}); \downarrow(\bar{R}\bar{t}, Wt, Rt)\}$ состоит из двух последовательных фаз, каждая из которых содержит операции чтения (R) из текущей ячейки ЗУ и записи (W) в эту ячейку ее содержимого $t \in \{0,1\}$ или инверсного значения \bar{t} . Символ \uparrow обозначает возрастающую последовательность адресов ячеек ЗУ, а символ \downarrow – убывающую [2]. Известно, что, основываясь на наличии степеней свободы маршевых тестов, в качестве начального состояния ЗУ может быть использовано любое начальное состояние его ячеек. Показано, что независимо от первоначального состояния ЗУ полнота покрытия маршевого теста $FC_{March_Test}(PNPSFk)$, определяемая как отношение обнаруженных им неисправностей к их общему числу, принимает постоянное значение [2, 33]. Например, тест MATS++, имеющий сложность теста $6N$, характеризуется полнотой покрытия $FC_{MATS++}(PNPSFk) = 1/2^{k-1} \cdot 100\%$. Соответственно, для PNPSF3 и PNPSF5 эти значения равняются $FC_{MATS++}(PNPSF3) = 25\%$ и $FC_{MATS++}(PNPSF5) = 6,25\%$. Невысокая полнота покрытия однократных маршевых тестов предопределила их многократное применение с изменяемыми начальными условиями, в том числе и начальным состоянием ячеек ЗУ. Эффективность l -кратного применения маршевого теста MATS++ при случайных начальных состояниях ячеек ЗУ оценивается соотношением для суммарной полноты покрытия PNPSF k неисправностей ЗУ [2, 33]:

$$FC_{MATS++}(PNPSFk, l) = (1 - (1 - 1/2^{k-1})^l) \cdot 100\%. \quad (6)$$

Выражение (6) определяет полноту покрытия $FC_{MATS++}(PNPSFk, l)$ неисправностей PNPSF k как результат последовательного применения теста MATS++ для l различных состояний ячеек ЗУ. Эти состояния формируются как случайные независимые значения, описываемые равномерным законом распределения. Соответственно, многократное тестирование ЗУ с использованием случайных начальных состояний можно интерпретировать как вероятностное тестирование со всеми его достоинствами и недостатками. Альтернативным решением l -кратного тестирования ЗУ является применение в качестве начальных состояний ячеек ЗУ тестовых наборов управляемого вероятностного теста. Для оценки предлагаемых авторами решений первоначально рассмотрим эффективность использования значений строк классических матриц Адамара в качестве начальных состояний ЗУ по отношению к состояниям ячеек ЗУ, формируемым случайнм образом.

В проведенных экспериментальных исследованиях использовалось авторское программное приложение, которое реализует полный цикл эксперимента, включая генерацию случайных равномерно распределенных двоичных векторов. Для формирования этих векторов применяется стандартный генератор псевдослучайных чисел языка C#, основанный на линейном конгруэнтном методе, обеспечивающем достаточную случайность и соответствующем требованиям равномерного распределения. Данное приложение неоднократно использовалось авторами в своих исследованиях (например, в работах [2, 25, 30, 33]) для получения равномерно распределенных двоичных векторов. В описанных ниже экспериментах приложение применяется для генерирования вероятностных тестов с требуемой разрядностью (*Random*), а также кодочувствительных неисправностей PNPSF k в части формирования адресов k ячеек ЗУ. В обоих случаях необходимо обеспечение равномерного распределения при формировании как тестовых наборов, так и адресов ячеек ЗУ, участвующих в неисправности.

Эксперимент 1. Реализовывалась процедура 16-кратного тестирования ЗУ емкостью $N = 8$ бит с применением маршевого теста MATS++. Первоначально определялась полнота покрытия неисправностей PNPSF5 для случая использования в качестве начальных состояний ячеек ЗУ равномерно распределенных случайных значений. В качестве альтернативного подхода оценивалась полнота покрытия при использовании кода Адамара, описываемого матрицами Адамара H_8 и $-H_8$, значения строк которого применялись в качестве начального состояния ЗУ. В терминах управляемых вероятностных тестов указанный код Адамара описывается как $CRT(16, 4, 8)$. В табл. 3 приведены экспериментальные результаты для одной реализации 16-кратного тестирования ЗУ при использовании случайных значений, которую можно интерпретировать как реализацию многократного вероятностного тестирования. В данном и следующем эксперименте формировались две разновидности кодочувствительных неисправностей PNPSF3 и PNPSF5, каждая из которых была протестирована для всех возможных сочетаний адресов ячеек ЗУ, участвующих в указанных неисправностях [2, 33]. В последующих экспериментах, где размер памяти составлял 24 бита и более, генерировались те же виды неисправностей PNPSF3 и PNPSF5, для каждой из которых случайным образом формировалось по 1000 неисправностей, отличающихся адресами участвующими в них ячеек ЗУ. Для сравнения в табл. 3 представлены аналогичные результаты, но для случая применения в качестве начальных состояний ЗУ тестовых наборов управляемого вероятностного теста $CRT(16, 4, 8)$.

Таблица 3

Полнота покрытия FC_{Random} и $FC_{Hadamard}$ неисправностей PNPSF5, %

Table 3

The faults coverage FC_{Random} and $FC_{Hadamard}$ of PNPSF5, %

<i>Random</i>	FC_{Random}	$Hadamar(H_8, -H_8)$	$FC_{Hadamard}$
$T_0= 1\ 1\ 1\ 1\ 0\ 1\ 0\ 1$	6,25	$T_0= 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0$	6,25
$T_1= 0\ 0\ 1\ 0\ 0\ 0\ 1$	12,41	$T_1= 0\ 1\ 0\ 1\ 0\ 1\ 0\ 1$	12,41
$T_2= 0\ 0\ 0\ 1\ 1\ 0\ 0\ 0$	18,57	$T_2= 0\ 0\ 1\ 1\ 0\ 0\ 1\ 1$	18,48
$T_3= 1\ 0\ 1\ 1\ 1\ 0\ 1\ 1$	24,55	$T_3= 0\ 1\ 1\ 0\ 0\ 1\ 1\ 0$	24,46
$T_4= 1\ 1\ 0\ 0\ 1\ 1\ 0\ 0$	30,63	$T_4= 0\ 0\ 0\ 0\ 1\ 1\ 1\ 1$	30,36
$T_5= 1\ 0\ 0\ 0\ 1\ 0\ 1\ 1$	35,27	$T_5= 0\ 1\ 0\ 1\ 1\ 0\ 1\ 0$	36,16
$T_6= 1\ 1\ 1\ 0\ 0\ 0\ 0\ 1$	38,68	$T_6= 0\ 0\ 1\ 1\ 1\ 1\ 0\ 0$	41,88
$T_7= 1\ 1\ 0\ 1\ 1\ 0\ 1\ 1$	42,32	$T_7= 0\ 1\ 1\ 0\ 1\ 0\ 0\ 1$	47,50
$T_8= 0\ 1\ 1\ 0\ 0\ 1\ 1\ 0$	48,21	$T_8= 1\ 1\ 1\ 1\ 1\ 1\ 1\ 1$	53,12
$T_9= 0\ 0\ 1\ 0\ 0\ 1\ 1\ 1$	51,61	$T_9= 1\ 0\ 1\ 0\ 1\ 0\ 1\ 0$	58,66
$T_{10}= 1\ 1\ 0\ 1\ 0\ 1\ 1\ 0$	56,07	$T_{10}= 1\ 1\ 0\ 0\ 1\ 1\ 0\ 0$	64,11
$T_{11}= 1\ 0\ 1\ 0\ 0\ 0\ 0\ 1$	56,98	$T_{11}= 1\ 0\ 0\ 1\ 1\ 0\ 0\ 1$	69,46
$T_{12}= 0\ 0\ 1\ 1\ 1\ 1\ 1\ 0$	61,43	$T_{12}= 1\ 1\ 1\ 1\ 0\ 0\ 0\ 0$	74,73
$T_{13}= 1\ 0\ 1\ 1\ 1\ 0\ 0\ 0$	64,29	$T_{13}= 1\ 0\ 1\ 0\ 0\ 1\ 0\ 1$	79,91
$T_{14}= 1\ 0\ 0\ 0\ 1\ 1\ 0\ 1$	67,23	$T_{14}= 1\ 1\ 0\ 0\ 0\ 0\ 1\ 1$	85,00
$T_{15}= 0\ 1\ 1\ 0\ 1\ 1\ 0\ 0$	70,09	$T_{15}= 1\ 0\ 0\ 1\ 0\ 1\ 1\ 0$	90,00

В каждой строке табл. 3 приведены значения начальных состояний ячеек ЗУ и суммарное значение полноты покрытия, полученной при текущей реализации теста MATS++ и всех предыдущих его реализаций. Видно, что полнота покрытия неисправностей PNPSF5 при выполнении 16-кратного тестирования ЗУ в случае теста $CRT(16, 4, 8)$ равняется 90 %. В то же время приведенная реализация вероятностного теста обеспечила только 70 %. Усредненное значение полноты покрытия по 1000 реализациям вероятностного тестирования оказалось равным 65,47 % (65,5 % для 10 000 реализаций, 65,55 % для 100 000 реализаций), что сравнимо с оценкой (6), принимающей значение 64,39 % для $k = 5$ и $l = 16$. Приведенные оценки свидетельствуют о заметном преимуществе использования в качестве начальных состояний ЗУ управляемых вероятностных тестов, построенных на матрицах Адамара, в сравнении с вероятностным подходом.

Эксперимент 2. Важным результатом данной статьи является двухмерное масштабирование управляемых вероятностных тестов с применением матриц Адамара. Для оценки эффективности такого подхода рассмотрим применение управляемого вероятностного теста $CRT(4, n \cdot 2, n \cdot 3)$ как

результата одномерного масштабирования шаблона $CRT(4, 2, 3)$, приведенного в табл. 1 [25]. Сама процедура одномерного масштабирования заключается в применении простейшего правила повторения наборов шаблона до требуемой разрядности $n \cdot 3$, соответствующей емкости N ЗУ, где $n = 2, 4, 8$. Выбор фиксированных значений n позволяет сопоставлять эффективность тестов $CRT(4, n \cdot 2, n \cdot 3)$ с тестами $CRT(n \cdot 4, (n \cdot 3)/2, n \cdot 3)$, полученными в результате двухмерного масштабирования того же шаблона с применением подхода, основанного на матрицах Адамара порядка $n = 2, 4, 8$. В табл. 4 приведены значения полноты покрытия $FC_{MATS++}(PNPSF3, l)$ в результате применения тестов $CRT(4, n \cdot 2, n \cdot 3) = CRT(4, 4, 6)$ и теста $CRT(n \cdot 4, (n \cdot 3)/2, n \cdot 3) = CRT(8, 3, 6)$, полученных в результате одномерного и двухмерного масштабирования исходного шаблона при $n = 2$. Тест $CRT(8, 3, 6)$ представлен его восемью наборами T_0, T_1, \dots, T_7 на рис. 4, а, а тест $CRT(4, 4, 6)$ описывается первыми четырьмя наборами T_0, T_1, T_2, T_3 теста $CRT(8, 4, 6)$ (см. тот же рис. 4, а). Для сравнительного сопоставления в табл. 4 приведены результаты эффективности применения вероятностного теста (*Random*) на основании оценки полноты покрытия $FC_{MATS++}(PNPSF3, l)$.

Таблица 4
Полнота покрытия $FC_{MATS++}(PNPSF3, l)$ для ЗУ емкостью 6 бит, %

Table 4
The faults coverage $FC_{MATS++}(PNPSF3, l)$ for 6-bit memory, %

T_i	T_0	T_1	T_2	T_3	T_4	T_5	T_6	T_7
$CRT(4, 4, 6)$	25	48,33	70,00	90,00	—	—	—	—
$CRT(8, 3, 6)$	25	48,33	70,00	90,00	95,00	98,33	100	100
<i>Random</i>	25	43,75	57,81	68,35	76,28	82,20	86,65	89,98

При реализации l -кратного применения теста MATS++ для случаев $CRT(4, 4, 6)$ и $CRT(8, 3, 6)$ начальные состояния ячеек ЗУ определяются наборами T_0, T_1, \dots (см. рис. 4). Таким образом, в первом случае реализовано четырехкратное применение теста и достигнута полнота покрытия, равная 90 %. Во втором случае при реализации восьмикратного применения теста MATS++ максимальная полнота покрытия, равная 100 %, достигается уже при седьмой ($l = 7$) итерации применения теста. В нижней строчке табл. 3 приведены средние значения $FC_{MATS++}(PNPSF3, l)$ при формировании начальных значений ячеек ЗУ случайным образом. Как видно из данных, приведенных в табл. 4, использование управляемых вероятностных значений, соответствующих тестам $CRT(4, 4, 6)$ и $CRT(8, 3, 6)$, позволяет достичь существенно большей полноты покрытия в сравнении с вероятностными значениями, сформированными случайным образом. Кроме того, в данном случае тест $CRT(8, 3, 6)$, полученный в результате двухмерного масштабирования для $n = 2$, превосходит по эффективности тест $CRT(4, 4, 6)$, являющийся результатом одномерного масштабирования исходного шаблона $CRT(4, 2, 3)$.

При проведении данного эксперимента тест $CRT(4, 4, 6)$ был построен с применением простейшего правила повторения, поэтому значения полноты покрытия для обоих случаев $CRT(4, 4, 6)$ и первых четырех наборов $CRT(8, 3, 6)$ должны быть одинаковыми (табл. 4). Применение других правил при реализации одномерного масштабирования исходных шаблонов позволяет ожидать неповторяющихся значений для $CRT(4, 4, 6)$ и $CRT(8, 3, 6)$, однако эти значения будут близкими по своим величинам. В общем случае при использовании шаблона $CRT(q, d, r)$ с $d > r/2$ эффективность применения первых q наборов теста $CRT(n \cdot q, (n \cdot r)/2, n \cdot r)$ будет сравнима с эффективностью теста $CRT(q, n \cdot d, n \cdot r)$. Отметим, что $CRT(n \cdot q, (n \cdot r)/2, n \cdot r)$ является результатом двухмерного масштабирования шаблона $CRT(q, d, r)$, а $CRT(q, n \cdot d, n \cdot r)$ – одномерного шаблона с применением различных правил, рассмотренных в работе [25].

Табл. 5 содержит данные полноты покрытия для управляемых вероятностных тестов, сформированных в результате двухмерного масштабирования ранее использованного шаблона $CRT(4, 2, 3)$ для $n = 4$ и 8. Для обоих тестов приводятся значения суммарной полноты покрытия при реализации l -кратного, $l \in \{1, 2, \dots, 11\}$, тестирования ЗУ соответствующей емкости.

Таблица 5

Полнота покрытия $FC_{MATS++}(PNPSF3, l)$ для ЗУ емкостью 12 и 24 бит, %

Table 5

The faults coverage $FC_{MATS++}(PNPSF3, l)$ for 12 and 24-bit memory, %

T_i	T_0	T_1	T_2	T_3	T_4	T_5	T_6	T_7	T_8	T_9	T_{10}
$CRT(16, 6, 12)$	25	47,73	68,18	86,36	90,91	93,94	95,45	95,45	97,73	99,24	100
$CRT(32, 12, 24)$	25	47,36	67,25	84,72	89,28	91,98	93,43	93,43	95,60	96,99	97,72
Random	25	43,75	57,81	68,35	76,28	82,20	86,65	89,98	92,49	94,36	95,77

Для сравнительного анализа эффективности управляемых вероятностных тестов, построенных с применением матриц Адамара, с классическим вероятностным тестированием в табл. 5 приводятся значения полноты покрытия для обоих случаев. Как видно из приведенных данных, предлагаемые авторами управляемые вероятностные тесты имеют заметно большую полноту покрытия неисправностей PNPSF3.

В табл. 6 приведены значения полноты покрытия $FC_{MATS++}(PNPSF5, l)$, аналогичные данным табл. 4 и 5, но для более сложных кодочувствительных неисправностей PNPSF5, в которых участвуют пять произвольных ячеек ЗУ при неисправном его поведении.

Таблица 6

Полнота покрытия $FC_{MATS++}(PNPSF5, l)$ для ЗУ емкостью 6, 12 и 24 бит, %

Table 6

The faults coverage $FC_{MATS++}(PNPSF5, l)$ for 6, 12 and 24-bit memory, %

T_i	T_0	T_1	T_2	T_3	T_4	T_5	T_6	T_7	T_8	T_9	T_{10}
$CRT(8, 3, 6)$	6,25	12,50	18,75	25,00	31,25	37,50	43,75	50,00	—	—	—
$CRT(16, 6, 12)$	6,25	12,49	18,71	24,92	30,42	35,90	41,36	46,82	51,55	56,28	60,98
$CRT(32, 12, 24)$	6,25	12,47	18,67	24,88	29,95	34,92	39,90	44,93	49,01	52,97	57,00
Random	6,25	12,11	17,60	22,75	27,58	32,11	36,35	40,33	44,06	47,55	50,83

Результаты, приведенные в табл. 6, подтверждают высокую эффективность управляемых вероятностных тестов, сформированных путем масштабирования исходных шаблонов, в сравнении с вероятностными тестами.

Эксперимент 3. В разд. 3 рассмотрено влияние свойства выбранного шаблона, выраженного в виде его среднего расстояния Хэмминга $aveHD_i(T_i, T_j)$, на аналогичную характеристику результирующего управляемого вероятностного теста (5). Было показано (см. табл. 2), что приведенное значение $aveHD_i(T_i, T_j)/(n \cdot r)$ среднего расстояния Хэмминга $aveHD_i(T_i, T_j)$ к одному разряду $(n \cdot r)$ -разрядного тестового набора, полученного в результате масштабирования с помощью матриц Адамара, достаточно быстро сводится с ростом величины n к значению 0,5. Это свидетельствует о том, что эффективность результирующего теста даже для небольших порядков n матриц Адамара, используемых для масштабирования, практически не зависит от свойств используемых шаблонов. Для трех ранее рассмотренных шаблонов $CRT(4, 2, 3)$, $CRT(4, 1, 3)$ и $CRT(4, 0, 3)$, результат масштабирования которых для $n = 2$ показан на рис. 4, была оценена их эффективность при обнаружении неисправностей PNPSF k для $k = 3$ и 5 маршевым тестом MATS++.

