ISSN 1729-7648 (Print)
ISSN 2708-0382 (Online)

ДОКЛАДЫ БГУИР

Том 23, № 3, 2025

Научный журнал выходит шесть раз в год

Учредитель
Учреждение образования «Белорусский государственный университет информатики и радиоэлектроники»

Журнал включен в Перечень ВАК Республики Беларусь, а также в базы данных:
CNKI, Dimensions, DOAJ, Google Scholar, Mendeley, Open Alex, НЭБ «КиберЛенинка», РИНЦ

СОДЕРЖАНИЕ

Ковальчук Н. С., Пилипенко В. А., Соловьёв Я. А. Влияние импульсной фотонной обработки в среде азота на оптические и электрофизические характеристики слоев двуокиси кремния и ее границы с кремнием 5

Махмутов Р. Т., Гапоненко Н. В., Лашковская Е. И., Меледин К. И., Кашко И. А., Малышев А. Д., Судник Л. В. Электрофизические свойства структур с пленками титаната бария и оксида алюминия, сформированными золь-гель-методом на титане 12

Туровец У. Е. Моделирование оптических процессов в тонкопленочном ИК-светодиоде на основе коллоидных квантовых точек PbS 19

Соловьев А. А., Певцов Е. Ф., Колчужин В. А. Системные математические макромодели логических МЭМС-вентилей 26

Мордачев В. И., Ционенко Д. А. Анализ электромагнитного фона, создаваемого созвездиями низкоорбитальных спутников связи, на основе прогноза образуемой ими территориальной плотности трафика 36

Шиманский Н. А., Баглов А. В., Мусаев Х. Б., Рузимурадов О. Н., Хорошко Л. С. Экспресс-анализ структурных и электронных свойств наноматериалов методами больших данных, больших языковых моделей и генеративного искусственного интеллекта 46

Мигалевич С. А., Герман Ю. О., Герман О. В. Метод автоматической генерации вопросов и ответов для систем тестирования знаний 54

Gasimov V. A., Zeynalli-Huseynzade L. R., Zahidova G. A. A Comparative Analysis of QoS, GoS, and QoE Metrics in Network Performance Evaluation (**Касумов В. А., Зейналлы-Гусейнзаде Л. Р., Захидова Г. А.** Сравнительный анализ показателей QoS, GoS и QoE при оценке производительности сети) 62

Усатов А. А., Недзьведь А. М., Го Цзижань. Оценка сходства между наборами данных с помощью векторных представлений 70

Петкевич М. Н., Коровко Д. В., Емельяненко Е. В. Математическая модель респираторных искажений патологического очага при визуализации на ПЭТ/КТ-изображениях 77

Вишняков В. А., Чуюэ Юй. Умный контракт блокчейна для поддержки терапии неврологических больных 86

Чиркова И. Н., Тумилович М. В. Влияние положения болюса на величину дозы ионизирующего излучения при моделировании облучения на линейном ускорителе 95

Сидоренко А. В., Высоцкая Е. А. Алгоритм шифрования информации с применением дискретной квантовой карты 102

Главный редактор
Вадим Анатольевич Богуш,
д. ф.-м. н., профессор, ректор Белорусского государственного университета информатики и радиоэлектроники (Минск, Беларусь)

Редакционная коллегия

Лыньков Л. М., д. т. н., профессор, Белорусский государственный университет информатики и радиоэлектроники, Минск, Беларусь – заместитель главного редактора

Стемпицкий В. Р., к. т. н., доцент, Белорусский государственный университет информатики и радиоэлектроники, Минск, Беларусь – заместитель главного редактора

Бойправ О. В., к. т. н., доцент, Белорусский государственный университет информатики и радиоэлектроники, Минск, Беларусь – ответственный секретарь редколлегии

Батура М. П., д. т. н., профессор, Белорусский государственный университет информатики и радиоэлектроники, Минск, Беларусь

Борисенко В. Е., д. ф.-м. н., профессор, Белорусский государственный университет информатики и радиоэлектроники, Минск, Беларусь

Гасанов М. Г., д. т. н., профессор, Азербайджанский технический университет, Баку, Азербайджан

Козлов С. В., д. т. н., профессор, Белорусский государственный университет информатики и радиоэлектроники, Минск, Беларусь

Лабунов В. А., акад. НАН Беларуси, д. т. н., профессор, Белорусский государственный университет информатики и радиоэлектроники, Минск, Беларусь

Писецкий Ю. В., д. т. н., доцент, Ташкентский университет информационных технологий имени Мухаммада ал-Хоразмий, Ташкент, Узбекистан

Сигов А. С., акад. РАН, д. ф.-м. н., профессор, МИРЭА – Российский технологический университет, Москва, Россия

Смирнов А. Г., д. т. н., профессор, Белорусский государственный университет информатики и радиоэлектроники, Минск, Беларусь

Татур М. М., д. т. н., профессор, Белорусский государственный университет информатики и радиоэлектроники, Минск, Беларусь

Цветков В. Ю., д. т. н., профессор, Белорусский государственный университет информатики и радиоэлектроники, Минск, Беларусь

Ответственный секретарь Т. В. Мироненко

Издание зарегистрировано в Министерстве информации Республики Беларусь, рег. номер 1087
Журнал издается с 2003 г. Дата последней перерегистрации 23 января 2010 г.

Подписано в печать 18.06.2025. Формат бумаги 60×84⅛. Бумага офисная. Отпечатано на ризографе. Гарнитура Таймс.
Усл. печ. л. 12,79. Уч.-изд. л. 11. Тираж 80 экз. Заказ 119.

Адрес редакции:
Белорусский государственный университет информатики и радиоэлектроники
220013, г. Минск, ул. П. Бровки, 6, ком. 329а
Тел.: +375 17 293-88-41
doklady@bsuir.by; http//:doklady.bsuir.by

Отпечатано в БГУИР. ЛП № 02330/264 от 24.12.2020.
220013, г. Минск, ул. П. Бровки, 6

Индекс для индивидуальной подписки 00787. Индекс для ведомственной подписки 007872

© УО «Белорусский государственный университет информатики и радиоэлектроники», оригинал-макет, оформление, 2025

ISSN 1729-7648 (Print)
ISSN 2708-0382 (Online)

DOKLADY BGUIR

V. 23, No 3, 2025

The scientific journal is published 6 times a year

Founder
Educational Establishment "Belarusian State University of Informatics and Radioelectronics"

The Journal is included in the List of Scientific Editions in the Republic of Belarus,
as well as the following databases:
CNKI, Dimensions, DOAJ, Google Scholar, Mendeley, Open Alex, CyberLeninka, RSCI

CONTENTS

Kovalchuk N. S., Pilipenko V. A., Solovjov Ja. A. Effect of Pulsed Photon Processing in Nitrogen Ambient on Optical and Electrophysical Characteristics of Silicon Dioxide Layers of Its Boundaries with Silicon 5

Makhmutov R. T., Gaponenko N. V., Lashkovskaya E. I., Meledin K. I., Kashko I. A., Malyshev A. D., Sudnik L. V. Electrophysical Properties of Structures with Barium Titanate and Oxide Films Formed by the Sol-Gel Method on Titanium 12

Turavets U. Y. Modeling of Optical Processes in Thin-Film IR Light-Emitting Diode Based on Colloidal PbS Quantum Dots..... 19

Solovev A. A., Pevtsov E. F., Kolchuzhin V. A. System-Level Mathematical Macro-Models of Logic MEMS Gates 26

Mordachev V. I., Tsyanenka D. A. Analysis of the Electromagnetic Background Created by Constellations of Low-Earth Orbit Satellites, Based on the Prediction Forecast of the Created Area Traffic Capacity 36

Shymanski N. A., Baglov A. V., Musaev Kh. B., Ruzimuradov O. N., Khoroshko L. S. Express Analysis of the Structural and Electronic Properties of Nanomaterials Using Big Data, Large Language Models and Generative Artificial Intelligence 46

Migalevich S. A., German Ju. O., German O. V. Method of Automatic Generation of Questions and Answers for Knowledge Testing Systems..... 54

Gasimov V. A., Zeynalli-Huseynzade L. R., Zahidova G. A. A Comparative Analysis of QoS, GoS, and QoE Metrics in Network Performance Evaluation 62

Usatoff A. A., Nedzved A. M., Guo Jiran. Assessing Similarity Between Datasets Using Vector Representations..... 70

Piatkevich M. N., Korovko D. V., Emelyanenko E. V. Mathematical Model of Respiratory Distortions of a Pathological Focus During Visualization on PET/CT Images..... 77

Vishniakou U. A., Chuyue Yu. Smart Contract Blockchain for Supporting Therapy of Neurological Patients..... 86

Chirkova I. N., Tumilovich M. V. Effect of Bolus Position on the Dose of Ionizing Radiation in a Linear Accelerator Simulation 95

Sidorenko A. V., Vysotskaya E. A. Information Encryption Algorithm Using a Discrete Quantum Map 102

Editor-in-Chief
Vadim A. Bogush, Dr. Sci. (Phys. and Math.), Professor,
Rector of the Belarusian State University
of Informatics and Radioelectronics (Minsk, Belarus)

Editorial Board

Leonid M. Lynkov, Dr. Sci. (Tech.), Professor, Belarusian State University of Informatics and Radioelectronics, Minsk, Belarus – Deputy Editor-in-Chief

Viktor R. Stempitsky, Cand. Sci. (Tech.), Associate Professor, Belarusian State University of Informatics and Radioelectronics, Minsk, Belarus – Deputy Editor-in-Chief

Olga V. Boiprav, Cand. Sci. (Tech.), Associate Professor, Belarusian State University of Informatics and Radioelectronics, Minsk, Belarus – Responsible Secretary of the Editorial Board

Mikhail P. Batura, Dr. Sci. (Tech.), Professor, Belarusian State University of Informatics and Radioelectronics, Minsk, Belarus