Таблица 7

Оценка полноты покрытия неисправностей PNPSF3 и PNPSF5, %

Table 7

Fault Coverage Assessment for PNPSF3 and PNPSF5, %

n	2		4		8		16		32	
k	3	5	3	5	3	5	3	5	3	5
$CRT(4, 2, 3)$	100	50,00	100	81,52	100	91,99	97,6	96,12	99,67	98,19
$CRT(4, 1, 3)$	83,33	46,67	91,67	69,55	95,65	84,32	95,62	91,84	98,62	95,94
$CRT(4, 0, 3)$	45,00	12,50	69,32	23,86	83,68	42,17	89,73	65,50	95,57	81,09

Аналогично, как и в предыдущих экспериментах, оценивалась обнаруживающая способность многократных маршевых тестов путем моделирования состоятельного множества рассматриваемых неисправностей. Соответственно, для $n = 2$ и 4 генерировалось все множество указанных неисправностей, а для остальных значений n генерировались те же виды неисправностей PNPSF3 и PNPSF5, для каждой из которых случайным образом формировалось по 1000 неисправностей. На численные значения полноты покрытия, приведенные в табл. 7, влияет как соотношение размерности ЗУ к количеству тестовых наборов, так и специфика используемых шаблонов. Это в первую очередь касается шаблона $CRT(4, 2, 3)$, который характеризуется максимальным значением $d = 2$. Из табл. 7 видно, что уже при $n = 32$ эффективность управляемого вероятностного теста как результата двухмерного масштабирования с применением матриц Адамара практически не зависит от свойств масштабируемого шаблона. Действительно, результат масштабирования весьма специфичного шаблона $CRT(4, 0, 3)$, состоящего только из повторяющихся нулевых значений, показывает неплохой результат в сравнении с шаблонами $CRT(4, 2, 3)$ и $CRT(4, 1, 3)$. Полнота покрытия неисправностей PNPSF3 и PNPSF5 тестом MATS++ при использовании указанных управляемых вероятностных тестов при $n = 32$ достигает максимально возможных значений.

Заключение. Рассмотрен подход к генерированию тестовых наборов при формировании управляемых вероятностных тестов с использованием матриц Адамара. Основой предложенного авторами метода является двухмерное масштабирование исходных шаблонов с применением указанных матриц. Показано, что использование различных шаблонов и их двухмерное масштабирование позволяют строить управляемые вероятностные тесты с требуемой разрядностью тестовых наборов и большим их количеством. Управляемые вероятностные тесты формируются без необходимости перечисления кандидатов в тестовые наборы, что сводит задачу синтеза управляемого вероятностного теста к формальной процедуре, основанной на операции повторения. Предложенная процедура масштабирования в отличие от классического подхода исключает необходимость вычисления меры (мер) различия и не требует вычислительных затрат. В сравнении с подходом, основанным на одномерном масштабировании, результирующий управляемый вероятностный тест в случае применения матриц Адамара порядка n содержит в n раз больше тестовых наборов. Количество наборов предлагаемых управляемых вероятностных тестов, а также их свойства, определяемые свойствами матриц Адамара, обеспечивают высокую эффективность подобных тестов. Для целей оценки эффективности предложенных решений были проведены экспериментальные исследования применимости рассмотренных тестов для целей тестирования ЗУ. Показаны их преимущества по отношению к вероятностным тестам при обнаружении кодочувствительных неисправностей ЗУ.

Вклад авторов. В. Н. Ярмолик предложил процедуру построения управляемых вероятностных тестов, основанную на применении двухмерного масштабирования исходных шаблонов. Н. А. Шевченко рассмотрел использование матриц Адамара для целей масштабирования и принял участие в анализе полученных результатов. В. В. Петровская провела большой объем экспериментальных исследований, приняла участие в обобщении, анализе и оформлении полученных результатов.

Список использованных источников

1. An orchestrated survey on automated software test case generation / S. Anand, E. Burke, T. Chen [et al.] // Journal of Systems and Software. – 2014. – Vol. C-39, no. 4. – P. 582–586.
2. Ярмолик, В. Н. Контроль и диагностика вычислительных систем / В. Н. Ярмолик. – Минск : Бест-принт, 2019. – 387 с.
3. Karmoge, S. P. Testing of embedded system, an issues and challenges / S. P. Karmoge, A. R. Mahajan // International Journal of Enhanced Research in Science, Technology & Engineering. – 2015. – Vol. 4, no. 8. – P. 181–186.
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, 8–12 Mar. 2010. – Dresden, Germany, 2010. – P. 323–326.

5. Malaiya, Y. K. The coverage problem for random testing / Y. K. Malaiya, S. Yang // Proc. of the Intern. Test Conf., Philadelphia, PA, USA, 16–18 Oct. 1984. – Philadelphia, 1984. – P. 237–242.
6. 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.
7. Garousi, V. 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.
8. Saini, D. K. Software testing for embedded systems / D. K. Saini // International Journal of Computer Applications. – 2012. – Vol. 43, no. 17. – P. 1–6.
9. Karmore, S. P. Universal methodology for embedded system testing / S. P. Karmore, A. R. Mahajan // Proc. of the 8th Intern. Conf. on Computer Science & Education (ICCSE 2013), Colombo, Sri Lanka, 26–28 Apr. 2013. – Colombo, 2013. – P. 567–572.
10. 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.
11. Chen, T. Y. 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.
12. Yarmolik, S. V. Controlled random tests / S. V. Yarmolik, V. N. Yarmolik // Automation and Remote Control. – 2012. – Vol. 73, no. 10. – P. 1704–1714.
13. Mrozek, I. Antirandom test vectors for BIST in hardware/software systems / I. Mrozek, V. N. Yarmolik // Fundamenta Informaticae. – 2012. – Vol. 119, no. 2. – P. 163–185.
14. 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.
15. Wu, S. H. Antirandom testing: A distance-based approach / S. H. Wu, S. Jandhyala, Y. K. Malaiya, A. P. Jayasumana // Hindawi Publishing Corporation VLSI Design. – 2008. – Vol. 2008, Article ID 165709. – 9 p. – DOI: 10.1155/2008/165709.
16. 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.
17. Xu, S. Orderly random testing for both hardware and software / S. Xu // Proc. of the 2008 14th IEEE Pacific Rim Intern. Symp. on Dependable Computing, Washington, D.C., USA, 15–17 Dec. 2008. – Washington, 2008. – P. 160–167.
18. 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.
19. Mrozek, I. Optimal controlled random tests / I. Mrozek, V. Yarmolik // Proc. of Computer Information Systems and Industrial Management: 16th IFIP TC8 Intern. Conf., CISIM 2017, Białystok, Poland, 16–18 June 2017. – Białystok, 2017. – P. 27–38.
20. 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.
21. Peterson, W. W. Error-Correction Codes / W. W. Peterson, E. J. Weldon. – Cambridge, Massachusetts, London England : The MIT Press, 1972. – 560 p.
22. Ярмолик, В. Н. Модификации способов определения расстояния Хэмминга для их применения в качестве мер различия при генерировании управляемых вероятностных тестов / В. Н. Ярмолик, В. В. Петровская, Н. А. Шевченко // Информатика. – 2024. – Т. 21, № 2. – С. 54–72.
23. Ярмолик, С. В. Управляемое случайное тестирование / С. В. Ярмолик, В. Н. Ярмолик // Информатика. – 2011. – Т. 29, № 1. – С. 79–88.
24. Plotkin, M. Binary codes with specified minimum distance / M. Plotkin // IRE Transactions on Information Theory. – 1960. – Vol. 6, no. 4. – P. 445–450.
25. Управляемые вероятностные тесты с фиксированным минимальным значением расстояния Хэмминга / В. Н. Ярмолик, В. В. Петровская, Д. В. Деменковец, В. А. Леванцевич // Информатика. – 2025. – Т. 22, № 1. – С. 7–26.
26. MacWilliams, F. J. The Theory of Error-Correcting Codes / F. J. MacWilliams, N. J. A. Sloane. – Amsterdam, The Netherland : Elsevier-North-Holland Pub. Co., 1977. – 762 p.
27. Hedayat, A. Hadamard matrices and their applications / A. Hedayat, W. D. Wallis // Annals of Statistics. – 1978. – Vol. 6, no. 6. – P. 1184–1238.
28. Seberry, J. Hadamard matrices: Constructions using number theory and linear algebra / J. Seberry, M. Yamada. – Hoboken, NJ, USA : John Wiley & Sons, Inc., 2020. – 384 p.
29. Mrozek, I. Multiple controlled random testing / I. Mrozek, V. Yarmolik // Fundamenta Informaticae. – 2016. – Vol. 144, no. 1. – P. 23–43.

30. 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.
31. Cheng, K.-L. Neighborhood pattern-sensitive fault testing and diagnostics for random-access memories / K.-L. Cheng, M.-F. Tsai, C. T. Wu // IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems. – 2002. – Vol. 21, no. 11. – P. 284–267.
32. Parvathi, M. Novel test methods for NPSF faults in SRAM / M. Parvathi, T. Hmiasree, T. Bhavyasree // Proc. of Intern. Conf. on Computational and Characterization Techniques in Engineering & Sciences (CCTES), Lucknow, India, 14–15 Sept. 2018. – Lucknow, 2018. – P. 112–118.
33. Ярмолик, В. Н. Построение и применение маршевых тестов для обнаружения кодочувствительных неисправностей запоминающих устройств / В. Н. Ярмолик, В. А. Леванцевич, Д. В. Деменковец, И. Мрозек // Информатика. – 2021. – Т. 18, № 1. – С. 25–42.

References

1. 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.
2. Yarmolik V. N. Control' i diagnostika vuchislitel'nuch system. *Computer Systems Testing and Diagnoses*. Minsk, Bestprint, 2019, 387 p. (In Russ.).
3. 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.
4. Krupp A., Muller W. A Systematic approach to the test of combined HW/SW systems. *Proceeding of the IEEE Conference on the Testing and Automation of Embedded Systems (DATE 2010)*, Dresden, Germany, 8–12 March 2010. Dresden, 2010, pp. 323–326.
5. Malaiya Y. K., Yang S. The coverage problem for random testing. *Proceeding of the International Test Conference, Philadelphia, PA, USA, 16–18 October 1984*. Philadelphia, 1984, pp. 237–242.
6. Arcuri A., Iqbal M. Z., Briand L. Random testing: Theoretical results and practical implications. *IEEE Transactions on Software Engineering*, 2011, vol. 38, no. 2, pp. 258–277.
7. 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.
8. Saini D. K. Software testing for embedded systems. *International Journal of Computer Applications*, 2012, vol. 43, no. 17, pp. 1–6.
9. Karmore S. P., Mahajan A. P. Universal methodology for embedded system testing. *Proceeding of the 8th International Conference on Computer Science & Education (ICCSE 2013)*, Colombo, Sri Lanka, 26–28 April 2013. Colombo, 2013, pp. 567–572.
10. 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.
11. 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.
12. Yarmolik S. V., Yarmolik V. N. Controlled random tests. *Automation and Remote Control*, 2012, vol. 73, no. 10, pp. 1704–1714.
13. Mrozek I., Yarmolik V. Antirandom test vectors for BIST in hardware/software systems. *Fundamenta Informaticae*, 2012, vol. 119, no. 2, pp. 163–185.
14. 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.
15. 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, article ID 165709, 9 p. DOI: 10.1155/2008/165709.
16. 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.
17. Xu S. Orderly random testing for both hardware and software. *Proceedings of the 2008 14th IEEE Pacific Rim International Symposium on Dependable Computing*, Washington, D.C., USA, 15–17 December 2008. Washington, 2008, pp. 160–167.
18. 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.

19. Mrozek I., Yarmolik V. Optimal controlled random tests. *Proceedings of Computer Information Systems and Industrial Management: 16th IFIP TC8 International Conference, CISIM 2017, Białystok, Poland, 16–18 June 2017*. Białystok, 2017, pp. 27–38.
20. Hamming R. W. Error detecting and error correcting codes. *The Bell System Technical Journal*, 1950, vol. 29, no. 2, pp. 147–160.
21. Peterson W. W., Weldon E. J. *Error-Correction Codes*. Cambridge, Massachusetts, London, England, The MIT Press, 1972, 560 p.
22. Yarmolik V. N., Petrovskaya V. V., Shevchenko N. A. *Dissimilarity measures based on the application of Hamming distance to generate controlled probabilistic tests*. Informatika [Informatics], 2024, vol. 21, no. 2, pp. 54–72 (In Russ.).
23. Yarmolik S. V., Yarmolik V. N. *Controlled random testing*. Informatika [Informatics], 2011, vol. 29, no. 1, pp. 79–88 (In Russ.).
24. Plotkin M. Binary codes with specified minimum distance. *IRE Transactions on Information Theory*, 1960, vol. 6, no. 4, pp. 445–450.
25. Yarmolik V. N., Petrovskaya V. V., Demenkovets D. V., Levantsevich V. A. *Controlled random tests with fixed minimal Hamming distance*. Informatika [Informatics], 2025, vol. 22, no. 1, pp. 7–26 (In Russ.).
26. MacWilliams F. J., Sloane N. J. A. *The Theory of Error-Correcting Codes*. Amsterdam, The Netherland, Elsevier-North-Holland Pub. Co., 1977, 762 p.
27. Hedayat A., Wallis W. D. Hadamard matrices and their applications. *Annals of Statistics*, 1978, vol. 6, no. 6, pp. 1184–1238.
28. Seberry J., Yamada M. *Hadamard Matrices: Constructions using Number Theory and Linear Algebra*. Hoboken, NJ, USA, John Wiley & Sons, Inc., 2020, 384 p.
29. Mrozek I., Yarmolik V. Multiple controlled random testing. *Fundamenta Informaticae*, 2016, vol. 144, no. 1, pp. 23–43.
30. 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.
31. Cheng K.-L., Tsai M.-F., Wu C. T. Neighborhood pattern-sensitive fault testing and diagnostics for random-access memories. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, 2002, vol. 21, no. 11, pp. 284–267.
32. Parvathi M., Hmiasree, T., Bhavyasree T. Novel test methods for NPSF faults in SRAM. *Proceedings of the International Conference on Computational and Characterization Techniques in Engineering & Sciences (CCTES), Lucknow, India, 14–15 September 2018*. Lucknow, 2018, pp. 112–118.
33. Yarmolik V. N., Levantsevich V. A., Demenkovets D. V., Mrozek I. Construction and application of march tests for pattern sensitive memory faults detection. Informatika [Informatics], 2021, vol. 18, no. 1, pp. 25–42 (In Russ.).

Информация об авторах

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

Шевченко Николай Алексеевич, студент, Дармштадтский технический университет.
E-mail: nik.sh.de@gmail.com

Петровская Вита Владленовна, магистр технических наук, Белорусский государственный университет информатики и радиоэлектроники.
E-mail: vita.petrovskaya@gmail.com

Information about the authors

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

Mikalai A. Shauchenka, Student, Darmstadt Technical University.
E-mail: nik.sh.de@gmail.com

Vita V. Petrovskaya, M. Sc. (Eng.), Belarusian State University of Informatics and Radioelectronics.
E-mail: vita.petrovskaya@gmail.com

БИОИНФОРМАТИКА

BIOINFORMATICS



УДК 57.087.1
DOI: 10.37661/1816-0301-2025-22-2-81-94

Оригинальная статья
Original Article

Программный комплекс для имитационного моделирования сайтов однонуклеотидного генетического полиморфизма

Н. Н. Яцков[✉], Д. Д. Сарнацкий, В. В. Скаун, В. В. Гринев

Белорусский государственный университет,
пр. Независимости, 4, Минск, 220030, Беларусь
[✉]E-mail: yatskou@bsu.by

Аннотация

Цели. В настоящее время высокопроизводительные методы секвенирования широко используются в фундаментальных и прикладных исследованиях различных заболеваний человека. Секвенирование функционально значимых регионов генома человека позволяет одновременно идентифицировать множество сайтов генетического полиморфизма, имеющих диагностическую и (или) прогностическую значимость в отношении генетических заболеваний человека. В числе приоритетных целей в этой области стоит разработка эффективных программных инструментов обработки геномных данных и идентификации сайтов однонуклеотидного полиморфизма с использованием методов компьютерного моделирования и анализа больших данных.

Методы. Разработан программный комплекс для имитационного моделирования и идентификации сайтов однонуклеотидного полиморфизма с использованием методов машинного обучения. Реализована методика подхода имитационного моделирования и анализа сайтов однонуклеотидного полиморфизма в молекулах ДНК на основе бета-распределения или нормального закона распределения, параметры которых определяются по имеющимся экспериментальным данным, и методов интеллектуального анализа, обученных на смоделированных данных и применяемых для точной идентификации сайтов однонуклеотидного полиморфизма. Комплекс включает R-пакет, веб-приложение и вспомогательные программные средства для обработки экспериментальных данных геномного секвенирования.

Результаты. Проверка работоспособности представленного программного комплекса проведена на наборах смоделированных и экспериментальных данных геномного секвенирования клеток человека. Выполнен сравнительный анализ наиболее эффективных алгоритмов идентификации сайтов однонуклеотидных полиморфизмов. Наилучшие результаты получены для моделей машинного обучения.

Заключение. Применение программного комплекса повышает точность определения сайтов генетического полиморфизма в ходе анализа больших данных геномного секвенирования. Комплекс может использоваться для моделирования синтетических данных по экспериментальным данным или самостоятельно с целью всестороннего тестирования и выбора наилучших алгоритмов идентификации однонуклеотидных полиморфизмов, а также для генеративного моделирования данных, используемых при обучении алгоритмов идентификации на основе методов интеллектуального анализа.

Ключевые слова: однонуклеотидный генетический полиморфизм, программный комплекс, имитационное моделирование, машинное обучение, интеллектуальный анализ данных, R-пакет, веб-приложение

Благодарности. Работа выполнена в рамках государственной программы научных исследований «Конвергенция-2025» (грант № 3.04.3.1, № гос. регистрации 20211918).

Для цитирования. Программный комплекс для имитационного моделирования сайтов однонуклеотидного генетического полиморфизма / Н. Н. Яцков, Д. Д. Сарнацкий, В. В. Скакун, В. В. Гринев // Информатика. – 2025. – Т. 22. – № 2. – С. 81–94. – DOI: 10.37661/1816-0301-2025-22-2-81-94.

Конфликт интересов. Авторы заявляют об отсутствии конфликта интересов.

Поступила в редакцию | Received 11.04.2025

Подписана в печать | Accepted 05.05.2025

Опубликована | Published 30.06.2025

Software complex for simulation modelling of single nucleotide genetic polymorphism sites

Mikalai M. Yatskou[✉], Dzianis D. Sarnatski, Victor V. Skakun, Vasily V. Grinev

Belarusian State University,

av. Nezavisimosti, 4, Minsk, 220030, Belarus

[✉]E-mail: yatskou@bsu.by

Abstract

Objectives. High-throughput sequencing methods have recently become widely used in the fundamental and applied research of various human diseases. Sequencing of functionally significant regions of the human genome enables the simultaneous identification of multiple genetic polymorphism sites that have diagnostic and/or prognostic significance for human genetic diseases. One of the key goals in this area is to develop efficient software tools for processing genomic data and identifying single nucleotide polymorphism sites using computer modelling and big data analysis methods.

Methods. A software complex has been developed for simulation modelling and identification of single nucleotide polymorphism sites using machine learning methods. The methods for the approach to simulation modelling and analysis of single nucleotide polymorphism sites in DNA molecules are implemented based on the beta or normal distributions, the parameters of which are determined from the available experimental data, and machine learning models trained on simulated data and used to accurately identify single nucleotide polymorphism sites. The software complex includes an R package, a web application, and auxiliary computational tools for processing experimental genomic sequencing data.

Results. The performance of the developed software complex was tested on sets of simulated and experimental data from human cell genomic sequencing. A comparative analysis of the most effective algorithms for identifying single nucleotide polymorphism sites was performed. The best results were obtained for machine learning models.

Conclusion. The use of the software complex increases the accuracy of identifying genetic polymorphism sites during the analysis of big genomic sequencing data. The software can be used for modelling synthetic data, based on experimental data or independently, for the purpose of comprehensive testing and selection of the best algorithms for identifying single nucleotide polymorphisms, as well as for generative data modelling used in training identification algorithms based on machine learning methods.

Keywords: single nucleotide genetic polymorphism, software complex, simulation modelling, machine learning, data mining, R package, web application

Acknowledgments. This work was carried out in the framework of the state programme of scientific research "Convergence-2025" (grant No. 3.04.3.1, state registration No. 20211918).

For citation. Yatskou M. M., Sarnatski D. D., Skakun V. V., Grinev V. V. *Software complex for simulation modelling of single nucleotide genetic polymorphism sites*. Informatika [Informatics], 2025, vol. 22, no. 2, pp. 81–94 (In Russ.). DOI: 10.37661/1816-0301-2025-22-2-81-94.

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

Введение. Генетический полиморфизм влияет на фенотип человека и других живых организмов [1]. Одним из наиболее распространенных типов генетических вариаций в геноме человека является однонуклеотидный полиморфизм (от англ. single nucleotide polymorphism), когда индивидуальные геномы отличаются по отдельным нуклеотидам. При этом идентификация таких мест (сайтов) в геноме, по которым наблюдаются однонуклеотидные различия, имеет не только фундаментальное, но и сугубо прикладное значение, в частности для диагностики и прогнозирования течения онкологических заболеваний человека [2].

Среди существующих способов определения сайтов однонуклеотидного полиморфизма (СОП) следует отметить статистические методы точного теста Фишера, биномиального распределения, на основе энтропии и машинного обучения [1, 3, 4]. Методы достаточно универсальны и просты для программной реализации, однако вычислительно затратны и трудно применимы при анализе экспериментальных данных с высоким уровнем шума и различными экспериментальными искажениями, являющимися источниками пропусков, повторов и прочих аномальных явлений [5]. Поэтому для выбора наиболее эффективного алгоритма идентификации сайтов, проверки конкурирующих методик анализа и оценки производительности конкретных экспериментальных планов исследования биомолекулярных систем зачастую используется имитационное моделирование [6, 7].

Так, имитационное моделирование применяется при генерации синтетических данных для методов машинного обучения (или интеллектуального анализа данных) с целью прямой идентификации СОП различных организмов [4]. В этом случае формирование смоделированных обучающих данных может иметь преимущества по точности и эффективности при анализе экспериментальных данных как с невысоким числом покрытий, так и с наличием пропусков, обусловленных экспериментальными искажениями. Существующие подходы к математическому моделированию СОП основаны на учете параметров экспериментального оборудования, применения вероятностных моделей и статистических подходов, использовании вспомогательной биологической информации [8, 9]. Однако из-за различий в типах генетических данных, методах моделирования, эволюционных характеристиках, форматах данных, терминологии и допущениях, принятых в существующих программных приложениях, выбор надежного инструмента для конкретного исследования является весьма сложным и субъективным процессом [10]. Следует отметить, что немногие методы моделирования используют экспериментальные результаты или наблюдаемые характеристики для моделирования, включающие сложные схемы моделирования с учетом структуры и особенностей экспериментального шума.

На данный момент разработаны различные подходы и программные средства к математическому моделированию сайтов генетического полиморфизма [6, 7, 11–15]. Основными ограничениями существующих программных решений являются: отсутствие вычислительных процедур анализа и моделирования распределений измеряемых характеристик нуклеотидных сайтов, учитывающих особенности конкретных экспериментов; необходимость дополнительной реализации процедур генерации выборок нуклеотидных последовательностей для обучения классификационных моделей с целью последующей идентификации однонуклеотидных полиморфизмов; отсутствие специализированного R-пакета, интегрирующего основные алгоритмы обработки экспериментальных данных, моделирования и идентификации СОП, для широкого использования биоинформатическим сообществом. Следовательно, разработка программной платформы имитационного моделирования и интеллектуального анализа генетических полиморфизмов по данным геномного секвенирования по-прежнему является актуальной задачей.

Целью настоящей работы является создание программного комплекса для имитационного моделирования и идентификации СОП с использованием моделей машинного обучения, параметры которых оцениваются на базе имеющихся экспериментальных данных. Комплекс включает R-пакет, веб-приложение и вспомогательные программные средства для обработки экспериментальных данных геномного секвенирования. Он позволяет смоделировать особенности конкретных экспериментов и сформировать эталонные выборки для обучения классификационных моделей, что обязательно в алгоритмах машинного обучения (интеллектуального анализа). Работоспособность разработанных программных средств подтверждена в ходе сравнитель-

ного анализа наиболее эффективных алгоритмов идентификации СОП на примерах смоделированных и экспериментальных данных секвенирования генома человека.

Программный комплекс. В программном комплексе реализована методика подхода имитационного моделирования и анализа СОП в молекулах ДНК на основе законов распределений, параметры которых определяются по имеющимся экспериментальным данным, и методов интеллектуального анализа, обученных на смоделированных данных и применяемых для точной идентификации СОП [16]. Схема методики подхода представлена на рис. 1, а и включает пять этапов: построения и аппроксимации гистограмм числа нуклеотидных покрытий с помощью выбранных функций законов распределений; имитационного моделирования нуклеотидных покрытий в каналах регистрации нуклеотидных прочтений сайтов; генерации информативных признаков сайтов, характеризующих покрытия нуклеотидов А, С, Г и Т; обучения моделей идентификации (интеллектуального анализа); идентификации СОП в смоделированных или экспериментальных данных с использованием настроенных моделей машинного обучения.

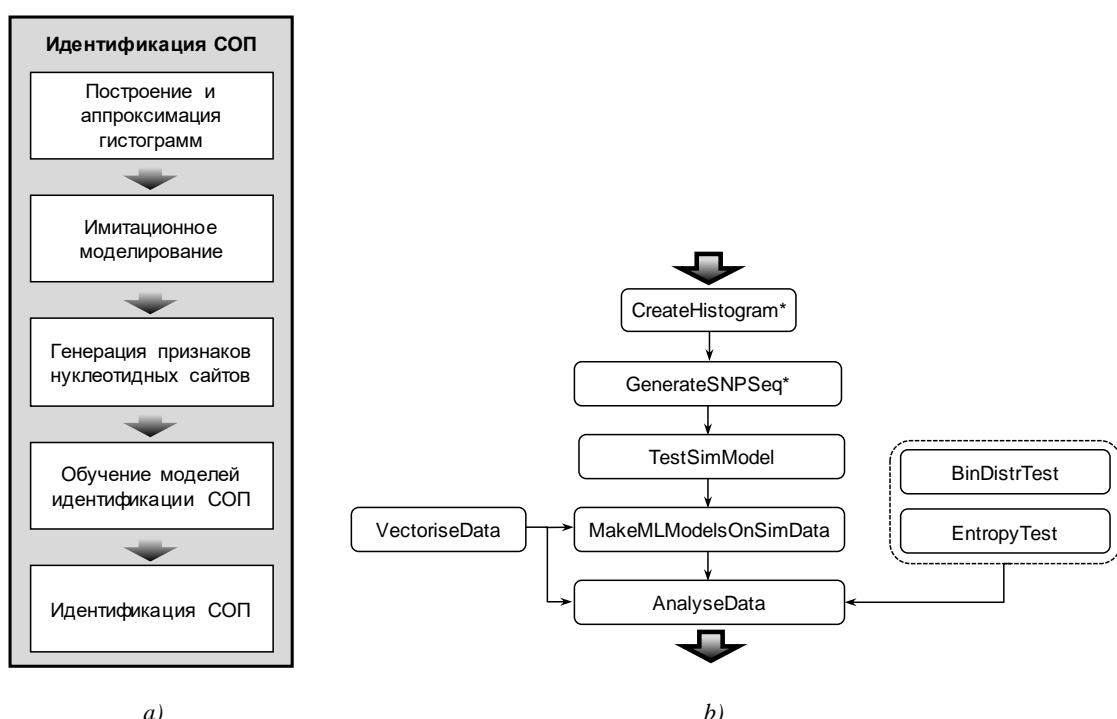


Рис. 1. Методика подхода имитационного моделирования и идентификации СОП нуклеотидных последовательностей (а) и схема организации взаимодействия функций в R-пакете (б).

*CreateHistogram** – обозначение функций *CreateHistogram1*, *CreateHistogram2*, *CreateHistogram3* и *CreateHistogramSNVs*; *GenerateSNPSeq** – обозначение функций *GenerateSNPSeqBeta*, *GenerateSNPSeqBetaGaussNoise*, *GenerateSNPSeqGauss* и *GenerateSNPSeqGaussGaussNoise*

Fig. 1. The methods of the approach to simulation modelling and identification of single nucleotide polymorphism (SNP) in nucleotide sequences (a) and the functional scheme in the R package (b).

*CreateHistogram** is the designation of the functions *CreateHistogram1*, *CreateHistogram2*, *CreateHistogram3* and *CreateHistogramSNVs*; *GenerateSNPSeq** is the designation of the functions *GenerateSNPSeqBeta*, *GenerateSNPSeqBetaGaussNoise*, *GenerateSNPSeqGauss* and *GenerateSNPSeqGaussGaussNoise*

В ходе работы выполняется анализ экспериментальных характеристик выбранного набора данных геномного секвенирования с целью определения закона распределения и оценки неизвестных параметров. Для аппроксимации чисел покрытий нуклеотидов сайтов предлагается использовать бета- или нормальное распределения. Результаты исследований позволяют сделать вывод о том, что бета-распределение удовлетворительно подходит для анализируемых интегральных характеристик экспериментального набора сайтов в геноме человека [17]. Нор-

мальное распределение – менее точное при исследовании генома человека [17], однако применимо к экспериментальным характеристикам, демонстрирующим гауссовость данных. Следует отметить возможность применения других типов вероятностных распределений или функций полиномиального сглаживания. Полученные экспериментальные оценки параметров распределений используются в имитационных моделях. Имитационное моделирование включает случайную генерацию позиций СОП в последовательности рассматриваемой молекулы, состоящей из фиксированного числа нуклеотидных сайтов, для каждого из которых воспроизводятся числа покрытий нуклеотидов по бета-распределению или нормальному распределению в заданном диапазоне. Экспериментальные гистограммы рассматриваются в качестве законов распределений. Для нуклеотидного сайта моделируется общее число покрытий четырех нуклеотидов, по которому затем генерируются числа покрытий для эталонного и альтернативного нуклеотидов. Аналогично моделируются покрытия для СОП. Принимается допущение о том, что каждый сайт может быть представлен не более чем двумя вариантами покрытий нуклеотидов – эталонным или альтернативным. Предложенный алгоритм позволяет воспроизводить наборы данных, максимально приближенные к эмпирическим условиям, задаваемым числами покрытий и законами их распределений, а также количеством полиморфных сайтов. Особенностями подхода являются применение алгоритмов имитационного моделирования для воспроизведения регистрируемых характеристик конкретных экспериментов и генерация «реалистичных» выборок для обучения классификационных моделей.

Для реализации программного обеспечения используются различные вычислительные платформы и технологии программирования [18]. Основным преимуществом среди статистического программирования R является возможность использования огромного набора биоинформационных алгоритмов, алгоритмов интеллектуального анализа данных и разнообразных статистических вычислительных ресурсов научного сообщества¹ [19]. Для разработки программного комплекса в работе выбрана вычислительная среда R.

Отдельное направление в разработке R-приложений связано с созданием «реактивных» веб-интерфейсов с помощью пакета Shiny и размещением программной реализации на ресурсе shinyapps.io, предоставляемом разработчиками открытого программного обеспечения RStudio². Достоинством данного подхода является возможность удаленной работы с приложением широкой научной аудитории пользователей в режиме онлайн через сеть Интернет. Для разработки веб-интерфейса приложения выбран пакет Shiny.

Программный комплекс включает R-пакет, веб-приложение и набор вспомогательных функций.

R-пакет. Разработанные R-функции, реализующие различные этапы анализа данных и построения имитационных моделей, интегрированы в отдельный R-пакет [20]. Назначение пакета – моделирование синтетических данных по экспериментальным данным или самостоятельно с целью всестороннего тестирования и выбора наилучших алгоритмов идентификации СОП, а также генеративного моделирования данных для обучения алгоритмов идентификации на основе методов машинного обучения. Функциональная схема и описание функций пакета представлены на рис. 1, *b* и в табл. 1.

Пакет включает функции: *CreateHistogram** – построения гистограмм чисел покрытий нуклеотидных сайтов, их аппроксимации с помощью бета- и нормального распределений (R-функции *dbeta* и *dnorm*) и оценки параметров распределений (R-функция *nls*); *GenerateSNPSeq** – моделирования нуклеотидных покрытий по бета- илициальному распределениям с добавлением аддитивного гауссова шума; *TestSimModel* – проверки разработанных имитационных моделей; *MakeMLModelsOnSimData* – обучения моделей интеллектуального анализа на смоделированных данных; *AnalyseData* – идентификации СОП с использованием предобученных моделей или классических методов тестов биномиального распределения и на основе вычисления

¹R Core Team. R: A language and Environment for Statistical Computing / R Foundation for Statistical Computing. – 2024. – URL: <http://www.R-project.org/> (date of access: 10.04.2025).

²Posit team. RStudio: Integrated Development Environment for R. Posit Software, PBC, Boston, MA. – 2024. – URL: <http://www.posit.co/> (date of access: 10.04.2025).

энтропии; *VectoriseData* – формирования четырех информативных признаков нуклеотидных сайтов для применения алгоритмов машинного обучения: X_1 – число покрытий эталонного нуклеотида, X_2 – X_4 – отсортированные в порядке убывания числа покрытий для альтернативных нуклеотидов (данные нормируются к общему числу покрытий сайта n). В качестве методов машинного обучения выбраны алгоритмы на основе деревьев условного вывода (функция *cTree* пакета *party*), классификации и регрессии построением дерева решений CART (функция *rpart* пакета *rpart*), машины опорных векторов (функция *svm* пакета *e1071*) и ансамблевого алгоритма XGBoost (функция *xgboost* пакета *xgboost*). В пакет включены модели *Classif_cTree_SimData.rds*, *Classif_rpart_SimData.rds*, *Classif_SVM_SimData.rds*, *Classif_xgboost_SimData.rds*, обученные на смоделированных данных по бета-закону распределения на выборке из 40 000 сайтов, из которых 20 000 являются СОП. Параметры моделей оценены на наборе данных геномного секвенирования молекул ДНК человека [21]. Эффективность алгоритмов оценивается с помощью мер точности *Precision*, чувствительности *Recall* и score F_1 , характеризующих свойства алгоритмов не включать ложноположительные события (*Precision*, неверно классифицированные сайты как СОП) и истинно положительные события (*Recall*, верно классифицированные как СОП), и их комбинированного вклада (F_1) [22].

Таблица 1
Функции R-пакета

Table 1
Functions of the R-package

R-функция <i>R function</i>	Описание <i>Description</i>	Результат <i>Value</i>
<i>AnalyseData</i>	Идентификация СОП в смоделированных или экспериментальных данных	Списки СОП
<i>BinRatioTest</i>	Программная реализация теста биномиального распределения	Список СОП, p -величины
<i>EntropyTest</i>	Программная реализация теста на основе энтропии	Список СОП, оценки энтропии E и p -величины
<i>CreateHistogram1</i> <i>CreateHistogram2</i> <i>CreateHistogram3</i> <i>CreateHistogramSNVs</i>	Построение гистограмм общего числа покрытий (*1), максимального числа покрытий (*2), разностей общего и максимального чисел покрытий (*3) нуклеотидных сайтов и СОП (*SNVs). Аппроксимация гистограмм производится с помощью бета- и нормального законов распределений, выполняется оценка параметров распределений	Оцененные параметры бета- и нормального распределений
<i>GenerateSNPSeqBeta</i> <i>GenerateSNPSeqBetaGaussNoise</i> <i>GenerateSNPSeqGauss</i> <i>GenerateSNPSeqGaussGaussNoise</i>	Имитационное моделирование сайтов и чисел их покрытий в нуклеотидной последовательности с использованием бета-распределения (*Beta), нормального распределения (*Gauss) и с добавлением гауссового шума (*BetaGaussNoise и *GaussGaussNoise)	Смоделированные данные, представленные символами референсного нуклеотида и числами покрытий сайта в каналах нуклеотидов A, C, G и T
<i>MakeMLModelsOnSimData</i>	Создание и обучение моделей идентификации СОП на смоделированных данных	Модели машинного обучения
<i>TestSimModel</i>	Построение гистограмм чисел покрытий нуклеотидных сайтов по смоделированным данным. Аппроксимация гистограмм с помощью бета- и нормального распределений. Оценка параметров распределений	Оцененные параметры бета- и нормального распределений
<i>VectoriseData</i>	Генерация информативных признаков сайтов, характеризующих покрытия нуклеотидов A, C, G и T, для обучения моделей классификации	Преобразованные к информативным признакам наборы данных

В ходе работы создан R-пакет SNPSimulatoR. Архив для установки – файл SNPSimulatoR_0.1.0.tar.gz – доступен по ссылке https://sstcenter.com/download/SNPSimulatoR_test_data. Примеры скриншотов подключения пакета SNPSimulatoR в RStudio показаны на рис. 2.