Victor E. Borisenko, Dr. Sci. (Phys. and Math.), Professor, Belarusian State University of Informatics and Radioelectronics, Minsk, Belarus

Mehman H. Hasanov, Dr. Sci. (Tech.), Professor, Azerbaijan Technical University, Baku, Azerbaijan

Sergei V. Kozlov, Dr. Sci. (Tech.), Professor, Belarusian State University of Informatics and Radioelectronics, Minsk, Belarus

Vladimir A. Labunov, Acad. of the NAS of Belarus, Dr. Sci. (Tech.), Professor, Belarusian State University of Informatics and Radioelectronics, Minsk, Belarus

Yuri V. Pisetsky, Dr. Sci. (Tech.), Associate Professor, Tashkent University of Information Technologies named after Muhammad al-Khwarizmi, Tashkent, Uzbekistan

Alexander S. Sigov, Acad. of the RAS, Dr. Sci. (Phys. and Math.), Professor, MIREA – Russian Technological University, Moscow, Russia

Alexander G. Smirnov, Dr. Sci. (Tech.), Professor, Belarusian State University of Informatics and Radioelectronics, Minsk, Belarus

Mikhail M. Tatur, Dr. Sci. (Tech.), Professor, Belarusian State University of Informatics and Radioelectronics, Minsk, Belarus

Viktor Yu. Tsvetkov, Dr. Sci. (Tech.), Professor, Belarusian State University of Informatics and Radioelectronics, Minsk, Belarus

Responsible Secretary T. Mironenka

The journal is registered in the Ministry of Information of the Republic of Belarus, Reg. No 1087.
The journal has been published since 2003. Last re-registration date January 23, 2010.

Signed for printing 18.06.2025. Format 60×84 ⅛. Office paper. Printed on a risograph. Type face Times.
Ed.-pr. l. 12,79. Ed.-ed. l. 11. Edition 80 copies. Order 119.

Editorial address
Belarusian State University of Informatics and Radioelectronics
6, P. Brovki St., Off. 329a, 220013, Minsk
Tel.: +375 17 293-88-41
doklady@bsuir.by; http//:doklady.bsuir.by

Printed in BSUIR. License LP No 02330/264 from 24.12.2020.
6, P. Brovki St., Room 329a, 220013, Minsk

Index for individual subscription 00787. Index for departmental subscription 007872

© Educational Establishment "Belarusian State University of Informatics and Radioelectronics", original layout, design, 2025

CC BY

http://dx.doi.org/10.35596/1729-7648-2025-23-3-5-11

УДК 621.382

ВЛИЯНИЕ ИМПУЛЬСНОЙ ФОТОННОЙ ОБРАБОТКИ В СРЕДЕ АЗОТА НА ОПТИЧЕСКИЕ И ЭЛЕКТРОФИЗИЧЕСКИЕ ХАРАКТЕРИСТИКИ СЛОЕВ ДВУОКИСИ КРЕМНИЯ И ЕЕ ГРАНИЦЫ С КРЕМНИЕМ

Н. С. КОВАЛЬЧУК, В. А. ПИЛИПЕНКО, Я. А. СОЛОВЬЁВ

ОАО «ИНТЕГРАЛ» – управляющая компания холдинга «ИНТЕГРАЛ» (Минск, Республика Беларусь)

Аннотация. Методами инфракрасной Фурье-спектрометрии, спектральной эллипсометрии, времяпролетной масс-спектроскопии вторичных ионов, исследованиями вольт-амперных и вольт-фарадных характеристик установлено влияние импульсной фотонной обработки в среде азота некогерентным потоком излучения от кварцевых галогенных ламп, направленным на нерабочую сторону пластины, обеспечивающей нагрев до 1150 °С за время около 7 с, на оптические свойства слоев двуокиси кремния толщиной 17,7 нм, сформированных пирогенным окислением кремния, легированного бором с ориентацией (100), и на электрофизические характеристики границы раздела ее с кремнием. Выявлено, что импульсная фотонная обработка слоев двуокиси кремния приводит к уплотнению и перестройке ее структуры, а также к образованию в двуокиси кремния связей Si–N, обеспечивающих нитридизацию SiO_2. На это указывают сдвиг, уменьшение полуширины и напряжения основной полосы поглощения Si–O-связи, уменьшение коэффициента преломления с 1,48 до 1,47 и увеличение толщины слоя до 18,2 нм. Показано, что нитридизация слоев SiO_2 при импульсной фотонной обработке в атмосфере азота приводит к уменьшению тока утечки диэлектрика в четыре раза и плотности его заряда в 3,43 раза. Это соответственно в 2,19 и 3,01 раза больше, чем при обработке в естественных атмосферных условиях за счет формирования на границе Si–SiO_2 слоя с увеличенной концентрацией азота. Полученные результаты могут быть использованы при создании диэлектрических покрытий в изделиях электронной техники.

Ключевые слова: двуокись кремния, импульсная фотонная обработка, пирогенное окисление, нитридизация, коэффициент преломления, электрофизические характеристики.

Конфликт интересов. Авторы заявляют об отсутствии конфликта интересов.

Благодарность. Авторы благодарят начальника сектора государственного центра «Белмикроанализ» ОАО «ИНТЕГРАЛ» Д. В. Жигулина за помощь в проведении экспериментальных исследований.

Для цитирования. Ковальчук, Н. С. Влияние импульсной фотонной обработки в среде азота на оптические и электрофизические характеристики слоев двуокиси кремния и ее границы с кремнием / Н. С. Ковальчук, В. А. Пилипенко, Я. А. Соловьёв // Доклады БГУИР. 2025. Т. 23, № 3. С. 5–11. http://dx.doi.org/10.35596/1729-7648-2025-23-3-5-11.

EFFECT OF PULSED PHOTON PROCESSING IN NITROGEN AMBIENT ON OPTICAL AND ELECTROPHYSICAL CHARACTERISTICS OF SILICON DIOXIDE LAYERS AND ITS BOUNDARIES WITH SILICON

NATALIA S. KOVALCHUK, VLADIMIR A. PILIPENKO, JAROSLAV A. SOLOVJOV

JSC "INTEGRAL" – Manager Holding Company "INTEGRAL" (Minsk, Republic of Belarus)

Abstract. The influence of pulsed photon processing in a nitrogen medium with an incoherent radiation flux from quartz halogen lamps directed to the non-working side of the wafer, providing heating to 1150 °C in about 7 s, on the optical properties of 17.7 nm thick silicon dioxide layers formed by pyrogenic oxidation of silicon doped with boron with the orientation (100), and on the electrophysical characteristics of the interface with silicon was established using the methods of infrared Fourier spectrometry, spectral ellipsometry, time-of-flight mass spectroscopy of secondary ions, and studies of the current-voltage and capacitance-voltage characteristics. It was revealed that pulsed photon processing of silicon dioxide layers leads to compaction and rearrangement of its structure, as well

as to the formation of Si–N bonds in silicon dioxide, providing nitridation of SiO_2. This is indicated by the shift, decrease in the half-width and voltage of the main absorption band of the Si–O bond, decrease in the refractive index from 1.48 to 1.47 and increase in the layer thickness to 18.2 nm. It is shown that nitridation of SiO_2 layers during pulsed photon processing in a nitrogen atmosphere leads to a decrease in the leakage current of the dielectric by four times and its charge density by 3.43 times. This is 2.19 and 3.01 times more, respectively, than during processing in natural atmospheric conditions due to the formation of a layer with an increased nitrogen concentration at the Si–SiO_2 boundary. The results obtained can be used to create dielectric coatings in electronic products.

Keywords: silicon dioxide, pulsed photon processing, pyrogenic oxidation, nitridation, refractive index, electrophysical characteristics.

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

Gratitude. The authors thank the head of the sector of the state center “Belmicroanalysis” of JSC “INTEGRAL” D. V. Zhigulin for assistance in conducting experimental studies.

For citation. Kovalchuk N. S., Pilipenko V. A., Solovjov Ja. A. (2025) Effect of Pulsed Photon Processing in Nitrogen Ambient on Optical and Electrophysical Characteristics of Silicon Dioxide Layers of Its Boundaries with Silicon. *Doklady BGUIR*. 23 (3), 5–11. http://dx.doi.org/10.35596/1729-7648-2025-23-3-5-11 (in Russian).

Введение

Процессы термического окисления кремния играют важную роль в технологии формирования изделий интегральной электроники. Вместе с тем существенное влияние на свойства слоев двуокиси кремния (SiO_2) и ее границы раздела с кремнием оказывают условия проведения процесса. Наиболее качественные слои двуокиси кремния, приближающиеся по свойствам к объемному материалу с минимальной плотностью зарядовых состояний на границе раздела с кремнием, получают путем окисления в сухом кислороде. Однако термическое окисление кремния в сухом кислороде необходимо проводить при температуре выше 1000 °C, что приводит к диффузионному перераспределению примесей в уже сформированных ионно-легированных областях и вызывает трудности при формировании подзатворного диэлектрика для субмикронных изделий микроэлектроники [1]. Для понижения температуры окисления широко применяют пирогенное окисление, представляющее собой процесс окисления кремния в кислороде, увлажненном молекулами H_2O, синтезированными из атомарно чистого кислорода и водорода. Данный подход обеспечил снижение температуры окисления кремния до 850 °C, тем самым уменьшил недостатки термического окисления кремния в сухом кислороде, обусловленные высокой температурой его проведения. Однако при таких низких температурах окисления это приводит к увеличению плотности зарядовых состояний, что обусловлено наличием оборванных связей кремния на границе раздела SiO_2–Si и присутствием на ней и в объеме двуокиси кремния гидроксильных групп и водорода.