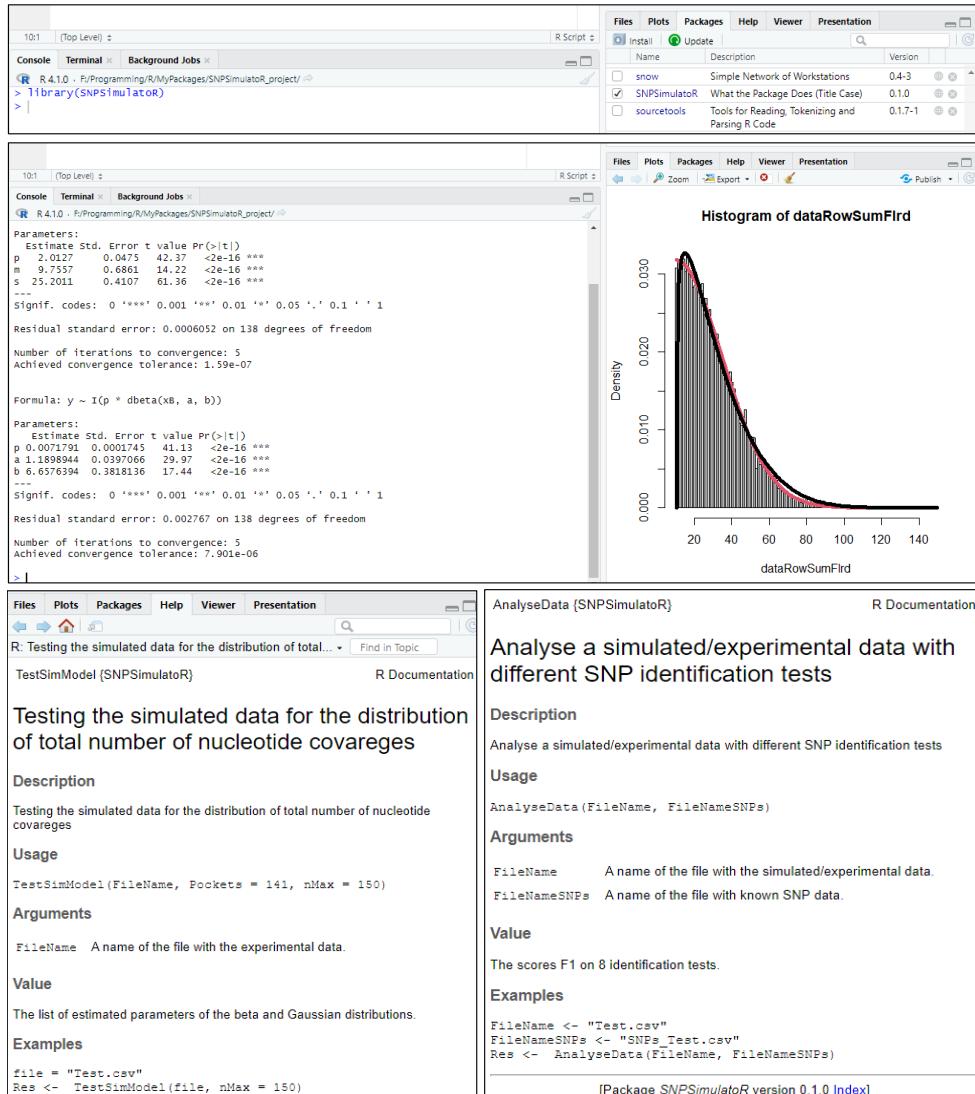


Рис. 2. Примеры подключения пакета, вывода результатов функции *CreateHistogram1* и представления информации о функциях *TestSimModel* и *AnalyseData*

Fig. 2. Examples of connecting the package, outputting the results of the function *CreateHistogram1*, and presenting information about the functions *TestSimModel* and *AnalyseData*

Веб-приложение. Разработанное веб-приложение SNPSimulatoR доступно по ссылке <https://dsa-cm.shinyapps.io/SNPSimulatoR>. Оно интегрирует реализованные алгоритмы вычислительного подхода. Примеры окон интерфейса веб-приложения представлены на рис. 3.

Главное окно интерфейса состоит из девяти панелей, соответствующих пяти этапам анализа: загрузки и моделирования данных; аппроксимации экспериментальных характеристик по бета- или нормальному распределениям; обучения моделей на генеративных данных; идентификации однонуклеотидных полиморфизмов с помощью моделей машинного обучения; визуализации и интерпретации результатов. На каждом этапе анализа пользователь должен установить системные параметры алгоритмов, при необходимости выбрать соответствующий файл предобученной модели идентификации СОП. Результаты анализа сохраняются в отдельный csv-файл. Приложение предназначено для анализа экспериментальных и смоделированных данных.

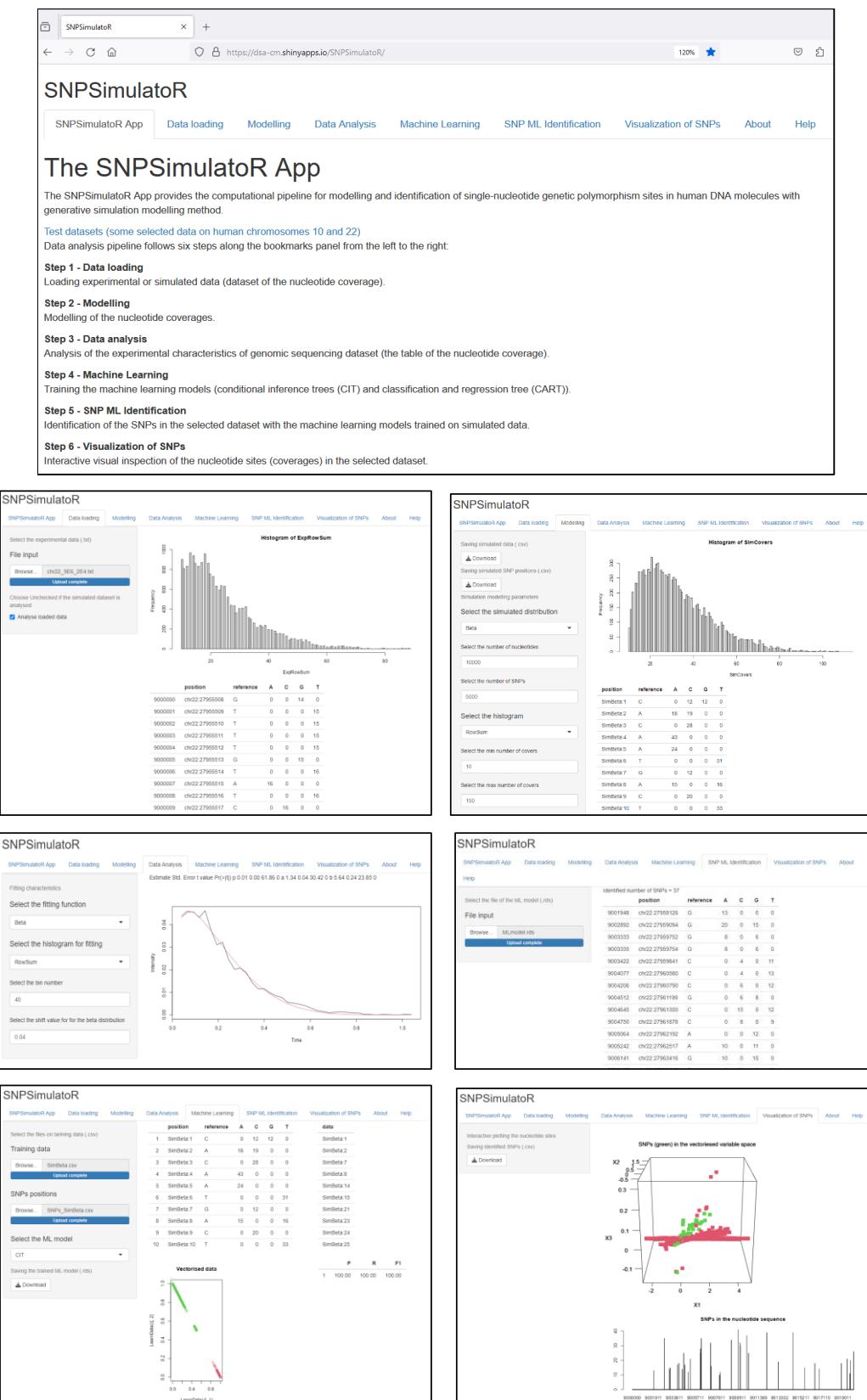


Рис. 3. Основные этапы исследования экспериментальных данных геномного секвенирования с использованием подхода на основе имитационного моделирования и интеллектуальной идентификации СОП

Fig. 3. Main stages of the study of experimental genomic sequencing data using the approach based on simulation modelling and intelligent identification of single nucleotide polymorphism (SNP)

Вспомогательные функции. Реализованы дополнительные функции для преобразования BAM(Binary Alignment Map)-файлов экспериментальных данных геномного секвенирования в частотные таблицы для непосредственной обработки в программном комплексе, фильтрации данных и сохранения результатов анализа в специализированном VCF(Variant Call Format)-формате.

Проверка работоспособности программного комплекса. Работоспособность разработанных средств программного комплекса подтверждена в ходе сравнительного анализа наиболее эффективных алгоритмов идентификации СОП на примерах смоделированных и экспериментальных данных геномного секвенирования [23, 24]. Для всестороннего исследования алгоритмов идентификации СОП смоделированы данные с учетом добавления гауссового шума с параметрами $\mu = 0$, $\sigma_l = q_l \cdot b_l$, $l = 1-4$ (индексы соответствуют нуклеотидам А, С, Г и Т), где $q_l > 0$, $\{b_1, b_2, b_3, b_4\}$ – набор из n прочтений нуклеотидных оснований А, С, Г или Т, покрывающих сайт j . Варьирование параметром σ_l изменяет уровень экспериментального шума, а именно регулирует информативность полезного сигнала, что позволяет всесторонне исследовать эффективность разрабатываемых или выбранных алгоритмов идентификации СОП и воссоздавать специальные экспериментальные условия. В качестве экспериментального набора рассмотрены эталонные данные геномного секвенирования молекул ДНК человека, предоставленные консорциумом GIAB (Genome in a Bottle Consortium) [21]. Выбор данных GIAB обусловлен тем, что на сегодняшний день это наиболее надежные бенчмарк-данные для решения задач, связанных с изучением геномного полиморфизма у человека (от разработки новых инструментальных методов «мокрой» биологии до сравнения алгоритмов обнаружения полиморфных сайтов). Набор данных содержит характеристики 29 633 768 нуклеотидных сайтов, из которых 36 150 являются истинными СОП. Фрагмент набора данных представлен в табл. 2. Проведено исследование наиболее эффективных алгоритмов идентификации сайтов: тестов на основе биномиального распределения, энтропии и модифицированного точного теста Фишера, базовых методов машинного обучения – деревьев условного вывода, классификации и регрессии построением дерева решений CART, машины опорных векторов с линейной разделяющей функцией, обученных на имитационно-смоделированных данных. Эффективность алгоритмов оценена с помощью меры точности F_1 .

Таблица 2
Фрагмент экспериментального набора данных

Table 2
Fragment of the experimental dataset

Хромосома : позиция <i>Chromosome : position</i>	Эталонный нуклеотид <i>Reference nucleotide</i>	Покрытие <i>Coverage</i>			
		A	C	G	T
chr22 : 47891620	T	0	0	0	27
chr22 : 47891621	G	0	0	28	0
chr22 : 47891622	T	0	0	0	30

Исследование смоделированных данных в условиях возрастающего шума позволяет сделать вывод о том, что для незашумленных данных предпочтительнее использовать алгоритмы машинного обучения, а в условиях существенного зашумления – статистический тест на основе энтропии и модель машинного обучения CART [23]. При исследовании экспериментальных данных точность идентификации СОП по мере F_1 на 2–5 % выше у метода на основе дерева решений условного вывода, чем у сравниваемых методов [24]. Прочие модели машинного обучения и классические тесты идентификации имеют сопоставимую точность. Следует отметить, что точность моделей идентификации СОП, обученных на экспериментальных данных, не превышала 60–70 %. Невысокая точность может быть обусловлена двумя факторами: ограниченностью обучающей экспериментальной выборки или тем, что имитационная модель действительно лучше формирует обучающие данные, фокусируясь на воспроизведении важных источников информации в данных, и не учитывает второстепенные сигналы, присутствующие в реальных данных.

Ограничения и дальнейшее развитие программного комплекса. Программный комплекс имеет ряд ограничений, таких как зависимость от сторонних R-библиотек, медленная инициализация, неоптимальная обработка больших файлов данных. Эти недостатки будут устранены путем построения специализированных библиотек C++, оптимизации и распараллеливания кодов промежуточной обработки данных и вычислительных алгоритмов для работы с большими данными [25], а также включения автоматического обновления моделей машинного обучения. В дальнейшем планируется разработка усовершенствованных имитационных моделей на основе дополнительных законов распределений (например, гамма-распределения и распределения Вейбулла) для воспроизведения чисел покрытий в четырех каналах регистрации прочтений нуклеотидных сайтов с учетом статистических взаимосвязей между каналами [26], способов генерации дополнительных информативных признаков сайтов нуклеотидных последовательностей [27]. Кроме того, будет существенно расширен список предобученных моделей идентификации СОП, в том числе включением моделей нейронных сетей глубокого обучения [28].

Заключение. Разработан программный комплекс для моделирования и анализа сайтов нуклеотидных последовательностей по экспериментальным наборам данных, основанный на генерации случайных событий бета-распределения или нормального закона распределения, параметры которых оцениваются по имеющимся экспериментальным данным, и методов интеллектуального анализа, обученных на смоделированных данных и применяемых для точной идентификации сайтов однонуклеотидных полиморфизмов. Комплекс имеет следующие преимущества в сравнении с существующими программными решениями: обеспечивает моделирование данных, близко воспроизводящих реальные экспериментальные условия, с целью изучения надежности конкретных экспериментов и оценки точности результатов, полученных в наблюдаемых экспериментальных условиях; генерирует синтетические данные для обучения методов интеллектуального анализа и последующего создания моделей идентификации однонуклеотидных полиморфизмов в конкретных экспериментальных наборах данных; моделирует наборы данных для тестирования и сравнения доступных методов идентификации однонуклеотидных полиморфизмов при анализе экспериментальных данных; предоставляет возможность удаленной работы с приложением широкой научной аудитории пользователей в режиме онлайн через глобальную сеть Интернет.

Проверка работоспособности разработанных моделей и методов программного комплекса проведена на примерах наборов смоделированных и экспериментальных данных геномного секвенирования молекул ДНК человека. Выполнен сравнительный анализ наиболее эффективных алгоритмов идентификации сайтов однонуклеотидных полиморфизмов. Наилучшие результаты получены для моделей машинного обучения – точность идентификации сайтов по мере F_1 на 2–5 % выше у обученных на смоделированных данных методов деревьев решений, чем у традиционных методов идентификации.

Комплекс может использоваться для моделирования синтетических данных по экспериментальным данным или самостоятельно с целью всестороннего тестирования и выбора наилучших алгоритмов идентификации СОП, а также для генеративного моделирования данных с целью обучения алгоритмов идентификации на основе методов машинного обучения (нейронных и байесовских сетей, ансамблевых алгоритмов и пр.).

Вклад авторов. Н. Н. Яцков предложил методику и разработал программный комплекс для имитационного моделирования и идентификации сайтов однонуклеотидного полиморфизма с использованием методов машинного обучения, подготовил текст статьи. Д. Д. Сарнацкий разработал R-функции программного комплекса. В. В. Скакун разработал веб-сайт, базу данных и репозиторий для сохранения результатов анализа. В. В. Гринев сформулировал задачу об идентификации сайтов однонуклеотидного полиморфизма и разработал R-функции программного комплекса. Все авторы приняли участие в обобщении, анализе и оформлении полученных результатов.

Список использованных источников

1. Sung, W. K. Algorithms for Next Generation Sequencing / W. K. Sung. – 1st ed. – N. Y. : Chapman & Hall / CRC, 2017. – 364 p.
2. Kappelmann-Fenzl, M. Next Generation Sequencing and Data Analysis / ed. M. Kappelmann-Fenzl. – 1st ed. – Cham : Springer, 2021. – 218 p.
3. Optimal design of low-density SNP arrays for genomic prediction: algorithm and applications / X. L. Wu, J. Xu, G. Feng [et al.] // PLoS ONE. – Sept. 2016. – Vol. 11, no 9. – P. e0161719. – DOI: 10.1371/journal.pone.0161719.
4. Machine learning as an effective method for identifying true single nucleotide polymorphisms in polyploid plants / W. Korani, J. P. Clevenger, Y. Chu, P. Ozias-Akins // Plant Genome. – Mar. 2019. – Vol. 12, iss. 1. – P. 180023. – DOI: 10.3835/plantgenome2018.05.0023.
5. Masoudi-Nejad, A. Next Generation Sequencing and Sequence Assembly. Methodologies and Algorithms / A. Masoudi-Nejad, Z. Narimani, N. Hosseinkhan. – 1st ed. – N. Y. : Springer, 2013. – 86 p.
6. Su, Z. HAPGEN2: simulations of multiple disease SNPs / Z. Su, J. Marchini, P. Donnelly // Bioinformatics. – 2011. – Vol. 27, iss. 16. – P. 2304–2305.
7. Oh, J. H. SITDEM: a simulation tool for disease/endpoint models of association studies based on single nucleotide polymorphism genotypes / J. H. Oh, J. O. Deasy // Computers in Biology and Medicine. – 2014. – Vol. 45. – P. 136–142.
8. A comparison of gene region simulation methods / A. E. Hendricks, J. Dupuis, M. Gupta [et al.] // PLoS ONE. – 2012. – Vol. 7, no 7. – P. e40925. – DOI: 10.1371/journal.pone.0040925.
9. Genetic Simulation Resources: a website for the registration and discovery of genetic data simulators / B. Peng, H. S. Chen, L. E. Mechanic [et al.] // Bioinformatics. – 2013. – Vol. 29, iss. 8. – P. 1101–1102.
10. Genetic data simulators and their applications: an overview / B. Peng, H. S. Chen, L. E. Mechanic [et al.] // Genetic Epidemiology. – 2015. – Vol. 39, iss. 1. – P. 2–10.
11. Tahmasbi, R. GeneEvolve: a fast and memory efficient forward-time simulator of realistic whole-genome sequence and SNP data / R. Tahmasbi, M. C. Keller // Bioinformatics. – 2017. – Vol. 33, iss. 2. – P. 294–296.
12. Posada, D. Simulating haplotype blocks in the human genome / D. Posada, C. Wiuf // Bioinformatics. – 2003. – Vol. 19, iss. 2. – P. 289–290.
13. DHOEM: a statistical simulation software for simulating new markers in real SNP marker data / L. Jacquin, T. V. Cao, C. Grenier, N. Ahmadi // BMC Bioinformatics. – Dec. 2015. – Vol. 16. – P. 404. – DOI: 10.1186/s12859-015-0830-7.
14. Meyer, H. V. PhenotypeSimulator: A comprehensive framework for simulating multi-trait, multi-locus genotype to phenotype relationships / H. V. Meyer, E. Birney // Bioinformatics. – 2018. – Vol. 34, iss. 17. – P. 2951–2956.
15. sim1000G: a user-friendly genetic variant simulator in R for unrelated individuals and family-based designs / A. Dimitromanolakis, J. Xu, A. Krol, L. Briollais // BMC Bioinformatics. – Jan. 2019. – Vol. 20, no 1. – P. 26. – DOI: 10.1186/s12859-019-2611-1.
16. Яцков, Н. Н. Генеративное имитационное моделирование сложных биофизических систем / Н. Н. Яцков, В. В. Апанасович, В. Н. Яцков // Компьютерные технологии и анализ данных (CTDA'2024) : материалы IV Междунар. науч.-практ. конф., Минск, 25–26 апр. 2024 г. – Минск : БГУ, 2024. – С. 211–214.
17. Simulation modelling for machine learning identification of single nucleotide polymorphisms in human genomes / M. M. Yatskou, E. V. Smolyakova, V. V. Skakun, V. V. Grinev // Pattern Recognition and Information Processing (PRIP'2023) : Proc. of the 16th Intern. Conf., Minsk, 17–19 Oct. 2023. – Minsk : BSU, 2023. – P. 49–53.
18. Яцков, Н. Н. Вычислительный подход и программный пакет RNAexploreR для группировки молекул РНК генов человека по их экзонным признакам / Н. Н. Яцков, В. В. Скаун, В. В. Гринев // Информатика. – 2019. – Т. 16, № 4. – С. 7–24.
19. Gentleman, R. Bioconductor: Open software development for computational biology and bioinformatics / R. Gentleman, V. J. Carey, D. M. Bates // Genome Biology. – 2004. – Vol. 5, no. 10, art. R80. – URL: <https://genomebiology.biomedcentral.com/articles/10.1186/gb-2004-5-10-r80> (date of access: 10.04.2025). – DOI: 10.1186/GB-2004-5-10-R80.
20. Программный пакет SNPSimulatoR для моделирования сайтов одноклеточного генетического полиморфизма / Н. Н. Яцков, Е. В. Смолякова, В. В. Скаун, В. В. Гринев // Квантовая электроника : материалы XIV Междунар. науч.-техн. конф., Минск, 21–23 нояб. 2023 г. – Минск : БГУ, 2023. – С. 510–515.
21. An open resource for accurately benchmarking small variant and reference calls / J. M. Zook, J. McDaniel, N. D. Olson [et al.] // Nature Biotechnology. – 2019. – Vol. 37, no. 5. – P. 561–566.

22. Идентификация сайтов однонуклеотидного генетического полиморфизма с использованием методов машинного обучения / Н. Н. Яцков, Е. В. Смолякова, К. И. Грудовик [и др.] // Квантовая электроника : материалы XIV Междунар. науч.-техн. конф., Минск, 21–23 нояб. 2023 г. – Минск : БГУ, 2023. – С. 504–509.
23. Identification of single nucleotide genetic polymorphism sites using machine learning methods / M. M. Yatskou, E. V. Smolyakova, V. V. Skakun, V. V. Grinev // Advances in Transdisciplinary Engineering. – 2023. – Vol. 42. – Р. 1031–1037.
24. Yatskou, M. M. Simulation modelling of single nucleotide genetic polymorphisms / M. M. Yatskou, V. V. Apanasovich, V. V. Grinev // Journal of the Belarusian State University. Mathematics and Informatics. – 2024. – No. 2. – Р. 104–112.
25. Яцков, Н. Н. Вычислительная платформа FluorSimStudio для обработки кинетических кривых застухания флуоресценции с использованием алгоритмов имитационного моделирования и интеллектуального анализа данных / Н. Н. Яцков, В. В. Апанасович // Журнал прикладной спектроскопии. – 2021. – Т. 88, № 3. – С. 452–461.
26. Сарнацкий, Д. Д. Имитационная модель генерации сайтов однонуклеотидного полиморфизма в молекулах ДНК человека / Д. Д. Сарнацкий, Н. Н. Яцков, В. В. Гринева // Компьютерные технологии и анализ данных (CTDA'2024) : материалы IV Междунар. науч.-практ. конф., Минск, 25–26 апр. 2024 г. – Минск : БГУ, 2024. – С. 265–268.
27. Сарнацкий, Д. Д. Исследование информативности признаков нуклеотидных сайтов при определении генетических полиморфизмов с использованием методов машинного обучения / Д. Д. Сарнацкий, Н. Н. Яцков, В. В. Гринев // Информационные технологии и системы (ИТС 2024) : материалы Междунар. науч. конф., Минск, 20 нояб. 2024 г. – Минск : БГУИР, 2024. – С. 69–70.
28. Яцков, Н. Н. Нейросетевое имитационное моделирование при анализе экспериментальных данных флуоресцентной спектроскопии / Н. Н. Яцков, В. В. Апанасович // Компьютерные технологии и анализ данных (CTDA'2024) : материалы IV Междунар. науч.-практ. конф., Минск, 25–26 апр. 2024 г. – Минск : БГУ, 2024. – С. 215–218.

References

1. Sung W. K. *Algorithms for Next Generation Sequencing*, 1st ed. New York, Chapman & Hall / CRC, 2017, 364 p.
2. Kappelmann-Fenzl M. (ed.). *Next Generation Sequencing and Data Analysis*, 1st ed. Cham, Springer, 2021, 218 p.
3. Wu X. L., Xu J., Feng G., Wiggans G. R., Taylor J. F., ..., Bauck S. Optimal design of low-density SNP arrays for genomic prediction: algorithm and applications. *PLoS ONE*, September 2016, vol. 11, no 9, p. e0161719. DOI: 10.1371/journal.pone.0161719.
4. Korani W., Clevenger J. P., Chu Y., Ozias-Akins P. Machine learning as an effective method for identifying true single nucleotide polymorphisms in polyploid plants. *Plant Genome*, March 2019, vol. 12, iss. 1, p. 180023. DOI: 10.3835/plantgenome2018.05.0023.
5. Masoudi-Nejad A., Narimani Z., Hosseinkhan N. *Next Generation Sequencing and Sequence Assembly. Methodologies and Algorithms*, 1st ed. New York, Springer, 2013, 86 p.
6. Su Z., Marchini J., Donnelly P. HAPGEN2: simulations of multiple disease SNPs. *Bioinformatics*, 2011, vol. 27, iss. 16, p. 2304–2305.
7. Oh J. H., Deasy J. O. SITDEM: a simulation tool for disease/endpoint models of association studies based on single nucleotide polymorphism genotypes. *Computers in Biology and Medicine*, 2014, vol. 45, pp. 136–142.
8. Hendricks A. E., Dupuis J., Gupta M., Logue M. W., Lunetta K. L. A comparison of gene region simulation methods. *PLoS ONE*, 2012, vol. 7, no 7, p. e40925. DOI: 10.1371/journal.pone.0040925.
9. Peng B., Chen H. S., Mechanic L. E., Racine B., Clarke J., ..., Feuer E. J. Genetic Simulation Resources: a website for the registration and discovery of genetic data simulators. *Bioinformatics*, 2013, vol. 29, iss. 8, pp. 1101–1102.
10. Peng B., Chen H. S., Mechanic L. E., Racine B., Clarke J., ..., Feuer E. J. Genetic data simulators and their applications: an overview. *Genetic Epidemiology*, 2015, vol. 39, iss. 1, pp. 2–10.
11. Tahmasbi R., Keller M. C. GeneEvolve: a fast and memory efficient forward-time simulator of realistic whole-genome sequence and SNP data. *Bioinformatics*, 2017, vol. 33, iss. 2, pp. 294–296.
12. Posada D., Wiuf C. Simulating haplotype blocks in the human genome. *Bioinformatics*, 2003, vol. 19, iss. 2, pp. 289–290.

13. Jacquin L., Cao T. V., Grenier C., Ahmadi N. DHOEM: a statistical simulation software for simulating new markers in real SNP marker data. *BMC Bioinformatics*, December 2015, vol. 16, p. 404. DOI: 10.1186/s12859-015-0830-7.
14. Meyer H. V., Birney E. PhenotypeSimulator: A comprehensive framework for simulating multi-trait, multi-locus genotype to phenotype relationships. *Bioinformatics*, 2018, vol. 34, iss. 17, pp. 2951–2956.
15. Dimitromanolakis A., Xu J., Krol A., Briollais L. sim1000G: a user-friendly genetic variant simulator in R for unrelated individuals and family-based designs. *BMC Bioinformatics*, January 2019, vol. 20, no. 1, p. 26. DOI: 10.1186/s12859-019-2611-1.
16. Yatskou M. M., Apanasovich V. V., Yatskou U. M. *Generative simulation modelling of complex biophysical systems*. Kompyuternye tehnologii i analiz dannyh (CTDA'2024): materialy IV Mezhdunarodnoj nauchno-prakticheskoy konferencii, Minsk, 25–26 aprelya 2024 g. [Computer Technologies and Data Analysis (CTDA'2024): Proceedings of the IV International Scientific Conference, Minsk, 25–26 April 2024]. Minsk, Belorusskij gosudarstvennyj universitet, 2024, pp. 211–214 (In Russ.).
17. Yatskou M. M., Smolyakova E. V., Skakun V. V., Grinev V. V. Simulation modelling for machine learning identification of single nucleotide polymorphisms in human Genomes. *Pattern Recognition and Information Processing (PRIP'2023) : Proceedings of the 16th International Conference, Minsk, 17–19 October 2023*. Minsk, Belarusian State University, 2023, pp. 49–53.
18. Yatskou M. M., Skakun V. V., Grinev V. V. A computational approach and software package RNAExploreR for grouping RNA molecules of human genes by exon features. *Informatika [Informatics]*, 2019, vol. 16, no. 4, pp. 7–24 (In Russ.).
19. Gentleman R., Carey V. J., Bates D. M. Bioconductor: Open software development for computational biology and bioinformatics. *Genome Biology*, 2004, vol. 5, no. 10, art. R80. Available at: <https://genomebiology.biomedcentral.com/articles/10.1186/gb-2004-5-10-r80> (accessed 10.04.2025). DOI: 10.1186/GB-2004-5-10-R80.
20. Yatskou M. M., Smolyakova E. V., Skakun V. V., Grinev V. V. *R-package SNPSimulatoR for modelling single nucleotide genetic polymorphism sites*. Kvantovaya elektronika: materialy XIV Mezhdunarodnoj nauchno-tehnicheskoy konferencii, Minsk, 21–23 noyabrya 2023 g. [Quantum Electronics: Proceedings of the XIV International Scientific and Practical Conference, Minsk, 21–23 November 2023]. Minsk, Belorusskij gosudarstvennyj universitet, 2023, pp. 510–515 (In Russ.).
21. Zook J. M., McDaniel J., Olson N. D., Wagner J., Parikh H., ..., Salit M. An open resource for accurately benchmarking small variant and reference calls. *Nature Biotechnology*, 2019, vol. 37, no 5, pp. 561–566.
22. Yatskou M. M., Smolyakova E. V., Grudovik K. I., Skakun V. V., Grinev V. V. *Identification of single nucleotide genetic polymorphisms using machine learning methods*. Kvantovaya elektronika: materialy XIV Mezhdunarodnoj nauchno-tehnicheskoy konferencii, Minsk, 21–23 noyabrya 2023 g. [Quantum Electronics: Proceedings of the XIV International Scientific and Practical Conference, Minsk, 21–23 November 2023]. Minsk, Belorusskij gosudarstvennyj universitet, 2023, pp. 504–509 (In Russ.).
23. Yatskou M. M., Smolyakova E. V., Skakun V. V., Grinev V. V. Identification of Single nucleotide genetic polymorphism sites using machine learning methods. *Advances in Transdisciplinary Engineering*, 2023, vol. 42, pp. 1031–1037.
24. Yatskou M. M., Grinev V. V., Apanasovich V. V. Simulation modelling of single nucleotide genetic polymorphisms. *Journal of the Belarusian State University. Mathematics and Informatics*, 2024, no. 2, pp. 104–112.
25. Yatskou M. M., Apanasovich V. V. Computational platform FluorSimStudio for processing kinetic curves of fluorescence decay using simulation modeling and data mining algorithms. *Journal of Applied Spectroscopy*, 2021, vol. 88, no. 3, pp. 452–461.
26. Sarnatski D. D., Yatskou M. M., Grinev V. V. *Simulation model of generation of single nucleotide polymorphism sites in human DNA molecules*. Kompyuternye tehnologii i analiz dannyh (CTDA'2024): materialy IV Mezhdunarodnoj nauchno-prakticheskoy konferencii, Minsk, 25–26 aprelya 2024 g. [Computer Technologies and Data Analysis (CTDA'2024): Proceedings of the IV International Scientific Conference, Minsk, 25–26 April 2024]. Minsk, Belorusskij gosudarstvennyj universitet, 2024, pp. 265–268 (In Russ.).
27. Sarnatski D. D., Yatskou M. M., Grinev V. V. *Study of the informativeness of nucleotide site features in determining genetic polymorphisms using machine learning methods*. Informacionnye tehnologii i sistemy (ITS 2024): materialy Mezhdunarodnoj nauchnoj konferencii, Minsk, 20 noyabrya 2024 g. [Information Technologies and Systems 2024 (ITS 2024): Proceedings of the International Scientific Conference, Minsk, 20 November 2024]. Minsk, Belorusskij gosudarstvennyj universitet informatiki i radioelektroniki, 2024, pp. 69–70 (In Russ.).
28. Yatskou M. M., Apanasovich V. V. *Neural network simulation modelling when analyzing experimental fluorescence spectroscopy data*. Kompyuternye tehnologii i analiz dannyh (CTDA'2024): materialy IV Mezhdunarodnoj nauchno-prakticheskoy konferencii, Minsk, 25–26 aprelya 2024 g. [Computer Technologies and Data Analysis (CTDA'2024): Proceedings of the IV International Scientific Conference, Minsk, 25–26 April 2024]. Minsk, Belorusskij gosudarstvennyj universitet, 2024, pp. 215–218 (In Russ.).

Информация об авторах

Яцков Николай Николаевич, кандидат физико-математических наук, доцент, доцент кафедры системного анализа и компьютерного моделирования, факультет радиофизики и компьютерных технологий, Белорусский государственный университет.

E-mail: yatskou@bsu.by

Сарнацкий Денис Дмитриевич, студент кафедры системного анализа и компьютерного моделирования, факультет радиофизики и компьютерных технологий, Белорусский государственный университет.

E-mail: denisiusarnatski@gmail.com

Скаакун Виктор Васильевич, кандидат физико-математических наук, доцент, заведующий кафедрой системного анализа и компьютерного моделирования, факультет радиофизики и компьютерных технологий, Белорусский государственный университет.

E-mail: skakun@bsu.by

Гринев Василий Викторович, кандидат биологических наук, доцент, доцент кафедры генетики, биологический факультет, Белорусский государственный университет.

E-mail: grinev_vv@bsu.by

Information about the authors

Mikalai M. Yatskou, Ph. D. (Phys.-Math.), Assoc. Prof., Department of Systems Analysis and Computer Modelling, Faculty of Radiophysics and Computer Technologies, Belarusian State University.

E-mail: yatskou@bsu.by

Dzianis D. Sarnatski, Student, Department of Systems Analysis and Computer Modelling, Faculty of Radiophysics and Computer Technologies, Belarusian State University.

E-mail: denisiusarnatski@gmail.com

Victor V. Skakun, Ph. D. (Phys.-Math.), Assoc. Prof., Head of Department of Systems Analysis and Computer Modelling, Faculty of Radiophysics and Computer Technologies, Belarusian State University.

E-mail: skakun@bsu.by

Vasily V. Grinev, Ph. D. (Biol.), Assoc. Prof., Department of Genetics, Faculty of Biology, Belarusian State University.

E-mail: grinev_vv@bsu.by

ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ INFORMATION TECHNOLOGIES



УДК 004.056.52
DOI: 10.37661/1816-0301-2025-22-2-95-110

Оригинальная статья
Original Article

Прототипирование системы беспарольного атрибутивного доступа к информационным ресурсам с использованием eID-карты Республики Беларусь и FIDO2-аутентификации

А. А. Жидович, А. А. Лубенько, И. С. Войтешенко[✉]

Белорусский государственный университет,
пр. Независимости, 4, Минск, 220030, Беларусь
[✉]E-mail: voit@bsu.by

Аннотация

Цели. Целью проведенной аналитической и исследовательской работы являются проектирование и реализация прототипа системы установки личности пользователя и его привилегий путем совместного использования беспарольной FIDO2-аутентификации и управления доступом на основе атрибутов. В качестве источника пользовательских атрибутов предложены средства электронной идентификации, соответствующие стандартам ICAO.

Методы. В исследовании применялись: систематизация и анализ литературы и технических спецификаций; системный подход к анализу существующих реализаций систем беспарольного атрибутивного доступа и теоретических моделей, используемых при их проектировании; SCn- и SCg-код технологии OSTIS для семантического описания основных понятий и концепций, связанных с FIDO2-аутентификацией; программные платформы и библиотеки.

Результаты. Результатом работы является прототип системы атрибутивного доступа к информационным ресурсам в цифровой среде с использованием eID-карты Республики Беларусь и FIDO2-аутентификации. Разработанное приложение было контейнеризовано и развернуто на онлайн-сервере, его работоспособность проверена с различных платформ с помощью распространенных браузеров.

Заключение. Представлено исследование по разработке и первоначальной оценке прототипа системы управления доступом к информационным ресурсам с помощью аутентификации по спецификации FIDO2 и модели управления доступом на основе атрибутов. При этом в качестве источника пользовательских атрибутов применяются средства электронной идентификации, удовлетворяющие стандартам Международной организации гражданской авиации, в том числе eID-карта Республики Беларусь.

Ключевые слова: беспарольная аутентификация, авторизация на основе атрибутов, спецификация FIDO2, протокол W3C WebAuthn, протокол «клиент-аутентификатор», eID-идентификация

Благодарности. Авторы выражают благодарность за содействие в работе профессору А. Н. Курбацкому и сотрудникам ЗАО «АВЕСТ».

Для цитирования. Жидович, А. А. Прототипирование системы беспарольного атрибутивного доступа к информационным ресурсам с использованием eID-карты Республики Беларусь и FIDO2-автентификации / А. А. Жидович, А. А. Лубенъко, И. С. Войтешенко // Информатика. – 2025. – Т. 22, № 2. – С. 95–110. – DOI: 10.37661/1816-0301-2025-22-2-95-110.

Конфликт интересов. Авторы заявляют об отсутствии конфликта интересов.

Поступила в редакцию | Received 17.03.2025

Подписана в печать | Accepted 01.04.2025

Опубликована | Published 30.06.2025

Prototyping of the system of passwordless attributive access to information resources using eID-card of the Republic of Belarus and FIDO2-authentication

Anton A. Zhidovich, Alexei A. Lubenko, Iosif S. Vojteshenko[✉]

*Belarusian State University,
av. Nezavisimosti, 4, Minsk, 220030, Belarus
✉E-mail: voit@bsu.by*

Abstract

Objectives. The purpose of the analytical and research work carried out is to design and implement a prototype system for establishing user identity and privileges based on the joint use of passwordless FIDO2 authentication and attribute-based access control. It is proposed that electronic identification means compliant with ICAO standards be used as a source of user attributes.

Methods. The following were used in this paper: systematization and analysis of literature and technical specifications; systematic approach to the analysis of existing implementations of passwordless attribute access systems and theoretical models used in their design; the SCn- and SCg-code of OSTIS technology for semantic description of basic concepts and concepts related to FIDO2-authentication; software platforms and libraries.

Results. The result of the work is a prototype of the system of attributive access to information resources in the digital environment using the eID-card of the Republic of Belarus and FIDO2-authentication. The developed application was containerized and deployed on the online server. Its performance was then tested from different platforms using standard browsers.

Conclusion. A study on the development and initial evaluation of a prototype of an information resource access control system based on authentication to the FIDO2 specification and an attribute-based access control model is presented. At the same time, as a source of user attributes the means of electronic identification that meet the standards of the International Civil Aviation Organization, including the eID-card of the Republic of Belarus, are used.

Keywords: passwordless authentication, attribute-based authorization, FIDO2 specification, protocol W3C WebAuthn, client-to-authenticator protocol, eID-identification

Acknowledgments. The authors express their gratitude to professor A. N. Kurbatsky and the staff of CJSC AVEST for their assistance in this work.