Как показано в [2–5], для устранения недостатков процесса пирогенного окисления кремния целесообразно проводить модификацию сформированной двуокиси кремния с применением импульсной фотонной обработки в естественных атмосферных условиях при температуре до 1100–1150 °C. Это позволило значительно улучшить электрофизические свойства как слоя SiO_2, так и его границы раздела с кремнием. Представляет интерес проведение такой обработки в атмосфере азота, поскольку эта среда может привести к нитридизации диэлектрика и, как следствие, к более значительным изменениям его структуры и электрофизических параметров. В процессе исследований установлено влияние импульсной фотонной обработки в среде азота слоев двуокиси кремния, полученных пирогенным окислением подложек монокристаллического кремния, на ее оптические свойства и электрофизические характеристики границы раздела SiO_2 с кремнием.

Методика проведения эксперимента

Слои двуокиси кремния толщиной 17,7 нм формировали путем пирогенного окисления подложек монокристаллического кремния марки КДБ12 диаметром 100 мм с ориентацией (100) при температуре 850 °C в течение 40 мин. Затем часть подложек подвергли импульсной фотонной обработке в среде азота, обеспечивающей нагрев до 1150 °C за время около 7 с некогерентным потоком излучения от кварцевых галогенных ламп, направленным на нерабочую сторону пластины.

Анализ оптических свойств проводили путем сопоставления спектров пропускания слоев двуокиси кремния, полученных после пирогенного окисления и импульсной фотонной обработки, с применением ИК-Фурье спектрометра VERTEX 70 (Bruker, Германия) и определения их коэффициентов преломления и толщины методом спектральной эллипсометрии с использованием спектрального эллипсометра UVISEL 2 (Horiba, Франция). Электрофизические характеристики границы раздела слоев двуокиси кремния с кремниевой подложкой оценивались путем анализа их пробивных напряжений, токов утечки и вольт-фарадных характеристик (*C–V*-характеристик) системы Si–SiO_2 с применением комплекса B1500 (Agilent, США) с зондовой станцией Summit 11000 AP (Cascade, США). Для анализа распределения концентрации связи Si–N по толщине слоев SiO_2 и поверхностного слоя кремния использовался масс-спектрометр вторичных ионов TOF.SIMS 5 (IONTOF, Германия).

Результаты исследований и их обсуждение

Согласно рис. 1, где изображены спектры пропускания двуокиси кремния до и после импульсной фотонной обработки, имеют место сдвиг основной полосы поглощения связи Si–O в коротковолновую область, уменьшение ее полуширины и увеличение интенсивности. Такие изменения говорят как об улучшении структуры двуокиси кремния, а именно об уменьшении ее пористости и о снижении напряжения данной связи, так и об увеличении его толщины. Кроме того, появляется пик на 1161,93 см$^{-1}$, свойственный связи Si–N. Данный эффект получается потому, что фотонная обработка при температуре 1150 °C в среде азота приводит к его диффузии в двуокись кремния и к формированию связей с оборванными связями кремния. Образованию таких незаполненных связей способствовал способ получения слоев двуокиси кремния при низких температурах, что подтверждают результаты их исследований методом спектральной эллипсометрии.

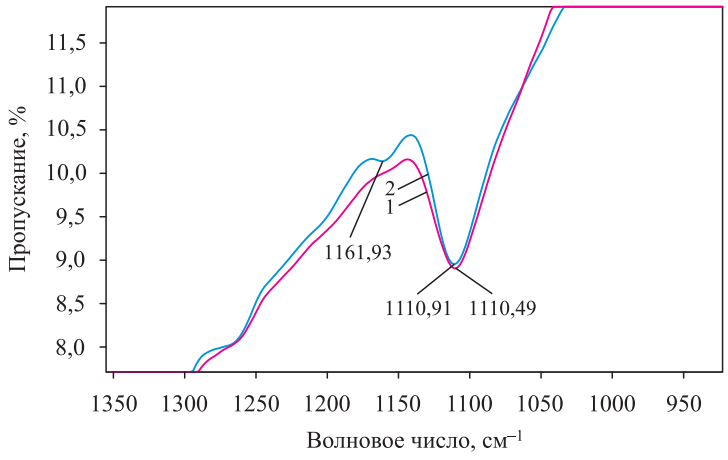


Рис. 1. Спектры пропускания слоев диоксида кремния:
1, 2 – до и после импульсной фотонной обработки
Fig. 1. Transmission spectra of silicon dioxide layers:
1, 2 – before and after pulsed photon processing

Установлено, что импульсная фотонная обработка слоя двуокиси кремния, полученного пирогенным окислением кремния, приводит к уменьшению коэффициента преломления с 1,48 до 1,47 и увеличению толщины пленки с 17,7 до 18,2 нм. Аналогичные результаты по увеличению толщины слоя (на 0,3 нм) двуокиси кремния при ее нитридизации получены и в [6]. Увеличение толщины слоя двуокиси кремния может объясняться тем, что при фотонной обработке в среде азота идут два процесса – уплотнение окисла за счет устранения группы OH из окисла, а также его рост при образовании связей Si–N, длина которых составляет ~0,1742 нм по сравнению с 0,1651 нм для Si–O. Последний процесс является преобладающим, что и вызывает рост толщины диэлектрика. Об этом также свидетельствует уменьшение коэффициента преломления пирогенной двуокиси кремния после импульсной фотонной обработки в среде азота с 1,48 до 1,47, что вызвано нарушением стехиометрии окисного слоя в результате образования дополнительных связей Si–N. Таким образом, проведение импульсной фотонной обработки двуокиси кремния, полученной пирогенным окислением, обеспечивающей нагрев до 1150 °C в среде азота, приво-

дит как к уплотнению, так и к перестройке ее структуры. На это указывают сдвиг, уменьшение полуширины и напряжения основной полосы поглощения Si–O-связи, уменьшение коэффициента преломления с 1,48 до 1,47, увеличение толщины слоя до 18,2 нм и образование в двуокиси кремния связей Si–N, обеспечивающих нитридизацию SiO_2.

Улучшение структуры слоя двуокиси кремния, полученного пирогенным окислением, после импульсной фотонной обработки в среде азота подтверждает анализ пробивных напряжений и токов утечки. Исследование данных характеристик показало, что проведение импульсной фотонной обработки пленок пирогенной двуокиси кремния приводит к увеличению ее напряжения пробоя с 17,3 до 18,0 В при уровне тока 1,0 мА и к уменьшению тока утечки с $1{,}0 \cdot 10^{-11}$ до $2{,}5 \cdot 10^{-12}$ А при напряжении 5,0 В. Поскольку толщина окисла увеличилась на 0,5 нм, такие изменения пробивного напряжения и токов утечки объяснить увеличением толщины диэлектрика невозможно. Полученные результаты можно связать лишь с образованием связей Si–N и уплотнением окисного слоя за счет снижения его пористости и перестройки структуры при импульсной фотонной обработке в среде азота.

Подтверждением влияния нитридизации слоя SiO_2 на напряжение пробоя и ток утечки является их сопоставление с параметрами, имеющими место при импульсной фотонной обработке пирогенной двуокиси кремния в естественных атмосферных условиях [3, 4]. В этом случае пробивное напряжение увеличивалось в 1,03 раза, а ток утечки уменьшился в 1,32 раза, в то время как при импульсной фотонной обработке в среде азота данные величины составили 1,04 и четыре раза соответственно. При этом следует отметить, что в [3, 4] толщина слоя SiO_2 составляла 42,5 нм и была в 2,4 раза больше толщины исследуемого диэлектрика. Приведенные данные показывают, что при обработке в азоте указанные выше параметры двуокиси кремния превышают аналогичные параметры, полученные при проведении процесса в естественных атмосферных условиях. То есть нитридизация SiO_2 оказала значительно большее воздействие на уменьшение тока утечки, что может быть связано только с образованием связей Si–N. Поскольку молекулы азота, не имеющие связей с атомами кремния в SiO_2, являются его дефектами, приводящими к увеличению тока утечки, отсутствие роста тока утечки позволяет утверждать, что азот в основном связан с кремнием, обеспечивая тем самым значительное уменьшение его величины. Такой результат может быть обусловлен как азотированием SiO_2 методом ионного легирования азотом с последующей импульсной фотонной обработкой, так и длительной термической обработкой в среде азота, компенсацией оборванных связей кремния за счет формирования с ними прочных ковалентных связей с азотом [7, 8].

Анализ вольт-фарадных характеристик (рис. 2) системы Si–SiO_2, сформированной на основе данных пленок, показал, что импульсная фотонная обработка в среде азота уменьшает плотность заряда с $1{,}82 \cdot 10^{11}$ до $5{,}30 \cdot 10^{10}$ см$^{-2}$, что полностью объясняется как перестройкой структуры пленки, которая происходит в ней во время такой обработки в азоте при нагреве до 1150 °С, так и устранением инверсного слоя в кремнии, сформированного зарядовыми состояниями в слое окисла.

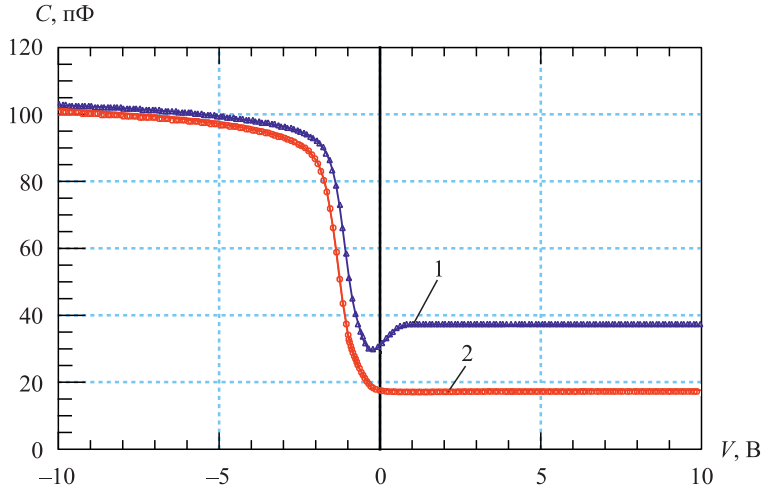


Рис. 2. Вольт-фарадные характеристики слоев SiO_2, полученных пирогенным окислением кремния: 1, 2 – до и после импульсной фотонной обработки

Fig. 2. Capacitance-voltage characteristics of SiO_2 layers obtained by pyrogenic oxidation of silicon: 1, 2 – before and after pulsed photon processing

Это подтверждается сопоставлением зарядовых свойств слоев SiO_2 после импульсной фотонной обработки в среде N_2 с результатами, полученными при обработке в естественных атмосферных условиях [3, 4]. Так, заряд в двуокиси кремния после обработки в азоте был в 2,19 раза меньше, чем его величина, приведенная в [3, 4]. Это можно объяснить тем, что при обработке в азоте образование связей Si–N может идти, как указывалось ранее, в местах оборванных связей кремния. Поскольку междоузельный кремний образуется как в объеме слоя двуокиси кремния за счет диссоциации группы OH на кислород и водород, так и в области межфазной границы Si–SiO_2 при оборванных связях Si на поверхности кремниевой пластины, нитридизация SiO_2 будет идти по всей ее толщине. При этом максимальная концентрация связей Si–N будет иметь место на межфазной границе Si–SiO_2 за счет наличия на ней значительно большей концентрации междоузельных атомов кремния, чем в остальном объеме двуокиси кремния.

Образование слоя с высокой концентрацией связей Si–N на межфазной границе экспериментально подтверждается исследованиями системы Si–SiO_2 времяпролетной масс-спектроскопией вторичных ионов (рис. 3). Кроме того, можно отметить увеличение концентрации связей Si–N на внешней поверхности слоя и их наличие в объеме диэлектрика. Увеличение концентрации таких связей на межфазной границе обеспечивает значительное уменьшение концентрации междоузельных атомов кремния.

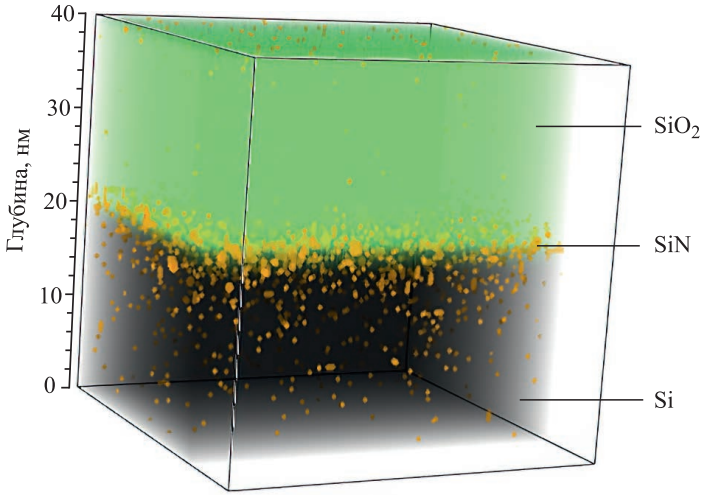


Рис. 3. 3D-распределение SiO_2, SiN и Si в системе Si–SiO_2 после импульсной фотонной обработки
Fig. 3. 3D distribution of the SiO_2, SiN, and Si in the Si–SiO_2 system after pulsed photon processing

Поскольку оборванные связи кремния являются основной причиной повышения фиксированного заряда в двуокиси кремния и заряда быстрых поверхностных состояний, это объясняет значительное уменьшение плотности заряда в слоях SiO_2 при их импульсной фотонной обработке в среде азота.

Таким образом, нитридизация слоев SiO_2 при импульсной фотонной обработке в атмосфере азота вызывает формирование на границе слоя Si–SiO_2 с увеличенной концентрацией азота. Сформированный слой обеспечивает уменьшение тока утечки диэлектрика в четыре раза и его заряда в 3,43 раза, что в 2,19 и 3,01 раза соответственно больше, чем при обработке в естественных атмосферных условиях.

Заключение

1. Проведение импульсной фотонной обработки двуокиси кремния, полученной пирогенным окислением, обеспечивающей нагрев до 1150 °C в среде азота, приводит как к уплотнению, так и к перестройке ее структуры. На это указывают сдвиг, уменьшение полуширины и напряжения основной полосы поглощения Si–O-связи, уменьшение коэффициента преломления с 1,48 до 1,47, увеличение толщины слоя с 17,7 до 18,2 нм и образование в двуокиси кремния связей Si–N, обеспечивающих нитридизацию SiO_2.

2. Нитридизация слоев SiO_2 при импульсной фотонной обработке в атмосфере азота приводит к уменьшениям тока утечки диэлектрика в четыре раза и его заряда в 3,43 раза, что в 2,19 и 3,01 раза соответственно больше, чем при обработке в естественных атмосферных условиях за счет формирования на границе слоя $Si–SiO_2$ с увеличенной концентрацией азота.

3. Полученные результаты могут быть использованы при создании диэлектрических покрытий в изделиях электронной техники.

Список литературы

1 Doering, R. Handbook of Semiconductor Manufacturing Technology, 2nd ed. / R. Doering, Y. Nishi. NY: CRC Press, 2007. https://doi.org/10.1201/9781420017663.
2. Влияние быстрой термической обработки исходных кремниевых пластин на процесс их пирогенного окисления / В. М. Анищик [и др.] // Журнал Белорусского государственного университета. Физика. 2018. № 2. С. 81–85.
3. Influence of the Rapid Thermal Treatment of the Initial Silicon Wafers on the Electro-Physical Properties of Silicon Dioxide, Obtained with Pyrogenous Oxidation / V. A. Pilipenko [et al.] // High Temperature Material Processes. 2019. Vol. 23, No 3. P. 283–290. DOI: 10.1615/HighTempMatProc. 2019031122.
4. Pilipenko, V. A. Influence of Fast Thermal Treatment on the Electrophysical Properties of Silicon Dioxide / V. A. Pilipenko, V. A. Solodukha, V. A. Gorushko // Journal of Engineering Physics and Thermophysics. 2018. Vol. 91, No 5. P. 1337–1341. https://doi.org/10.1007/s10891-018-1866-0.
5. Солодуха, В. А. Влияние быстрой термической обработки подзатворного диэлектрика на параметры мощных полевых MOSFET транзисторов / В. А. Солодуха, В. А. Пилипенко, В. А. Горушко // Доклады БГУИР. 2018. № 5. С. 99–103.
6. Красников, Г. Конструктивно-технологические особенности субмикронных МОП-транзисторов / Г. Красников. М.: Техносфера, 2002.
7. Влияние ионной имплантации азота на электрофизические свойства подзатворного диэлектрика силовых МОП-транзисторов / В. Б. Оджаев [и др.] // Журнал Белорусского государственного университета. Физика. 2020. № 3. С. 55–64. https://doi.org/10.33581/2520-2243-2020-3-55-64.
8. Локализация атомов азота в структурах $Si–SiO_2$ / В. Б. Оджаев [и др.] // Вестник Полоцкого государственного университета. Серия С. Фундаментальные науки. 2022. Т. 39, № 11. С. 65–79. https://doi.org/10.52928/2070-1624-2022-39-11-65-79.

Поступила 10.03.2025 Принята в печать 17.04.2025

References

1 Doering R., Nishi Y. (2008) *Handbook of Semiconductor Manufacturing Technology*. NY, CRC Press Publ. https://doi.org/10.1201/9781420017663.
2. Anishchik V. M., Harushka V. A., Pilipenka U. A., Ponariadov V. V., Saladukha V. A. (2018) Influence of the Rapid Thermal Treatment of the Initial Silicon Wafers on the Process of Their Pyrogenic Oxidation. *Journal of the Belarusian State University. Physics*. 2, 81–85 (in Russian).
3. Pilipenko V. A., Solodukha V. A., Zharin A., Gusev O. K., Vorobey R., Pantsialeyeu K., et al. (2019) Influence of the Rapid Thermal Treatment of the Initial Silicon Wafers on the Electro-Physical Properties of Silicon Dioxide, Obtained with Pyrogenous Oxidation. *High Temperature Material Processes*. 23 (3), 283–290. DOI: 10.1615/HighTempMatProc. 2019031122.
4. Pilipenko V. A., Solodukha V. A., Gorushko V. A. (2018) Influence of Fast Thermal Treatment on the Electrophysical Properties of Silicon Dioxide. *Journal of Engineering Physics and Thermophysics*. 91 (5), 1337–1341. https://doi.org/10.1007/s10891-018-1866-0.
5. Solodukha V. A., Pilipenko V. A., Gorushko V. A. (2018) Influence of Rapid Thermal Treatment of the Gate Dielectric on the Parameters of Power Field MOSFET Transistors. *Doklady BGUIR.* (5), 99–103 (in Russian).
6. Krasnikov G. (2002) *Design and Technological Features of Submicron MOSFETs.* Moscow, Technosphera Publ. (in Russian).
7. Odzaev V. B., Panfilenka A. K., Pyatlitski A. N., Prasalovich U. S., Kovalchuk N. S., Soloviev Ya. A., et al. (2020) Influence of Nitrogen Ion Implantation on the Electrophysical Properties of the Gate Dielectric of Power MOSFETs. *Journal of the Belarusian State University. Physics*. (3), 55–64. https://doi.org/10.33581/2520-2243-2020-3-55-64 (in Russian).

8. Odzaev V., Prasalovich U., Pyatlitski A., Kovalchuk N., Soloviev Ya., Zhigulin D., et al. (2022) Localization of Nitrogen Atoms in Si–SiO_2 Structures. *Herald of Polotsk State Univercity. Series C. Fundamental Sciences*. 39 (11), 65–79. https://doi.org/10.52928/2070-1624-2022-39-11-65-79 (in Russian).

Received: 10 March 2025 Accepted: 17 April 2025

Вклад авторов / Authors' contribution

Авторы внесли равный вклад в написание статьи / The authors contributed equally to the writing of the article.