For citation. Zhidovich A. A., Lubenko A. A., Vojteshenko I. S. *Prototyping of the system of passwordless attributive access to information resources using eID-card of the Republic of Belarus and FIDO2-authentication*. Informatika [Informatics], 2025, vol. 22, no. 2, pp. 95–110 (In Russ.). DOI: 10.37661/1816-0301-2025-22-2-95-110.

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

Введение. Для большинства киберпреступлений характерна кража учетных данных пользователя. Например, получив доступ к учетным данным жертвы, злоумышленник может продать данные или совершить противоправные действия различной направленности.

В основе систем безопасности в цифровой среде лежит процесс установки личности пользователя и его привилегий. Таким образом, вопросы идентификации, аутентификации и авторизации являются одними из самых важных при построении безопасной информационной системы.

Современные многокомпонентные системы, состоящие из множества взаимодействующих ресурсов, характеризуются высокой сложностью в контексте требований к авторизации пользователей. Организация эффективного и безопасного механизма установки личности пользователя и его привилегий зачастую становится приоритетной задачей для разработчиков. Сложность системы авторизации обуславливает необходимость создания интеграционных решений, обеспечивающих унификацию подходов к идентификации и управлению доступом. Это требует дополнительных усилий на этапе проектирования, что увеличивает общую стоимость и временные затраты на разработку таких систем.

Активное развитие электронных идентификационных документов (eID) открывает новые горизонты для безопасного и эффективного доступа, например, к финансовым, медицинским, муниципальным и государственным информационным системам. В условиях растущей необходимости в защите персональных данных и соблюдении соответствующих требований законодательства интеграция eID в такие системы может стать ключевым шагом к повышению уровня безопасности, удобства и доверия пользователей.

Одним из трендов в информационной безопасности является процесс перехода к беспарольным методам аутентификации. В частности, все большее внимание уделяется спецификациям FIDO2, повышающим безопасность и конфиденциальность пользователя по сравнению с классическими парольными методами. На данный момент в мире ряд организаций предоставляют унифицированный компонент, позволяющий интегрировать FIDO2-аутентификацию в различные платформы. Стоит отметить, что этот подход пользуется доверием среди компаний в области финансовых технологий, где уязвимость может привести к серьезным потерям.

Таким образом, унифицированная система установки личности пользователя и его привилегий, интегрирующая возможности eID-идентификации, гибкой авторизации и FIDO2-аутентификации, способна сократить время и стоимость разработки сложных информационных систем, повысить степень защиты конфиденциальных данных, а также упростить доступ для пользователей.

Целями настоящей работы являются проектирование и реализация прототипа системы установки личности пользователя и его привилегий на основе совместного применения беспарольной FIDO2-аутентификации и атрибутивного управления доступом. В качестве источника пользовательских атрибутов предложено использование средств электронной идентификации, соответствующих стандартам Международной организации гражданской авиации (ICAO). Было показано, что на данный момент это решение является перспективным и удобным в использовании.

Теоретические и технологические предпосылки проведения исследований и их практической реализации

Технология беспарольной FIDO2-аутентификации. В отличие от доступа с помощью пароля беспарольные решения с самого начала разрабатывались с учетом современного удобства использования и современных ландшафтов атак. Примерами беспарольного доступа являются доступы с применением Windows Hello¹, отпечатка пальца, сканера сетчатки глаза, FIDO-токенов и т. д. Все эти решения по своей сути являются многофакторными и безопасными с точки зрения протоколов.

Несмотря на то что технология FIDO2 уже находит достаточно широкое практическое применение, научные исследования и практические разработки в этой области активно развиваются [1–6].

¹Windows Hello. – URL: <https://www.microsoft.com/en-us/windows/tips/windows-hello> (date of access: 12.01.2025).

Стандарт FIDO2, разработанный FIDO-альянсом², представляет собой совокупность спецификаций в двух частях. Первая часть называется стандартом веб-аутентификации W3C WebAuthn³. Эта спецификация представляет собой JavaScript API, с помощью которого удаленные веб-сайты могут запрашивать учетные данные открытого ключа. Вторая часть – это спецификация протокола «клиент-аутентификатор» (client-to-authenticator protocol, CTAP). Протокол «клиент-аутентификатор» описывает, как приложение и операционная система устанавливают связь с совместимым устройством аутентификации через USB-, NFC- или BLE-средства связи.

Взаимодействие между платформой и аутентификатором по данному протоколу можно описать следующим образом:

1. Платформа устанавливает соединение с аутентификатором.
2. Платформа получает информацию об аутентификаторе и его возможностях.
3. Платформа отправляет команду для операции, поддерживаемой средством аутентификации.
4. Аутентификатор отвечает данными ответа или ошибкой.

Технология, лежащая в основе FIDO2, базируется на асимметричной криптографии. В этом процессе каждый закрытый и открытый ключ генерирует пару ключей. Открытый ключ получателя не является секретом и может быть обменен или опубликован в любое время с партнерами по коммуникации для шифрования сообщений. Для расшифровки сообщения требуется закрытый ключ. Он должен храниться получателем в безопасности, поскольку третьи стороны могли бы расшифровать сообщение, если бы они его перехватили.

Каждый аутентификатор, доступный для FIDO2, имеет сертификат X.509, также известный как сертификат аттестации, сохраняемый при изготовлении устройства. Закрытый ключ записывается на устройство и не может быть экспортирован или иным образом изменен. Кроме того, устройство криптографически сертифицировано. Это означает, что, если злоумышленники попытаются перехватить запрос на регистрацию и заменить его своим собственным, они не смогут обменять сгенерированный открытый ключ на другой ключ, поскольку подписи подтверждения не будут совпадать. Схема взаимодействия компонентов FIDO2 изображена на рис. 1.

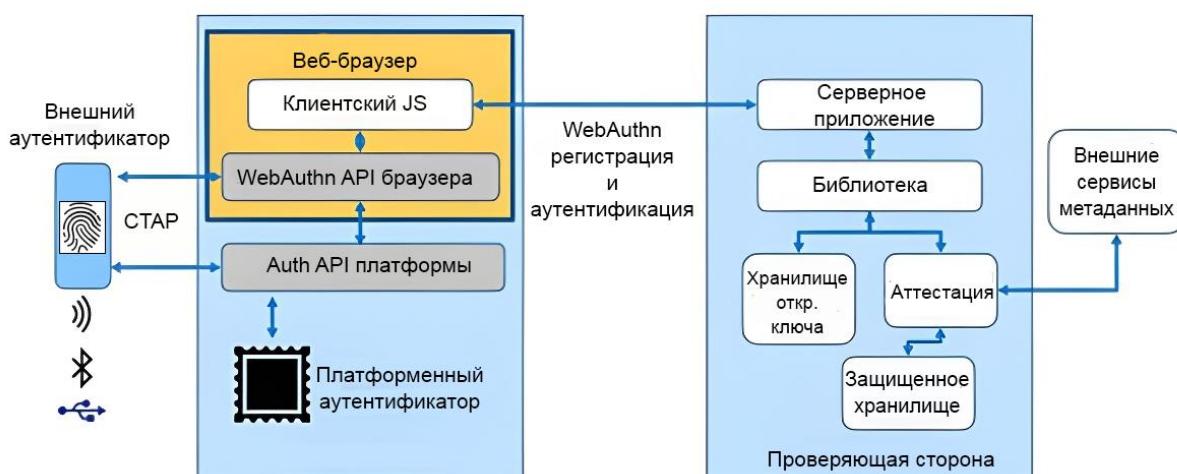


Рис. 1. Компоненты FIDO2 и их взаимодействие

Fig. 1. FIDO2 components and their interaction

²Fido Alliance. – URL: <https://fidoalliance.org> (date of access: 12.01.2025).

³WebAuthn API specification. – URL: <https://w3c.github.io/webauthn/> (date of access: 12.01.2025).

Платформозависимый и платформонезависимый аутентификаторы действуют как защищенный объект, интегрированный в конечное устройство пользователя (например, Windows Hello в качестве платформозависимого аутентификатора) или подключенный к нему (например, аппаратный токен USB в качестве платформонезависимого аутентификатора). Служба метаданных FIDO помогает подтвердить подлинность аутентификатора, регулярно извлекая цепочки сертификатов аутентификатора.

Производители генерируют и хранят глобально уникальный идентификатор подтверждения подлинности случайным образом, который идентифицирует модель аутентификации. Процесс регистрации уникального идентификатора (ключа) подробно описан в работе [4].

Для аутентификации пользователь подтверждает свою личность проверяющей стороне через клиента WebAuthn с помощью средства аутентификации.

После того как пользователь зарегистрировал свой ключ, он может применять его для беспарольной аутентификации. Порядок действий при аутентификации во многом совпадает с процессом регистрации, однако имеет свои особенности [7]. Рассмотрим подробнее процесс аутентификации:

1. Когда пользователь переходит к интернет-сервису, поддерживающему WebAuthn, и хочет пройти аутентификацию, для клиента WebAuthn отправляется запрос на аутентификацию проверяющей стороне.
2. Проверяющая сторона получает запрос на аутентификацию и отправляет сгенерированный запрос обратно клиенту WebAuthn.
3. Клиент перенаправляет полученный запрос аутентификатору. В запросе содержатся информация о необходимом типе аутентификатора, идентификаторы зарегистрированных ключей и идентификатор пользователя.
4. Аутентификатор находит сохраненные ключи и предлагает пользователю выбрать один из них. После этого выполняется проверка подлинности пользователя. В случае успеха аутентификатор создает ответ, называемый *assertion response*, который содержит в себе подпись, созданную закрытым ключом, и другую информацию, необходимую для проверяющей стороны.
5. Аутентификатор отправляет ответ клиенту.
6. Клиент создает и отправляет ответ, содержащий в себе *assertion response*, проверяющей стороне.
7. Проверяющая сторона проверяет доступность открытого ключа в базе данных и подпись, находящуюся в *assertion response*, с помощью открытого ключа, сохраненного во время регистрации.
8. Проверяющая сторона дает доступ к данным пользователя.

Совместное использование FIDO2-аутентификации и управления доступом на основе атрибутов может существенно повысить уровень гибкости и масштабируемости системы безопасности. Важно понимать, что это два разных процесса, причем базовая спецификация FIDO2 (в частности, WebAuthn) не предусматривает подобную интеграцию. Следовательно, совместная настройка ABAC и FIDO2 может требовать дополнительных усилий и ресурсов, специфических для каждой системы или задачи. Например, если в системе присутствуют динамические атрибуты (такие, как местоположение), то интеграция с FIDO2-аутентификацией потребует внедрения механизмов сбора и обновления этих данных, при этом требования спецификации не нарушаются.

Беспарольный доступ на основе атрибутов с использованием медиатора. В работе [2] предложено комплексное и готовое к интеграции в промышленные системы решение для использования анонимных учетных данных с локальной или удаленной аттестацией в виде фреймворка FIDO-AC, который является расширением базовой спецификации FIDO2. Там же представлены оценка безопасности и конфиденциальности, предоставляемых полученной системой, и реализация рабочего прототипа.

Суть полученного в работе [2] решения заключается в создании дополнительной стороны – посредника (медиатора), который отвечает за валидацию данных о пользователе, полученных, например, с eID-документа. Таким образом, в FIDO-AC выделены следующие стороны – участники процесса авторизации:

- FIDO-сервер (проверяющая сторона), формирующий запросы к FIDO-клиенту согласно спецификации FIDO2 и проводящий верификацию полученного ответа;
- FIDO-аутентификатор, использующийся для работы базовой спецификации FIDO2;
- клиент (браузер/приложение с поддержкой WebAuthn) для взаимодействия FIDO-сервера с аутентификатором и приложением FIDO-AC;
- медиатор (посредник), отвечающий за аттестацию данных, полученных с eID;
- приложение FIDO-AC для считывания данных с электронного идентификатора пользователя;
- электронный идентификатор (eID) для удостоверения личности со встроенным запоминающим устройством (чипом), хранящим данные пользователя и подпись эмитента.

Медиатор и приложение FIDO-AC могут быть как отдельными сторонами, так и реализованными модулями одного приложения.

Процесс начинается со считывания пользователем зашифрованных данных и криптографической подписи со своего eID с помощью приложения FIDO-AC. В случае с электронным паспортом требуется ввести номер документа, дату своего рождения и срок действия, на основании которых генерируется ключ для прохождения базового контроля доступа.

Затем при попытке доступа к защищенному ресурсу системы FIDO-сервер создает запрос к клиенту. Он должен соответствовать спецификации WebAuthn и содержать криптографически случайный буфер байтов challenge. Получив запрос, клиент перенаправляет его в приложение FIDO-AC вместе с данными, полученными с электронного документа (а именно данные пользователя, подпись эмитента и открытый ключ, связанный с данными). После этого медиатор должен провести пассивную и активную фазы аутентификации.

На следующем шаге приложение FIDO2 с использованием доказательства с нулевым разглашением (авторами предложены ZK-SNARK и ZK-STARK [8]) на основе данных с eID создает подпись, подтверждающую соответствие пользователя политике ресурса. Эта подпись, а также подпись медиатора передаются клиенту и добавляются к буферу challenge. Затем происходит стандартная FIDO2-аутентификация. Полученный в процессе объект аттестации передается проверяющей стороне.

На завершающем этапе проверяющая сторона должна провести две верификации: проверку подписи, изданной аутентификатором FIDO, и подписи медиатора. Так как значение challenge было изменено, проверяющая сторона должна применить аналогичные изменения к буферу challenge, с которым производится сравнение, а именно присоединить хеш-значение подписи медиатора.

Предлагаемое решение на основе интеграции eID-идентификации, FIDO2-аутентификации и авторизации на основе атрибутов

Функциональные и нефункциональные требования. В соответствии со спецификацией WebAuthn API приложение должно поддерживать регистрацию ключа (платформозависимого) и аутентификацию с его использованием. Прежде всего для регистрации FIDO2-аутентификатора пользователь должен быть авторизованным, т. е. приложение должно иметь возможность аутентификации каким-либо из классических способов. В качестве такого способа был выбран парольный. После регистрации ключа его можно применять для беспарольного доступа в свой аккаунт. При этом желательно, чтобы в результате успешной авторизации устанавливались аутентификационные Cookie. Они создаются на стороне сервера и используются при последующих входах с того же клиента, не требуя от пользователя никаких действий.

Платформа, на которой разработано веб-приложение, должна поддерживать библиотеки, позволяющие выступать ей в роли проверяющей стороны. Они необходимы для создания соответствующих стандарту запросов регистрации ключа и создания подписи для аутентификации с WebAuthn API.

Для того чтобы обеспечить доступ к разработанному приложению с различных устройств, должна быть возможность разместить его на онлайн-сервисе. Также для более удобной разработки желательно, чтобы выбранный сервис поддерживал непрерывное развертывание.

Для развертывания приложение должно поддерживать возможность запуска на UNIX-подобных операционных системах Linux для последующей контейнеризации.

При попытке доступа к защищенному ресурсу, которому соответствует некоторая политика, должно осуществляться считывание данных о пользователе с электронного документа. Для считывания данных электронного документа, проведения валидации представленных на нем данных и проверки соответствия атрибутов пользователя политике, соответствующей конечной точке, разработано клиентское Android-приложение. Оно описано ниже. Таким образом, разрабатываемое веб-приложение должно иметь возможность обращаться к Android-приложению, передавая требуемую политику, а затем получать результат проведенных проверок.

Можно выделить следующие функциональные требования к веб-приложению:

1. Приложение должно иметь хранилище пользовательских данных.
2. Должна быть реализована поддержка работы с политиками доступа к конечным точкам веб-приложения.
3. Должна быть реализована возможность аутентификации пользователя в приложении по паролю с установкой аутентификационных Cookie.
4. В приложении должен быть реализован функционал регистрации FIDO2-аутентификатора для последующего его использования.
5. Доступ на страницу регистрации ключа должен быть запрещен для неавторизованных пользователей.
6. Если пользователь авторизован, то он не должен иметь доступ к странице входа в учетную запись.
7. Для авторизованных пользователей должна быть возможность выхода из учетной записи.
8. Приложение должно иметь возможность обращаться к клиентскому Android-приложению для nfc-подключения к электронному документу, проведения базового контроля доступа, пассивной и активной аутентификации, извлечения данных и проверки их соответствия требуемой сервером политике.

Помимо функциональных требований к приложению были выделены следующие требования к платформе, на которой ведется разработка:

1. Платформа должна поддерживать библиотеки, позволяющие серверу выступать в качестве доверенной стороны.
2. Платформа должна иметь возможности развертывания на Linux и контейнеризации, иметь необходимые компоненты.

Обоснование выбора платформ и средств разработки.

Фреймворк ASP.NET Core. В соответствии с разработанными требованиями в качестве платформы был выбран ASP.NET Core 8.0 – кроссплатформенный высокопроизводительный веб-фреймворк с открытым исходным кодом для создания современных приложений, подключенных к Интернету.

Фреймворк позволяет создавать как веб-API, в таком случае клиентская часть может быть реализована отдельно, так и приложение MVC, разрабатывая единое решение для пользовательского интерфейса и стороны сервера. Это удобно в данном случае, так как приложение должно поддерживать ввод и обработку запросов от пользователя, а также принимать и отправлять данные клиенту, реализуя спецификацию WebAuthn API.

ASP.NET Core содержит встроенный контейнер внедрения зависимостей, отвечающих за добавление того или иного функционала в веб-приложение. Эти зависимости также называются сервисами. В частности, для выполнения аутентификации фреймворк содержит сервис IAuthenticationService, который легко регистрируется на этапе конфигурации приложения с помощью метода builder.Services.AddAuthentication().

В контексте данной работы также важно, что ASP.NET Core предлагает встроенные средства авторизации и управления доступом, которые позволяют гибко настраивать политики безопасности. Во фреймворк встроены атрибуты для применения различных политик авторизации к отдельным контроллерам или действиям внутри контроллеров. Это дает возможность тонкой настройки требований к аутентификации и авторизации для каждой конечной точки приложения.

RSK FIDO2-компонент для ASP.NET Core. В качестве библиотеки, позволяющей разрабатываемому веб-приложению выступать в качестве проверяющей стороны, был выбран компонент, выпущенный компанией Rock Solid Knowledge (член альянса FIDO) для платформы ASP.NET Core 8.0⁴.

Разработанная RSK FIDO2-библиотека позволяет веб-приложению выступать в роли проверяющей стороны WebAuthn, предоставляя сервис IFidoAuthentication, содержащий методы для формирования запросов к клиенту, который взаимодействует с WebAuthn API. Библиотека представляет собой компонент, предназначенный для промышленной разработки, и для его использования требуется предоставление компанией официальной лицензии⁵.

Платформа Somee Web Hosting. Разработанное приложение было контейнеризовано (написаны соответствующие файлы Docker, Docker Compose) и развернуто на Somee Web Hosting.

Платформа позволяет загружать любое приложение и имеет возможность гибкой настройки серверной части. Somee позволяет упрощать и ускорять цикл разработки, снижает потребности в сложной работе с сервером, подходит для работы с нагруженными приложениями и быстрого масштабирования проектов.