Сведения об авторах

Ковальчук Н. С., канд. техн. наук, доц., зам. ген. дир. – гл. инж., ОАО «ИНТЕГРАЛ» – управляющая компания холдинга «ИНТЕГРАЛ» (ОАО «ИНТЕГРАЛ»)

Пилипенко В. А., д-р техн. наук, проф., зам. нач. государственного центра «Белмикроанализ», ОАО «ИНТЕГРАЛ»

Соловьёв Я. А., д-р. техн. наук, доц., зав. отраслевой лабораторией новых технологий и материалов, ОАО «ИНТЕГРАЛ»

Information about the authors

Kovalchuk N. S., Cand. Sci. (Tech.), Associate Professor, Deputy of General Director – Chief Engineer, JSC "INTEGRAL" – Manager Holding Company "INTEGRAL" (JSC "INTEGRAL")

Pilipenko V. A., Dr. Sci (Tech.), Professor, Deputy Head of the State Center "Belmicroanalysis", JSC "INTEGRAL"

Solovjov Ja. A., Dr. Sci. (Tech.), Associate Professor, Head of the Laboratory of New Technologies and Materials, JSC "INTEGRAL"

Адрес для корреспонденции

220108, Республика Беларусь,
Минск, ул. Казинца, 121а
ОАО «ИНТЕГРАЛ» –
управляющая компания холдинга «ИНТЕГРАЛ»
Тел.: +375 17 398-14-03
E-mail: jsolovjov@integral.by
Соловьёв Ярослав Александрович

Address for correspondence

220108, Republic of Belarus,
Minsk, Kazintsa St., 121a
JSC "INTEGRAL" –
Manager Holding Company "INTEGRAL"
Tel.: +375 17 398-14-03
E-mail: jsolovjov@integral.by
Solovjov Jaroslav Aleksandrovich

CC BY

http://dx.doi.org/10.35596/1729-7648-2025-23-3-12-18

УДК 621.357.74-022.532:669.056.91

ЭЛЕКТРОФИЗИЧЕСКИЕ СВОЙСТВА СТРУКТУР С ПЛЕНКАМИ ТИТАНАТА БАРИЯ И ОКСИДА АЛЮМИНИЯ, СФОРМИРОВАННЫМИ ЗОЛЬ-ГЕЛЬ-МЕТОДОМ НА ТИТАНЕ

Р. Т. МАХМУТОВ[1], Н. В. ГАПОНЕНКО[1], Е. И. ЛАШКОВСКАЯ[1], К. И. МЕЛЕДИН[1], И. А. КАШКО[1], А. Д. МАЛЫШЕВ[1], Л. В. СУДНИК[2]

[1]Белорусский государственный университет информатики и радиоэлектроники (Минск, Республика Беларусь)
[2]Институт порошковой металлургии имени академика О. В. Романа (Минск, Республика Беларусь)

Аннотация. Золь-гель-методом сформированы пленки оксида алюминия и титаната бария на подложках из титана и окисленного монокристаллического кремния. Исследована зависимость емкости и тангенса угла диэлектрических потерь от частоты конденсаторной структуры $Al_2O_3/BaTiO_3/Ni$ в диапазоне 200 Гц–200 кГц на подложке из титана с использованием подложки в качестве нижнего электрода. В области высоких частот (20–200 кГц) среднеквадратичное отклонение емкости составило 95–97 пФ, низких частот (0,2–10,0 кГц) – 80–94 пФ. Значения емкости и тангенса угла диэлектрических потерь сохранялись в пределах установленных значений среднеквадратичного отклонения в течение года. Приведены примеры изображений микродисков и волноводов, полученных методами фотолитографии и химического травления структур с пленками Al_2O_3, SiO_2 и $BaTiO_3$ для последующих разработок планарных волноводов и электрооптических модуляторов.

Ключевые слова: золь-гель-метод, конденсатор, титанат бария, титан, микродиски, волноводы, фотолитография.

Конфликт интересов. Авторы заявляют об отсутствии конфликта интересов.

Благодарность. Авторы выражают благодарность начальнику сектора государственного центра «Белмикроанализ» ОАО «ИНТЕГРАЛ» Д. В. Жигулина за анализ образцов методом сканирующей электронной микроскопии.

Для цитирования. Электрофизические свойства структур с пленками титаната бария и оксида алюминия, сформированными золь-гель-методом на титане / Р. Т. Махмутов [и др.] // Доклады БГУИР. 2025. Т. 23, № 3. С. 12–18. http://dx.doi.org/10.35596/1729-7648-2025-23-3-12-18.

ELECTROPHYSICAL PROPERTIES OF STRUCTURES WITH BARIUM TITANATE AND OXIDE FILMS FORMED BY THE SOL-GEL METHOD ON TITANIUM

RUSLAN T. MAKHMUTOV[1], NIKOLAI V. GAPONENKO[1], EKATERINA I. LASHKOVSKAYA[1], KIRILL I. MELEDIN[1], IVAN A. KASHKO[1], ALEXANDER D. MALYSHEV[1], LARISA V. SUDNIK[2]

[1]Belarusian State University of Informatics and Radioelectronics (Minsk, Republic of Belarus)
[2]Powder Metallurgy Institute (Minsk, Republic of Belarus)

Abstract. Aluminum oxide and barium titanate films were formed on titanium and oxidized single-crystal silicon substrates using the sol-gel method. The dependence of the capacitance and the dielectric loss tangent on the frequency of the $Al_2O_3/BaTiO_3/Ni$ capacitor structure was studied in the range of 200 Hz–200 kHz on a titanium substrate using the substrate as the lower electrode. In the high-frequency region (20–200 kHz), the standard deviation of the capacitance was 95–97 pF, and in the low-frequency region (0.2–10.0 kHz), it was 80–94 pF. The capacitance and dielectric loss tangent values remained within the established standard deviation values for a year. Examples of images of microdisks and waveguides obtained by photolithography and chemical etching of struc-

tures with Al_2O_3, SiO_2 and $BaTiO_3$ films are given for subsequent development of planar waveguides and electro-optical modulators.

Keywords: sol-gel method, capacitor, barium titanate, titanium, microdisks, waveguides, photolithography.

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

Gratitude. The authors express their gratitude to the head of the sector of the state center "Belmicroanalysis" of JSC "INTEGRAL" D. V. Zhigulin for the analysis of samples using scanning electron microscopy

For citation. Makhmutov R. T., Gaponenko N. V., Lashkovskaya E. I., Meledin K. I., Kashko I. A., Malyshev A. D., Sudnik L. V. (2025) Electrophysical Properties of Structures with Barium Titanate and Oxide Films Formed by the Sol-Gel Method on Titanium. *Doklady BGUIR*. 23 (3), 12–18. http://dx.doi.org/10.35596/1729-7648-2025-23-3-12-18 (in Russian).

Введение

В настоящее время развиваются технологии формирования пленок с высокой диэлектрической проницаемостью не только на кремнии и других полупроводниках, но и на металлах. Покрытия на металлах представляют интерес для разработки металлических электродов широкого назначения – например, в конструкции развязывающих конденсаторов [1], в устройствах для разложения воды импульсным электрическим полем [2] и др. Исследования по формированию диэлектрических пленок на металлах проводятся несмотря на высокий уровень шероховатости поверхности по сравнению с подложками материалов микроэлектроники.

С другой стороны, интерес к пленкам титаната бария возрастает за счет электрооптического эффекта, что обеспечивает возможность электрического переключения и модуляции оптического сигнала [3]. В этой связи актуальны разработки фотолитографии и способов химического травления планарных структур и канальных волноводов из пленок титаната бария и других, совместимых с ним, материалов буферных слоев на различных подложках. Известно, что канальные волноводные структуры для разработки компонентов оптоэлектроники могут быть получены золь-гель-методом, фотолитографией и химическим травлением. Однако в большинстве публикаций изложены данные, полученные на сплошных пленках. Высокое качество легированных эрбием золь-гель-пленок может обеспечить низкие волноводные потери и достижение коэффициента усиления в оптическом диапазоне около 1 дБ/см, что дает возможность разрабатывать волноводные разветвители на их основе с компенсацией потерь [4].

В статье приведены результаты исследования емкости и тангенса угла диэлектрических потерь в зависимости от частоты для конденсаторной структуры на подложке титана ($Ti/Al_2O_3/BaTiO_3/Ni$). Проведен анализ изображения планарных микроструктур на основе пленок оксида алюминия и титаната бария, полученных золь-гель-методом.

Проведение эксперимента

Для формирования пленок оксида алюминия использовался золь на основе изопропоксида алюминия, изопропилового спирта, воды и азотной кислоты. Пленки оксида алюминия наносили на подложки кремния, а также подложки с оксидом кремния (Si/SiO_2) и подложки из титана. Для формирования конденсатора на титане четыре слоя ксерогеля оксида алюминия формировали центрифугированием золя со скоростью 2000 об/мин с последующей сушкой каждого слоя при 200 °C и отжигом на воздухе при 450 °C в течение 30 мин. Общая толщина всех четырех слоев составляла около 110 нм. Затем наносили золь титаната бария. Его синтезировали на основе изопропоксида титана $Ti(OC_3H_7)_4$, ацетата бария ($Ba(CH_3COO)_2$), ацетилацетона ($CH_3–CO–CH_2–CO–CH_3$) и уксусной кислоты (CH_3COOH) [5].

Для измерения электрофизических параметров пленочных структур на титане было приготовлено два образца. Первый формировался для измерения емкости и тангенса угла диэлектрических потерь. Для этого пять слоев ксерогеля $BaTiO_3$ формировались центрифугированием на структуре Ti/Al_2O_3 со скоростью 2900 об/мин с последующей сушкой каждого слоя при 200 °C в течение 10 мин и отжигом на воздухе при 450 °C в течение 30 мин после сушки последнего слоя. Окончательная термообработка проводилась при 700 °C. Затем методом ионно-лучевого распыления выполнялось напыление контактов из никеля квадратной формы площадью около 1,4 мм^2. Второй образец на титане формировался аналогично первому и использовался для измерения пробивных напряжений. Пробивные напряжения измерялись после формирования пяти и десяти слоев ксерогеля титаната бария.

Волновод и микродиски $BaTiO_3$ формировали на подложке Si/SiO_2 размерами 20×20 мм. Образец обрабатывали в хромовой смеси и наносили методом центрифугирования пять слоев золь-гель-покрытия титаната бария с промежуточной сушкой при 200 °С и термическим отжигом при 450 °С. Рисунок волновода и микродисков формировали методом контактной фотолитографии с использованием позитивного фоторезиста DSAM-3020. Слои $BaTiO_3$ и SiO_2 травили в буферном травителе, содержащем HF и NH_4F.

Емкость *C* и тангенс угла диэлектрических потерь tgδ измеряли с использованием измерителя иммитанса Е7-20 в диапазоне частот 200 Гц–200 кГц. Измерения проводились в течение одного года после формирования образца.

Для исследования пробивных напряжений использовался источник постоянного напряжения Б5-84/1 (МНИПИ, РБ). Напряжение подавалось между медным контактом, прижатым к пленке, и титановой подложкой в качестве второго контакта. Напряжение пробоя фиксировалось при появлении тока через пленку. Морфологию полученных образцов исследовали с помощью сканирующего электронного микроскопа (СЭМ) S-4800 (Hitachi, Япония). СЭМ-анализ использовался для контроля толщины пленок ксерогелей, сформированных на титане и кремнии в аналогичных условиях.

Результаты исследований и их обсуждение

Семислойные структуры, сформированные на подложках кремния из золя на основе изопропоксида алюминия с концентрацией 7 мг/мл, не растрескиваются при температуре отжига 300 °С и имеют толщину приблизительно 280–240 нм, уменьшающуюся при увеличении скорости центрифугирования от 2000 до 3000 об/мин (рис. 1).

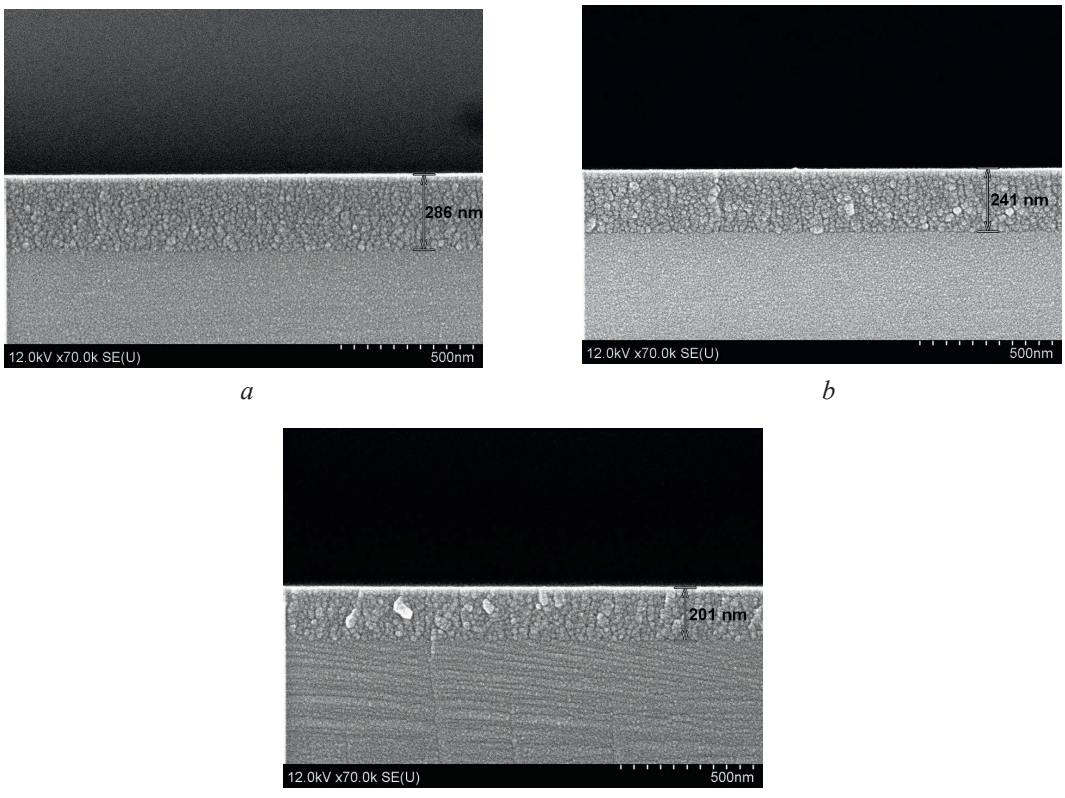


Рис. 1. СЭМ-изображения семислойных пленок Al_2O_3,
полученных при разных скоростях центрифугирования, об/мин: *a* – 2000; *b* – 2500; *c* – 3000
Fig. 1. SEM images of seven-layer Al2O3 films obtained at different spin speeds, rpm:
a – 2000; *b* – 2500; *c* – 3000

Толщина однослойных пленок титаната бария, сформированных центрифугированием, составляла около 80 нм [5]. Исходя из данных СЭМ-анализа на кремнии и титане, на рис. 2 приведена схема конденсаторной структуры на титане с указанием толщины многослойных пленок: Al_2O_3 – 110 нм, $BaTiO_3$ – 330 нм.

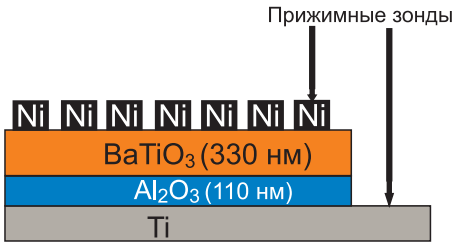
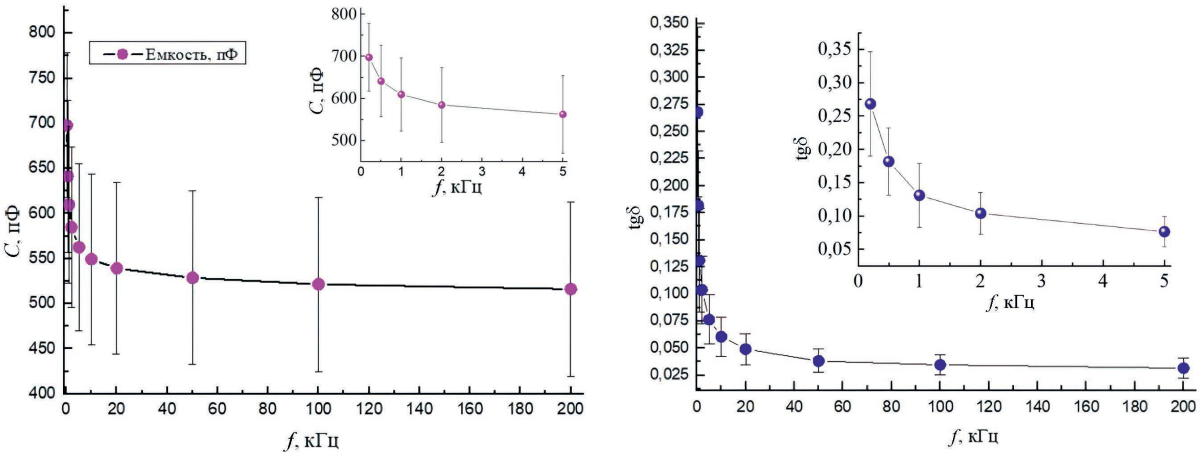


Рис. 2. Схематическое изображение конденсаторной структуры на титане
Fig. 2. Schematic representation of a capacitor structure on titanium

На рис. 3 изображены частотные (*f*) зависимости с доверительным интервалом емкости *C* и тангенса угла диэлектрических потерь tgδ конденсатора на подложке из титана с пленками ксерогелей оксида алюминия и титаната бария. Численные средние значения *C* и tgδ для 17 измерений (объем выборки), а также значения среднеквадратичных отклонений σ этих величин приведены в табл. 1.



a *b*

Рис. 3. Частотные зависимости с доверительным интервалом емкости (*a*)
и тангенса угла диэлектрических потерь конденсатора (*b*)
на подложке из титана с пленками ксерогелей оксида алюминия и титаната бария
Fig. 3. Frequency dependences with confidence interval of capacitance (*a*)
and tangent of dielectric loss angle of capacitor (*b*)
on titanium substrate with films of aluminum oxide and barium titanate xerogels

Таблица 1. Характеристики конденсатора на титановой подложке для различных частот измерения
Table 1. Characteristics of a capacitor on a titanium substrate for measuring different frequencies

Частота *f*, кГц	Среднее значение *C*, пФ	Среднеквадратичное отклонение σ_C	Среднее значение tgδ	Среднеквадратичное отклонение $\sigma_{tg\delta}$
0,2	698	80	0,27	0,08
0,5	641	85	0,18	0,05
1	610	87	0,13	0,05
2	585	89	0,10	0,03
5	563	92	0,08	0,02
10	549	94	0,06	0,02
20	539	95	0,05	0,01
50	529	96	0,04	0,01
100	522	97	0,03	0,01
200	516	97	0,03	0,01

С увеличением частоты f средние значения емкости C и tgδ уменьшаются соответственно от 698 до 516 пФ и от 0,27 до 0,03. При этом в области высоких частот 20–200 кГц среднеквадратичное отклонение емкости σ_C составляет 95–97, в области низких частот 0,2–10,0 кГц – несколько меньше – 80–94. Среднеквадратичное отклонение тангенса угла диэлектрических потерь $\sigma_{tg\delta}$, наоборот, уменьшается в области высоких частот.