Спецификация FIDO2 требует в процессе регистрации и аутентификации защищенное соединение между пользователем и клиентом. Поэтому для работы веб-приложения необходим SSL-сертификат – цифровой сертификат, удостоверяющий подлинность веб-сайта и позволяющий использовать зашифрованное соединение. Платформа Somee предоставляет возможности получения как бесплатного SSL-сертификата центра сертификации Let's Encrypt, так и сертификатов от сторонних поставщиков, таких как Comodo, Thawte или GeoTrust.

Взаимодействие с биометрическими документами, удостоверяющими личность. В биометрические документы встроена интегральная микросхема (чип), содержащая электронное средство биометрической идентификации с персональными данными владельца биометрического документа в соответствии с требованиями ICAO.

ICAO разработала LDS⁶ – стандартизированную структуру данных, которая обеспечивает глобальную интероперабельность и семантическую совместимость. LDS содержит информацию о владельце электронного идентификатора и подразделяется на 16 групп данных. Каждая группа данных содержит определенные личные сведения, необходимые для идентификации личности. Это позволяет хранить и обмениваться информацией о личности в стандартизированном формате, обеспечивая эффективный обмен данными. Например, первая группа данных включает код документа, государство или организацию выдачи, имя владельца документа (имя владельца обычно представляется в двух частях – основной определитель и вторичный определитель), дату рождения, гражданство, пол, дату истечения срока действия документа, контрольные цифры; вторая группа данных содержит глобально интероперабельные биометрические параметры, необходимые для подтверждения личности с помощью eID: кодировку(и) биометрических характеристик лица, количество записанных кодировок, биометрический подтип, дату и время создания, срок действия и др.

Использование идентификации eID способствует совместимости на нескольких уровнях, что делает ее эффективным и удобным средством для подтверждения личности и обмена информацией в цифровой среде.

Для обеспечения высокого уровня защиты и целостности данных, представленных на электронном документе, стандартом ICAO предусмотрены следующие протоколы защиты⁷:

1. Базовый контроль доступа (BAC – Basic Access Control) предназначен для гарантии того, что доступ к данным карты возможен только при физическом доступе к ней. При считывании документа требуется ввести изображенные на карте номер, дату рождения владельца и срок

⁴ RSK FIDO2 for ASP.NET documentation. – URL: <https://www.identityserver.com/documentation/fido2/> (date of access: 12.01.2025).

⁵ Компания Rock Solid Knowledge предоставила авторам статьи официальную лицензию на компонент RSK FIDO2 для ASP.NET Core.

⁶ Doc 9303 // ICAO. – URL: <https://www.icao.int/publications/pages/publication.aspx?docnum=9303> (date of access: 12.01.2025).

⁷ URL: <https://www.icao.int/publications/pages/publication.aspx?docnum=9303>

действия либо пин-код. На рис. 2, а изображена семантическая модель базового контроля доступа [7] с использованием спецификации SCg-кода [9].

2. Пассивная аутентификация позволяетчитывающему устройству определять подлинность данных eID с помощью объекта защиты документа, проверяя подпись эмитента и хеш-значение полученных данных. Семантическая модель пассивной аутентификации представлена в работе [7].

3. Активная аутентификация, или аутентификация с чипом, предназначена для защиты данных документа от изменения и клонирования. Для этого при издании электронного документа создается пара открытого и закрытого ключей. Открытый ключ передается с данными считающему устройству, а закрытый помещается в защищенное хранилище документа. На рис. 2, б с использованием SCg-кода изображена семантическая модель активной аутентификации.

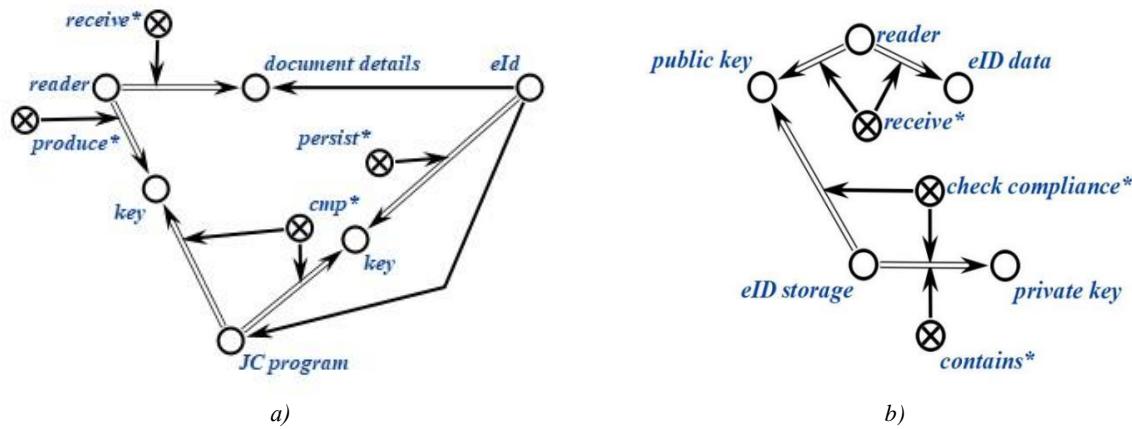


Рис. 2. Модель базового контроля доступа (а); активной аутентификации (б)

Fig. 2. Basic access control model (a); active authentication model (b)

Для целей разработки и тестирования компонента системы, обеспечивающего взаимодействие с биометрическими документами, компанией ЗАО «АВЕСТ» – ведущим производителем систем защиты электронного документооборота в Республике Беларусь – был предоставлен полнофункциональный макет электронного идентификатора гражданина Республики Беларусь. Макет содержал необходимые тестовые данные о владельце и поддерживал описанные выше протоколы защиты.

Требования к мобильному компоненту системы беспарольного доступа. В качестве клиента для разрабатываемой системы было выбрано мобильное приложение. На основе открытости и широкого спектра компонентов и средств разработки была выбрана операционная система Android. Большинство современных android-устройств оснащены NFC-считывателями. Следовательно, от разрабатываемого приложения необходимо требовать поддержку обработки NFC-событий и извлечения из них данных.

Приложению следует корректно и достаточно быстро считывать и декодировать (расшифровывать) данные, находящиеся на электронном идентификаторе. Оно должно вызываться из браузера и иметь возможность взаимодействовать с сервером, выступающим в роли проверяющей стороны. Приложение должно использовать защищенное хранилище на android-устройстве. Передача атрибутов пользователя на сервер должна производиться без раскрытия содержимого этих атрибутов на основе использования протоколов и методов с нулевым разглашением информации. Приложение должно соответствовать стандартам ICAO, определяющим требования к электронным документам, их обработке, аутентификации и безопасности, обеспечивая надежность и совместимость с международными стандартами и регуляторными нормами. Оно должно осуществлять проверку целостности и подлинности электронных документов, проводить пассивную и активную аутентификацию с целью предотвращения подделки, изменения и несанкционированного доступа к информации.

Основываясь на вышеперечисленных требованиях, для реализации мобильного приложения был выбран язык программирования Kotlin с использованием дополнительных открытых библиотек для работы с электронными идентификаторами SCUBA⁸, JMRTD⁹ в среде разработки Android Studio (рис. 3).

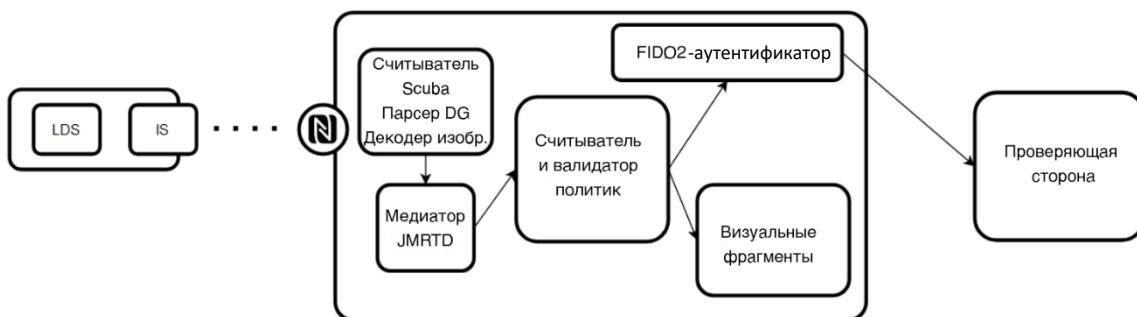


Рис. 3. Схема мобильного компонента системы

Fig. 3. Scheme of the mobile component of the system

Установка подключения к электронному идентификатору. Для прохождения базового контроля доступа приложению предоставляются номер электронного документа, срок его действия, дата рождения владельца.

Процесс извлечения данных из электронного документа состоит из нескольких шагов:

1. Приложение настраивается на обнаружение и реагирование на прикладывание электронного идентификатора к устройству.
2. После обнаружения eID приложение взаимодействует с NFC-чипом и считывает данные с использованием соответствующей технологии чтения NFC (в данном случае IsoDep).
3. После извлечения данных из электронного документа приложение выполняет базовый контроль доступа, сравнивая введенные пользователем данные (номер электронного документа, срок действия и дату рождения) с данными, полученными с электронного идентификатора. Если данные совпадают и проходят проверку, приложение может предоставить доступ к пользовательским атрибутам.
4. Выполняется пассивная и активная аутентификации.

Структура решения, полученного на основе интеграции eID-идентификации, FIDO2-аутентификации и атрибутивной авторизации. Структура решения с функциональной точки зрения изображена на рис. 4.

Для достижения модульности, гибкости и масштабируемости в приложении реализована трехуровневая архитектура (рис. 5). Она разделяет приложение на три основных уровня: представления, бизнес-логики и доступа к данным. Каждый уровень выполняет определенные функции и взаимодействует с другими уровнями для обеспечения эффективной работы всего приложения.

Для корректного взаимодействия между уровнями приложения и тестируемости использован встроенный в ASP.NET Core механизм внедрения зависимостей. При этом для каждой зависимости (например, сервиса или репозитория) определен интерфейс, а затем класс, который его реализует. На этапе конфигурирования приложения интерфейс и реализация регистрируются, после чего зависимость может быть получена из контейнера.

Для хранения учетных данных пользователей применяется СУБД MS SQL Server. Для взаимодействия веб-приложения с выбранной СУБД используется Entity Framework Core 6. EF Core оптимизирует взаимодействие с базой данных с помощью технологии «ленивой загрузки» (lazy loading) и отложенного выполнения запросов, что позволяет повышать производительность приложения, особенно при работе с большими объемами данных и высокой нагрузкой.

⁸SCUBA. – URL: <https://scuba.sourceforge.net/> (date of access: 12.01.2025).

⁹JMRTD. – URL: <https://jmrtd.org/> (date of access: 12.01.2025).



Рис. 4. Функциональная схема решения

Fig. 4. Functional scheme of the solution

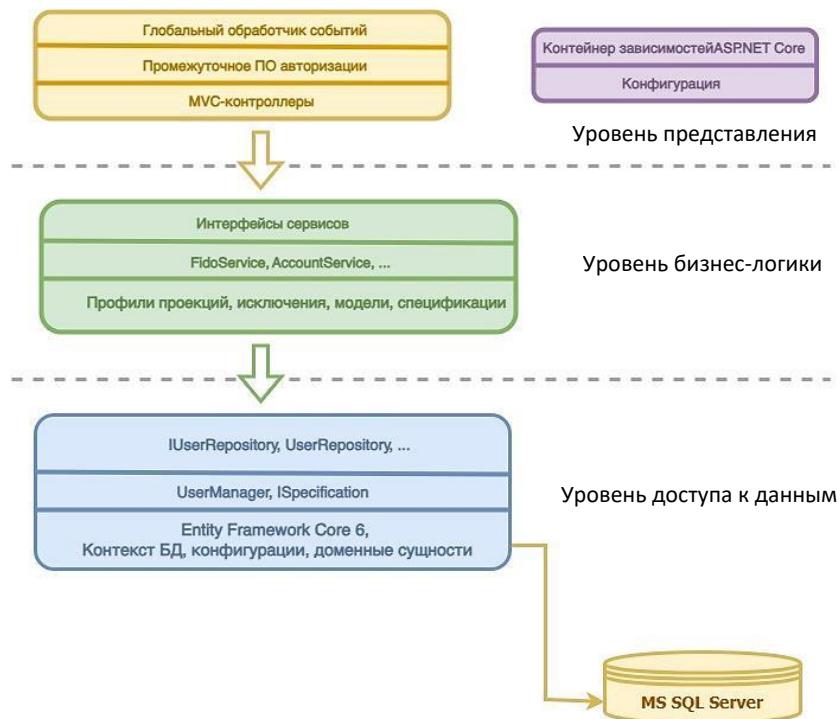


Рис. 5. Трехуровневая архитектура стороны сервера

Fig. 5. Three-tier server-side architecture

В базе данных не предусмотрено хранение чувствительных данных пользователя. Данные о году рождения, месте жительства и др. будут получаться при считывании данных с электронного документа. Обрабатываться на стороне клиента и отправляться стороне сервера они не будут.

Работа с сущностями и контекстом базы данных осуществляется через класс UserManager, что повышает защищенность данных о пользователе, так как пароль хранится в виде хеш-значения, а в качестве «соли» применяется SecurityStamp. При этом использование UserManager инкапсулировано с помощью паттерна «Репозиторий», что делает доступ к данным из уровня бизнес-логики более простым и удобным. Сам объект класса UserManager внедряется в репозиторий через конструктор.

Интерфейс ISpecification<T> и метод расширения ApplySpecification<T> дают возможность осуществлять поиск пользователей по различным критериям.

Уровень бизнес-логики содержит интерфейсы и реализации сервисов для работы с пользовательскими учетными записями, аутентификации и реализации FIDO2. Одним из основных сервисов здесь является FidoService.

На уровне бизнес-логики осуществляется работа не с конкретной реализацией интерфейса UserRepository, обеспечивающего регистрацию пользователей, обновление данных, выборку пользователей, аутентификацию с использованием пароля и т. д., а с интерфейсом IUserRepository, что дает возможность при необходимости сменить хранилище данных, а также облегчает тестируемость.

Уровень представления. За FIDO-аутентификацию и аутентификацию с помощью логина и пароля отвечают соответственно два основных контроллера: FidoController и AccountController. AccountController содержит get-метод Login, который отправляет пользователю страницу с формой, в которую нужно ввести логин и пароль. Post-метод Login принимает данные из заполненной формы, идентифицирует пользователя (осуществляет поиск по логину) и сравнивает пароль с имеющимся. FidoController отвечает за формирование запросов к клиенту для создания новых учетных данных и аутентификации. При этом сам контроллер не содержит в себе бизнес-логику, а только отвечает за запуск нужных методов соответствующих сервисов.

Реализация процесса регистрации FIDO2-ключа. Для регистрации новых учетных данных в FidoController реализованы три метода: StartRegistration(), Register() и CompleteRegistration(). Регистрация ключа может быть выполнена только для авторизованных пользователей. Действия контроллера заключаются в том, чтобы сформировать запрос к клиенту на создание новых учетных данных и после получения ответа проверить его и зарегистрировать (рис. 6).

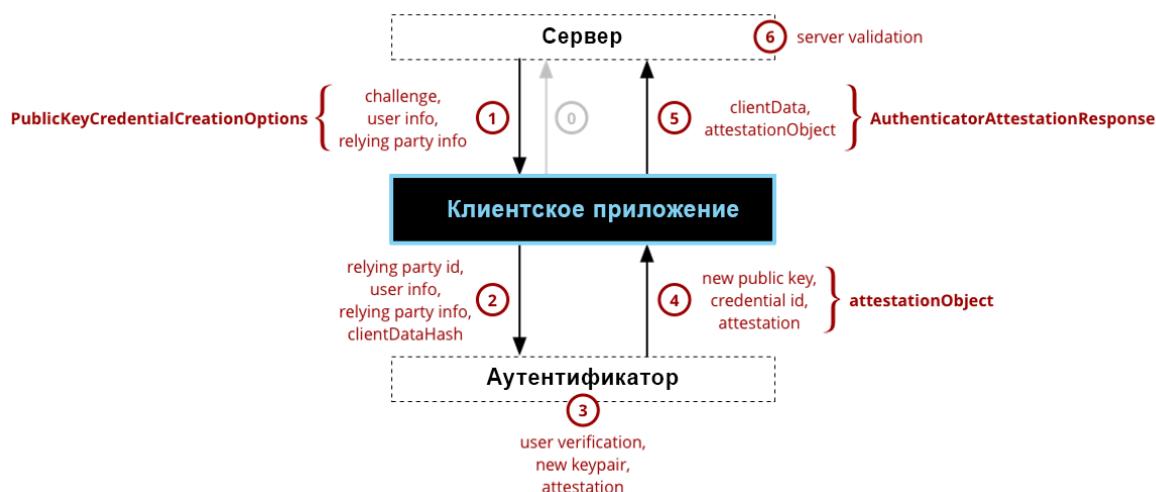


Рис. 6. Схема реализации процесса регистрации FIDO2-ключа

Fig. 6. Scheme of realization of the FIDO2-key registration process

Инициализация регистрации:

1. StartRegistration возвращает страницу с формой для ввода имени ключа.
2. Введенное имя передается в Register(), где идет обращение к сервису для работы с FIDO.

3. Генерируется challenge, создается объект запроса, содержащий данные о пользователе, проверяющей стороне и сам буфер challenge.

4. Запрос кодируется в base64 и передается в представление Register.

Обработка на клиенте:

1. JavaScript-скрипт на странице регистрации запускается при нажатии submit.

2. Декодируется challenge, формируются структуры:

grp (проверяющая сторона) – указывает на домен сервера;

user – содержит данные пользователя (userId, displayName, userHandle);

pubKeyCredParams – указывает на приоритетные типы аутентификаторов (EC SHA-256, RSA SHA-256);

authenticatorSelection – указывает на использование платформенных ключей;

attestation – значение параметра (none) означает, что сервер не требует аттестации;

timeout – указывает время на выполнение аутентификации.

3. Объект publicKey передается в функцию navigator.credentials.create (является частью WebAuthn API).

Аутентификация и отправка данных:

1. Аутентификатор запрашивает подтверждение пользователя.

2. После успешной верификации возвращается PublicKeyCredentials, содержащий id, rawId, response (с clientDataJSON и attestationObject).

3. Данные кодируются в base64 и отправляются серверу.

Завершение регистрации:

1. Контроллер Fido CompleteRegistration завершает этап регистрации.

2. Сервер проверяет и регистрирует учетные данные.

3. Данные хранятся в InMemoryKeyStore в памяти сервера, возможно подключение внешней базы данных.

Семантическая модель регистрации FIDO2-ключа была представлена авторами ранее в работе [10].

Реализация процесса аутентификации пользователя с FIDO2-ключом. На стороне сервера процесс аутентификации пользователя в контроллере Fido представлен двумя методами: Login и CompleteLogin. Действия контроллера заключаются в том, чтобы проверить, существуют ли требуемые учетные данные в хранилище, и сформировать запрос к клиенту. После этого, получив ответ, содержащий подпись, проверить его и подтвердить аутентификацию пользователя.