Пробивное напряжение пленочной структуры $Ti/Al_2O_3/BaTiO_3$ составляло 14 В для пятислойной пленки $BaTiO_3$, а после наращивания добавочной пятислойной пленки $BaTiO_3$ увеличилось до 38 В. Полагая, что толщина добавочной пятислойной пленки $BaTiO_3$ в структуре $Ti/Al_2O_3/BaTiO_3$ такая же, как и на кремнии, и по данным СЭМ составляет 340 нм, оценивали, что электрическая прочность (напряженность электрического поля, при которой достигается пробой) ксерогеля титаната бария должна составлять ~70 В/мкм.

Многослойные пленки ксерогелей в сочетании с фотолитографией и мокрым химическим травлением пригодны для формирования фотолитографических рисунков. Примеры таких изображений приведены на рис. 4, 5. Для формирования канального волновода из ксерогеля Al_2O_3 методом центрифугирования на кремниевую подложку наносили 15 слоев золя Al_2O_3, термообработку каждого слоя проводили при 300 °С. СЭМ-изображения канального волновода Al_2O_3 на кремнии приведены на рис. 4. Структуры канального волновода и микродиска, изображенные на рис. 5, были сформированы на термически окисленных кремниевых подложках нанесением пяти слоев золя $BaTiO_3$ методом центрифугирования с последующей сушкой при температуре 200 °С и термообработкой при 450 °С.

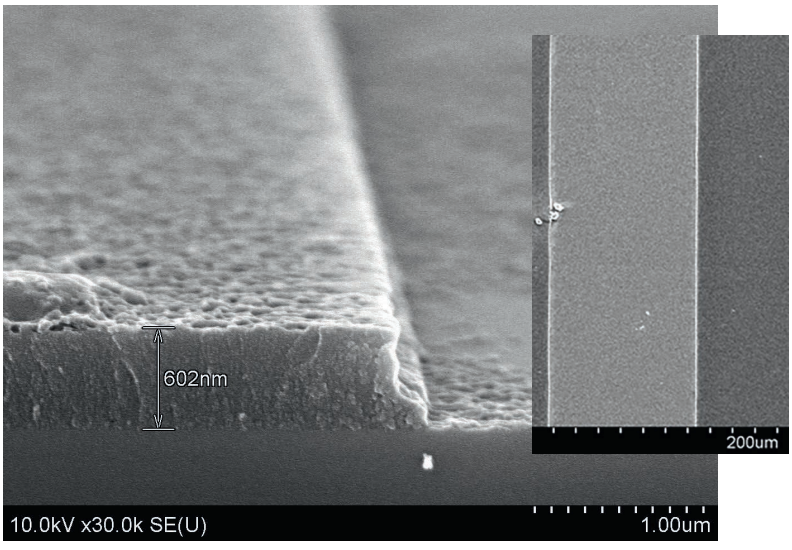


Рис. 4. СЭМ-изображение дорожек на сколе образца с 15 слоями Al_2O_3 (температура отжига каждого слоя 300 °С): на вставке – вид сверху
Fig. 4. SEM image of tracks on a cleavage of a sample with 15 layers of Al_2O_3 (annealing temperature of each layer is 300 °C): inset – top view

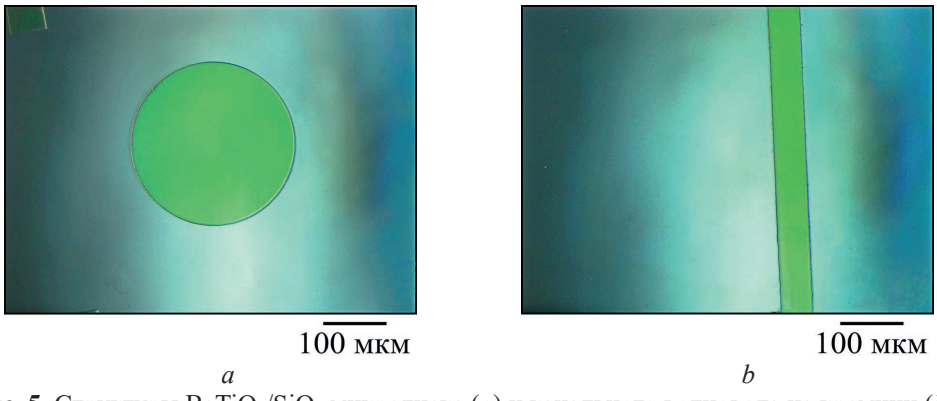


Рис. 5. Структуры $BaTiO_3/SiO_2$ микродиска (*a*) и канального волновода на кремнии (*b*)
Fig. 5. $BaTiO_3/SiO_2$ structures of microdisc (*a*) and channel waveguide on silicon (*b*)

Для корректной оценки волноводных потерь и их снижения в канальных волноводах на различных подложках потребуется формирование качественного слоя с низким показателем преломления между подложкой и волноводом, что является предметом дальнейших исследований. Развитие технологии микродисков и волноводов перспективно для элементной базы оптоэлектроники. Микродиски, легированные люминесцирующими ионами и синтезированные золь-гель-методом, могут найти применение в микролазерах с генерацией моды шепчущей галереи, в том числе в сочетании с канальными волноводами на одной подложке [6].

Заключение

Золь-гель-методом синтезированы многослойные пленочные структуры из пленок ксерогелей титаната бария на подложках из титана и кремния. Для структур на подложках из титана с дополнительным слоем из оксида алюминия среднее значение емкости, измеренное в течение года хранения образца, составило 700 пФ для частоты 200 Гц и 500 пФ – для 200 кГц. Электрическая прочность ксерогеля титаната бария составила приблизительно 70 В/мкм. Показана возможность изготовления планарных структур из микродисков и канальных волноводов методом фотолитографии и мокрого химического травления, что в дальнейшем будет использовано для изучения волноводных потерь и перераспределения оптических мод в структуре волновод/микрорезонатор.

Работа выполнена при финансовой поддержке задания 2.02 ГПНИ «Материаловедение, новые материалы и технологии», задания 1.4 ГПНИ «Химические процессы, реагенты и технологии, биорегуляторы и биооргхимия» и проекта Белорусского республиканского фонда фундаментальных исследований № Т23РНФ-147.

Список литературы

1. Hironori, Hatono. Application of $BaTiO_3$ Film Deposited by Aerosol Deposition to Decoupling Capacitor / Hironori Hatono, Tomokazu Ito, Akihiko Matsumura // Jpn. J. Appl. Phys. 2007. Vol. 46, No 10B. P. 6915–6919. https://doi.org/10.1143/JJAP.46.6915.
2. Устройство разложения воды на кислород и водород электромагнитными полями: пат. Рос. Федерации № 2 645 504 С2 / Г. Л. Багич. Опубл. 21.02.2018.
3. Ferroelectric $BaTiO_3$ Thin-Film Optical Waveguide Modulators / A. Petraru [et al.] // Appl. Phys. Lett. 2002. Vol. 81. P. 1375–1377. https://doi.org/10.1063/1.1498151.
4. Huang, W. Sol-Gel Silica-on-Silicon Buried-Channel EDWAs / W. Huang, R. R. A. Syms // J. Light. Technol. 2003. Vol. 21, No 5. P. 1339–1349. https://doi.org/10.1109/JLT.2003.812417.
5. Optical Properties and Upconversion Luminescence of $BaTiO_3$ Xerogel Structures Doped with Erbium and Ytterbium / E. I. Lashkovskaya [et al.] // Gels. 2022. Vol. 8. P. 347–361. https://doi.org/10.3390/gels8060347.
6. Vahala, K. J. Optical Microcavities / K. J. Vahala // Nature. 2003. Vol. 424. P. 839–846. https://doi.org/10.1038/nature01939.

Поступила 04.03.2025 Принята в печать 08.05.2025

References

1. Hironori Hatono, Tomokazu Ito, Akihiko Matsumura (2007) Application of $BaTiO_3$ Film Deposited by Aerosol Deposition to Decoupling Capacitor. *Jpn. J. Appl. Phys*. 46 (10B), 6915–6919. https://doi.org/10.1143/JJAP.46.6915.
2. Bagich G. L. (2018) Device for Decomposing Water into Oxygen and Hydrogen Using Electromagnetic Fields. *Russian Federation Patent No 2 645 504 C2*. Published 21.02.2018 (in Russian).
3. Petraru A., Schubert J., Schmid M., Buchal Ch. (2002) Ferroelectric $BaTiO_3$ Thin-Film Optical Waveguide Modulators. *Appl. Phys. Lett*. 81, 1375–1377. https://doi.org/10.1063/1.1498151.
4. Huang W., Syms R. R. A. (2003) Sol-Gel Silica-on-Silicon Buried-Channel EDWAs. *J. Light. Technol*. 21 (5), 1339–1349. https://doi.org/10.1109/JLT.2003.812417.
5. Lashkovskaya E. I., Gaponenko N. V., Stepikhova M. V., Yablonskiy A. N., Andreev B. A., Zhivulko V. D., et al. (2002) Optical Properties and Upconversion Luminescence of $BaTiO_3$ Xerogel Structures Doped with Erbium and Ytterbium. *Gels*. 8, 347–361. https://doi.org/10.3390/gels8060347.
6. Vahala K. J. (2003) Optical Microcavities. *Nature*. 424, 839–846. https://doi.org/10.1038/nature01939.

Received: 4 March 2025 Accepted: 8 May 2025

Вклад авторов / Authors' contribution

Авторы внесли равный вклад в написание статьи / The authors contributed equally to the writing of the article.