Разработанное приложение было развернуто на сервисе Somee, также был установлен SSL-сертификат, предоставляемый платформой. Его работоспособность была проверена с помощью устройств на Windows, Android и iOS. При этом использовались браузеры Chrome, Opera, Edge, Mozilla и Safari.

Передача политики клиентской стороне. Важным вопросом является формат, в котором политика будет передаваться от сервера клиенту. Выбранный формат должен поддерживать отношения «больше», «меньше», «равно», например «возраст больше 25», а также конъюнкцию и дизъюнкцию условий. В качестве такого формата был выбран OData¹⁰. OData (Open Data Protocol) – открытый протокол, определяющий стандартные способы запроса и изменения данных через REST API, а также стандартный формат для запросов и ответов, который основан на широко используемых протоколах и форматах данных. Формат запроса OData можно использовать для передачи политики конечной точки, к которой пользователь запрашивает доступ, приложению, взаимодействующему с электронным документом.

Таким образом, для того чтобы на странице веб-приложения открыть Android-приложение для взаимодействия с электронным документом, а также передать необходимую политику, JavaScript-код страницы должен направить запрос по адресу вида `app://fido-abac-eid-demo.com/check-policy?$filter=age gt 35 and contains(nationality, 'USA')`.

¹⁰Документация. – URL: <https://www.odata.org/documentation/> (дата обращения: 12.01.2025).

Конфигурация разработанного приложения. После добавления библиотеки Rsk.AspNetCore.Fido нужно подключить соответствующий сервис с помощью метода builder.Services.AddFido(), в который требуется передать лицензионный ключ и тип лицензии.

Лицензионный ключ является конфиденциальной информацией и, следовательно, требует более безопасного хранения, чем стандартный файл appsettings.json, являющийся частью большинства ASP.NETCore-приложений. При развертывании на предприятии распространенным решением является использование хранилища Azure Key Vault, созданного для этих целей. На этапе же разработки целесообразно использовать хранилище секретов Windows (User Secrets).

Для возможности аутентификации с помощью логина и пароля следует подключить аутентификацию на основе Cookie. Для этого вызывается метод builder.Services.AddAuthentication(), с помощью параметра устанавливается использование схемы аутентификации на основе Cookie. Также на данном этапе указывается, что приложение будет использовать MVC.

Демонстрация разработанного веб-приложения. Рассмотрим основные элементы полученной реализации. Веб-приложение представляет собой набор конечных точек (страниц и ресурсов), также для удобства навигации реализованы навигационная панель и подвал.

После парольной аутентификации пользователь переходит на домашнюю (тестовую) страницу, содержащую информацию об электронных идентификаторах. На тестовой странице размещены ссылки, ведущие к двум защищенным ресурсам, доступ к которым ограничен разными политиками. Первый требует от пользователя возраст более 18 лет или гражданство Республики Беларусь. Для доступа ко второму необходимо иметь возраст более 35 лет и гражданство Соединенных Штатов Америки.

При переходе по каждой из ссылок запускается Android-приложение, в которое передается политика соответствующего ресурса. Приложение проверяет подлинность личности и соответствие пользовательских атрибутов требуемой политике. Результат проверки возвращается на веб-ресурс.

На рис. 7 изображен снимок экрана мобильного приложения после считывания данных с тестовой eID-карты, установки личности с использованием FIDO2-аутентификации и проверки атрибутов пользователя, а также страница в веб-приложении с результатом авторизации.

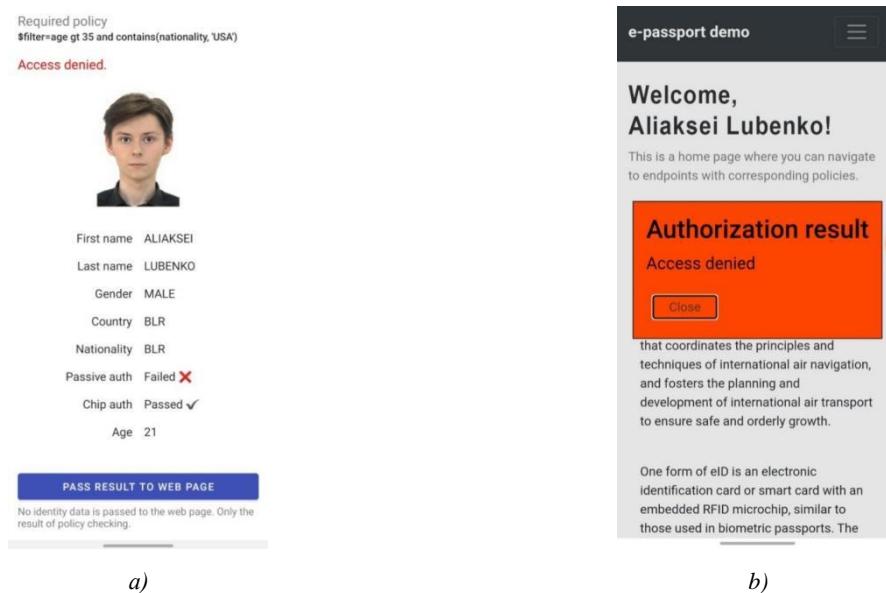


Рис. 7. Результат проверки подлинности личности и соответствия пользовательских атрибутов требуемой политике доступа к ресурсу на основании тестовой eID-карты Республики Беларусь:
a) экран мобильного устройства с NFC-считывателем; b) веб-приложение

Fig. 7. Result of identity authentication and compliance of user attributes with the required resource access policy based on the test eID-card of the Republic of Belarus: a) mobile device screen with NFC reader; b) web application

Можно заметить, что пользователю отказано в доступе. В первую очередь он не соответствует требуемой политике (указана в верхней части снимка экрана). Также не была пройдена пассивная аутентификация. Дело в том, что на тестовой eID-карте с данными одного из разработчиков отсутствует подпись издателя (если бы она была, это был бы настоящий паспорт). Поэтому при работе с тестовой картой была исключена пассивная аутентификация.

Пример с прохождением пассивной аутентификации на основании подлинной eID-карты Республики Беларусь здесь не приводится из-за необходимости защиты персональных данных.

Заключение. В настоящей публикации авторы представили исследование по разработке и первоначальной оценке прототипа системы управления доступом к информационным ресурсам путем аутентификации по спецификации FIDO2 и модели управления доступом на основе атрибутов. При этом в качестве источника пользовательских атрибутов применяются средства электронной идентификации, удовлетворяющие стандартам Международной организации гражданской авиации, в том числе eID-карта Республики Беларусь.

Разработанное приложение было контейнеризовано и развернуто на онлайн-сервере, его работоспособность проверена с различных платформ с помощью распространенных браузеров.

В форме SCn- и SCg-кода технологии OSTIS представлены основные понятия и концепции, связанные с FIDO2-аутентификацией, предложена семантическая сеть, описывающая взаимодействие сервера, клиента и аутентификатора при регистрации WebAuthn-ключа и аутентификации. Это позволяет обеспечить гибкий и масштабируемый подход к аутентификации на основе онтологии.

Показано, что разработанная система безопасности способствует совместимости с информационными системами на нескольких уровнях, в том числе техническом и семантическом. Предполагается, что построенная система должна обладать гибкостью, масштабируемостью и интероперабельностью вследствие поддержки стандартов ICAO и будет адаптирована к использованию в корпоративных и муниципальных системах.

Вклад авторов. *А. А. Жидович* внес существенный вклад в реализацию работы и написание текста статьи; *А. А. Лубенько* реализовал мобильный компонент системы беспарольного доступа, участвовал в написании текста; *И. С. Войтешенко* обосновал концепцию и актуальность статьи, осуществил анализ и интерпретацию результатов работы, проанализировал содержание текста, утвердил окончательный вариант статьи.

Список использованных источников

1. Angelogianni, A. How many FIDO protocols are needed? Analysing the technology, security and compliance / A. Angelogianni, I. Politis, C. Xenakis // ACM Computing Surveys. – 2024. – Vol. 56, iss. 8. – P. 1–51. – DOI: 10.1145/3654661.
2. Challenges with Passwordless FIDO2 in an Enterprise Setting: A Usability Study / M. Kepkowski, M. Machulakian, I. Wood, D. Kaafar. – URL: <https://arxiv.org/abs/2308.08096> (date of access: 08.11.2024).
3. FIDO2 passwordless authentication for remote devices / S. Dixit, A. Gupta, R. Jain [et al.] // Networks and Systems in Cybernetics : Proc. of 12th Computer Science On-line Conf. 2023. – Springer, 2023. – Vol. 2. – P. 349–362. – DOI: 10.1007/978-3-031-35317-8_32.
4. Hoefling, D. Understanding How FIDO Makes Passwordless Authentication Possible / D. Hoefling. – URL: <https://practical365.com/understanding-how-fido-makes-passwordless-authentication-possible/> (date of access: 27.11.2024).
5. Fast IDentity online with anonymous credentials (FIDO-AC) / W.-Z. Yeoh, M. Kepkowski, G. Heide [et al.] // Proc. of the 32nd USENIX Security Symp., Anaheim, CA, USA, 9–11 Aug. 2023. – Anaheim, 2023. – P. 3029–3046.
6. Brodsky, Z. Using MITM to bypass FIDO2 phishing-resistant protection / Z. Brodsky. – URL: <https://www.silverfort.com/blog/using-mitm-to-bypass-fido2/> (date of access: 08.11.2024).
7. Zhidovich, A. Semantic notation of access control technology based on eID identification, FIDO2-authentication and attribute-based authorization in digital environment / A. Zhidovich, A. Lubenko, I. Vojteshenko // Open Semantic Technologies for Intelligent Systems. – 2024. – No. 8. – P. 371–376.

8. Garoffolo, A. Zendoo: A ZK-SNARK verifiable cross-chain transfer protocol enabling decoupled and decentralized sidechains / A. Garoffolo, D. Kaidalov, R. Oliynykov // 2020 IEEE 40th Intern. Conf. on Distributed Computing Systems (ICDCS), Singapore, 29 Nov. – 01 Dec. 2020. – Singapore, 2020. – P. 1257–1262.
9. Технология комплексной поддержки жизненного цикла семантически совместимых интеллектуальных компьютерных систем нового поколения / под общ. ред. В. В. Голенкова. – Минск : Бестпринт, 2023. – 1064 с.
10. Semantic approach to designing applications with passwordless authentication according to the FIDO2 specification / A. Zhidovich, A. Lubenko, I. Vojteshenko, A. Andrushevich // Open Semantic Technologies for Intelligent Systems. – 2023. – No. 7. – P. 311–316.

References

1. Angelogianni A., Politis I., Xenakis C. How many FIDO protocols are needed? Analysing the technology, security and compliance. *ACM Computing Surveys*, 2024, vol. 56, iss. 8, pp. 1–51. DOI: 10.1145/3654661.
2. Kepkowski M., Machulakian M., Wood I., Kaafar D. Challenges with Passwordless FIDO2 in an Enterprise Setting: A Usability Study. Available at: <https://arxiv.org/abs/2308.08096> (accessed 08.11.2024).
3. Dixit S., Gupta A., Jain R., Joshi R., Gonge S., Kotecha K. FIDO2 passwordless authentication for remote devices. *Networks and Systems in Cybernetics: Proceedings of 12th Computer Science On-line Conference 2023*. Springer, 2023, vol. 2, pp. 349–362. DOI: 10.1007/978-3-031-35317-8_32.
4. Hoefling D. Understanding How FIDO Makes Passwordless Authentication Possible. Available at: <https://practical365.com/understanding-how-fido-makes-passwordless-authentication-possible/> (accessed 27.11.2024).
5. Yeoh W.-Z., Kepkowski M., Heide G., Kaafar D., Hanzlik L. Fast IDentity online with anonymous credentials (FIDO-AC). *Proceedings of the 32nd USENIX Security Symposium, Anaheim, CA, USA, 9–11 Aug. 2023*. Anaheim, 2023, pp. 3029–3046.
6. Brodsky Z. Using MITM to bypass FIDO2 phishing-resistant protection. Available at: <https://www.silverfort.com/blog/using-mitm-to-bypass-fido2/> (accessed 08.11.2024).
7. Zhidovich A., Lubenko A., Vojteshenko I. Semantic notation of access control technology based on eID identification, FIDO2-authentication and attribute-based authorization in digital environment. *Open Semantic Technologies for Intelligent Systems*, 2024, no. 8, pp. 371–376.
8. Garoffolo A., Kaidalov D., Oliynykov R. Zendoo: A ZK-SNARK verifiable cross-chain transfer protocol enabling decoupled and decentralized sidechains. *2020 IEEE 40th International Conference on Distributed Computing Systems (ICDCS), Singapore, 29 November – 01 December 2020*. Singapore, 2020, pp. 1257–1262.
9. Golenkov V. V. (ed.). Tehnologija kompleksnoj podderzhki zhiznennogo cikla semanticheski sovmestimyh intellektual'nyh kompjuternyh sistem novogo pokolenija. *Technology for Integrated Life Cycle Support of Semantically Compatible Intelligent Computer Systems of the New Generation*. Minsk, Bestprint, 2023, 1064 p.
10. Zhidovich A., Lubenko A., Vojteshenko I., Andrushevich A. Semantic approach to designing applications with passwordless authentication according to the FIDO2 specification. *Open Semantic Technologies for Intelligent Systems*, 2023, no. 7, pp. 311–316.

Информация об авторах

Жидович Антон Андреевич, выпускник кафедры технологий программирования факультета прикладной математики и информатики, Белорусский государственный университет.

E-mail: anton.zhidovich@gmail.com

Лубенко Алексей Анатольевич, выпускник кафедры технологий программирования факультета прикладной математики и информатики, Белорусский государственный университет.

E-mail: alexeilubenko02@gmail.com

Войтешенко Иосиф Станиславович, кандидат технических наук, доцент, доцент кафедры технологий программирования факультета прикладной математики и информатики, Белорусский государственный университет.

E-mail: voit@bsu.by

<https://orcid.org/0000-0002-0134-1793>

Information about the authors

Anton A. Zhidovich, Graduate of the Department of Programming Technologies, Faculty of Applied Mathematics and Informatics, Belarusian State University. E-mail: anton.zhidovich@gmail.com

Alexei A. Lubenko, Graduate of the Department of Programming Technologies, Faculty of Applied Mathematics and Informatics, Belarusian State University. E-mail: alexeilubenko02@gmail.com

Iosif S. Vojteshenko, Ph. D. (Eng.), Assoc. Prof., Assoc. Prof. of the Department of Programming Technologies of the Faculty of Applied Mathematics and Informatics, Belarusian State University.

E-mail: voit@bsu.by

<https://orcid.org/0000-0002-0134-1793>

Правила для авторов

Редакция журнала «Информатика» просит авторов руководствоваться приведенными ниже правилами.

I. Статьи принимаются в редакцию через электронную систему подачи по адресу <http://inf.grid.by> в формате файлов текстовых редакторов Microsoft Word. Объем оригинальной статьи – от 8 до 16 стр., включая рисунки, таблицы и достаточное количество наиболее актуальных ссылок; объем обзорной статьи – от 16 до 32 стр., включая все основные ссылки. Текст набирается с переносами, шрифт Times New Roman 11 пт, интервал между строками одинарный, абзацный отступ 0,5 см, поля по 2,5 см со всех сторон.

Материал статьи должен быть четко структурированным: Введение; основные разделы, в которых изложены цели и задачи, методы, результаты; Заключение (выводы).

II. Статьи о результатах работ, проведенных в научных учреждениях, должны иметь разрешение на публикацию (сопроводительное письмо за подписью руководителя или выписку из заседания ученого совета, отдела или кафедры, акт экспертизы).

III. Статьи в обязательном порядке должны включать аннотацию, ключевые слова, список литературы, информацию об авторах на русском и английском языках.

На заглавной странице располагаются следующие метаданные:

1. Индекс по универсальной десятичной классификации (УДК); на русском и английском языках тип статьи (оригинальная или обзорная), название статьи, инициалы и фамилии всех авторов, полное наименование учреждений, где работают авторы, с указанием почтового адреса, при наличии указывается ученая степень и ORCID, e-mail ответственного лица.

2. Аннотация (Abstract) объемом 150–250 слов в оригинальной статье должна быть структурирована отдельными подразделами: Цели, Методы, Результаты, Заключение, а также максимально характеризовать содержательную часть рукописи. Сюда не следует включать впервые введенные термины, аббревиатуры (за исключением общезвестных), ссылки на литературу.

3. Ключевые слова (Keywords) – наиболее значимые слова или словосочетания по теме работы, отражающие специфику темы, объекты и результаты исследования; перечень ключевых слов должен содержать 5–10 слов.

4. В разделе Благодарности (Acknowledgements) указываются все источники финансирования исследования, а также благодарности людям, которые участвовали в работе над статьей.

5. Автор обязан уведомить редакцию о реальном или потенциальном конфликте интересов, включив информацию в раздел Конфликт интересов (Conflict of interest).

6. Формулы, рисунки, таблицы в статье нумеруются в соответствии с порядком их упоминания в тексте. Ссылки на рисунки и таблицы в тексте обязательны. Рисунки должны быть выполнены с хорошим разрешением в масштабе, позволяющем четко различать надписи и обозначения. Цветные иллюстрации печатаются только в том случае, когда это необходимо для понимания излагаемого материала. Подрисуночные подписи с расшифровкой всех позиций, представленных на рисунке, и названия таблиц набираются шрифтом гарнитуры основного текста размером 9 пт. Перевод подрисуночной подписи и пояснений к рисунку, а также перевод названия таблицы, заголовки строк или столбцов располагаются курсивом после русскоязычной версии.

7. Набор формул выполняется в формульном редакторе Microsoft Equation или Math Type. Прямыми 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).

8. Список использованной литературы оформляется в соответствии с требованиями Высшей аттестационной комиссии Республики Беларусь (ГОСТ 7.5–2008). Номер литературной ссылки в тексте дается порядковым номером в квадратных скобках. Ссылаясь на неопубликованные работы не допускается.

9. 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.

10. 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).

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

IV. 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. После проверки на соответствие правилам статья направляется рецензенту с указанием сроков рецензирования.

V. При наличии замечаний рецензента автору предоставляется определенное время на доработку рукописи. Статьи, направляемые на доработку, должны быть возвращены в исправленном виде с ответами на все замечания. Окончательное решение о публикации или отклонении рукописи принимается редакцией журнала. При положительном заключении рецензента статья передается редактору для подготовки к печати. Редакция оставляет за собой право на редакционные изменения, не искажающие основное содержание статьи.

VI. Редакция журнала предоставляет возможность первоочередного опубликования статей, представленных лицами, которые осуществляют послевузовское обучение (аспирантура, докторантуре, соискательство) в год завершения обучения.

VII. Авторы несут ответственность за направление в редакцию статей, уже опубликованных ранее или принятых к публикации другими изданиями.

ИНДЕКСЫ

00827

для индивидуальных
подписчиков

008272

для предприятий
и организаций

2 | 2025

Том ■ Vol. 22

Информатика ■ INFORMATICS

ISSN 1816-0301 (Print)



9 771816 030000