Сведения об авторах

Махмутов Р. Т., асп. каф. микро- и наноэлектроники, Белорусский государственный университет информатики и радиоэлектроники (БГУИР)

Гапоненко Н. В., д-р физ.-мат. наук, проф., зав. науч.-исслед. лаб. «Нанофотоника» (Лаб. 4.5), БГУИР

Лашковская Е. И., мл. науч. сотр. Лаб. 4.5, БГУИР

Меледин К. И., асп. каф. микро- и наноэлектроники, БГУИР

Кашко И. А., науч. сотр. науч.-исслед. лаб. «Интегрированные микро- и наносистемы», БГУИР

Малышев А. Д., студ., БГУИР

Судник Л. В., д-р техн. наук, зам. дир. по научной работе Научно-исследовательского института импульсных процессов с опытным производством, Институт порошковой металлургии имени академика О. В. Романа

Information about the authors

Makhmutov R. T., Postgraduate at the Department of Micro- and Nanoelectronics, Belarusian State University of Informatics and Radioelectronics (BSUIR)

Gaponenko N. V., Dr. Sci. (Phys. and Math.), Professor, Head of the R&D Laboratory "Nanophotonics" (Lab 4.5), BSUIR

Lashkovskaya E. I., Junior Researcher at the Lab 4.5, BSUIR

Meledin K. I., Postgraduate at the Department of Micro- and Nanoelectronics, BSUIR

Kashko I. A., Researcher at the R&D Laboratory "Integrated Micro- and Nanosystems", BSUIR

Malyshev A. D., Student, BSUIR

Sudnik L. V., Dr. Sci. (Tech.), Deputy Director for Research at the Research Institute of Pulse Processes with Pilot Production, Powder Metallurgy Institute

Адрес для корреспонденции

220013, Республика Беларусь,
Минск, ул. П. Бровки, 6
Белорусский государственный университет
информатики и радиоэлектроники
Тел.: +375 17 293-88-75
E-mail: gaponenko@bsuir.by
Гапоненко Николай Васильевич

Address for correspondence

220013, Republic of Belarus,
Minsk, P. Brovki St., 6
Belarusian State University
of Informatics and Radioelectronics
Tel.: +375 17 293-88-75
E-mail: gaponenko@bsuir.by
Gaponenko Nikolai Vasilievich

CC BY

http://dx.doi.org/10.35596/1729-7648-2025-23-3-19-25

УДК 535.376; 535-15; 628.9.038; 628.9.037

МОДЕЛИРОВАНИЕ ОПТИЧЕСКИХ ПРОЦЕССОВ В ТОНКОПЛЕНОЧНОМ ИК-СВЕТОДИОДЕ НА ОСНОВЕ КОЛЛОИДНЫХ КВАНТОВЫХ ТОЧЕК PbS

У. Е. ТУРОВЕЦ

Белорусский государственный университет информатики и радиоэлектроники (Минск, Республика Беларусь)

Аннотация. Проведено моделирование оптических процессов в структуре тонкопленочного ИК-светодиода методом конечной разности во временной области. Исследованы такие параметры, как пропускание, эффективность распространения электромагнитных волн в диапазоне 1,2–1,4 мкм. Смоделировано контрольное устройство со слоем оксида индия-олова (ITO) в качестве прозрачного проводящего электрода. Рассмотрена замена слоя ITO на слой оксида олова, легированного фтором (FTO). Установлено, что при такой замене коэффициент пропускания при прохождении ИК-излучением функциональных слоев FTO увеличивается до 70 %, а угловое распределение E^2 – на 10° в сравнении с устройством со слоем ITO. В связи с чем целесообразно заменять слой прозрачного проводящего электрода ITO на слой FTO.

Ключевые слова: моделирование, метод конечной разности во временной области, ИК-светодиод, напряженность поля, пропускание, эффективность извлечения света.

Конфликт интересов. Автор заявляет об отсутствии конфликта интересов.

Для цитирования. Туровец, У. Е. Моделирование оптических процессов в тонкопленочном ИК-светодиоде на основе коллоидных квантовых точек PbS / Доклады БГУИР. 2025. Т. 23, № 3. С. 19–25. http://dx.doi.org/10.35596/1729-7648-2025-23-3-19-25.

MODELING OF OPTICAL PROCESSES IN THIN-FILM IR LIGHT-EMITTING DIODE BASED ON COLLOIDAL PbS QUANTUM DOTS

ULYANA Y. TURAVETS

Belarusian State University of Informatics and Radioelectronics (Minsk, Republic of Belarus)

Abstract. The optical processes in the structure of a thin-film IR LED have been simulated using the finite difference method in the time domain. The parameters such as transmission and propagation efficiency of electromagnetic waves in the range of 1.2–1.4 μm have been investigated. A control device with an indium tin oxide (ITO) layer as a transparent conducting electrode has been simulated. The replacement of the ITO layer with a fluorine-doped tin oxide (FTO) layer has been considered. It has been found that with such a replacement, the transmittance of IR radiation passing through the FTO functional layers increases to 70 %, and the angular distribution of E^2 increases by 10° compared to a device with an ITO layer. Thus, it is advisable to replace the layer of the transparent conducting ITO electrode with the FTO layer.

Keywords: modeling, finite difference time domain method, IR LED, field strength, transmittance, light extraction efficiency.

Conflict of interests. The author declares no conflict of interests.

For citation. Turavets U. Y. (2025) Modeling of Optical Processes in Thin-Film IR Light-Emitting Diode Based on Colloidal PbS Quantum Dots. *Doklady BGUIR*. 23 (3), 19–25. http://dx.doi.org/10.35596/1729-7648-2025-23-3-19-25 (in Russian).

Введение

Внешняя квантовая эффективность (ВКЭ) – один из важнейших параметров ИК-светодиодов (ИКС) на основе коллоидных квантовых точек (ККТ) PbS. Ограничение значений ВКЭ может быть связано с низким коэффициентом эффективности излучения света (ЭИС). Для получения качественного и количественного анализа данного параметра моделируют оптические процессы, происходящие в структуре светодиода, что позволяет рассчитывать рассеиваемую мощность и оценивать степень извлечения света из устройства [1, 2]. Эти параметры служат для первичной оценки эффективности работы функциональных слоев устройства. Для моделирования различных светоизлучающих элементов, как правило, применяется 2D-моделирование, поскольку оно выполняется значительно быстрее и требует меньших вычислительных мощностей [3].

В статье на основании результатов моделирования процессов распространения света рассмотрена оптимизация ВКЭ ИКС на основе квантовых точек сульфида свинца путем подбора материала проводящего прозрачного электрода (ППЭ).

Теоретическое обоснование

Для оценки распределения электромагнитных волн в структуре ИКС выполняли расчет распределения интенсивности дальнего поля, равной напряженности эклектического поля в квадрате (E^2). Данные параметры служат для первичной оценки эффективности ИКС, их расчет выполняли путем решения уравнений Максвелла с использованием метода конечной разности во временной области (FDTD) на базе программного комплекса Ansys Lumerical.

Конструкция ИКС представляет собой тонкопленочные функциональные слои [4, 5]. На рис. 1 приведена 3D-модель структуры ИКС в соответствии с порядком слоев, а толщина материалов слоев – в табл. 1. Значения комплексного показателя преломления для ближнего ИК-диапазона были взяты из открытых баз данных [6, 7].

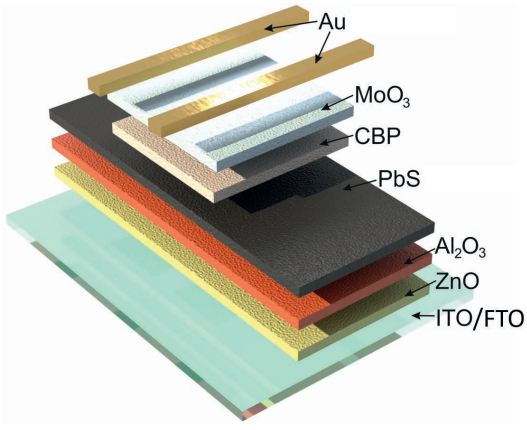


Рис. 1. Схематическое представление конструкции инфракрасного светодиода на коллоидных квантовых точках PbS
Fig. 1. Schematic representation of the infrared light-emitting diode design based on colloidal PbS quantum dots

Таблица 1. Параметры функциональных слоев инфракрасного светодиода
Table 1. Infrared light-emitting diode functional layers' parameters

Функциональный слой	Материал	Толщина, нм
Катод	Au	100
Дырочно-инжекционный слой (ДИС)	MoO_3	3
Дырочно-транспортный слой (ДТС)	4,4'–N,N'–дикарбазолил–бифенил (CBP)	60
Активный слой (ККТ)	PbS	47,5
Электронно-блокирующий слой (ЭБС)	Al_2O_3	3
Электронно-транспортный слой (ЭТС)	ZnO	25
Анод (ППЭ)	Оксид индия-олова ITO/оксид олова, легированный фтором FTO	70/70
Подложка	SiO_2	>1000

Для сравнительной оценки характеристик использовали такие параметры, как ЭИС, оптическое пропускание и распределение электромагнитных волн в структуре ИКС для предполагаемых функциональных слоев и их толщин. Параметр ЭИС определяет отношение фотонов, испускаемых светодиодом, ко всем фотонам, испускаемым активной областью, и определяется из формулы [1]

$$\text{ВКЭ} = \text{ЭИС} \cdot \text{ВнКЭ}, \tag{1}$$

где $\text{ЭИС} = 1 - \sqrt{1 - \left(\frac{n_{\text{ЭТС}}^2}{n_{\text{акт.сл}}^2}\right)}$ [2]; ВнКЭ – внутренняя квантовая эффективность.

ЭИС зависит от коэффициента преломления активного слоя и ЭТС. В рассматриваемом случае активный слой – ККТ PbS и ZnO.

С учетом ЭБС коэффициент ЭИС = 0,34 %, без учета ЭБС – ЭИС = 0,64 %. То есть ЭБС препятствует распространению света на выходе из устройства. Однако, по полученным экспериментальным данным, использование ЭБС необходимо для выравнивания неравномерной подвижности носителей заряда в активную область [4], что значительно улучшает выходные электрические характеристики. Поэтому стратегия выбора слоев для ИКС требует комплексного решения и компромисса при оптимизации слоев. Так, слой оксида индия-олова (ITO) непрозрачен в ИК-области, поэтому блокирует прохождение излучения.

Проведем сравнение двух моделей: контрольной структуры, изображенной на рис. 1, и модели с заменой ППЭ, где будет выполнена замена слоя ITO на слой оксида олова, легированного фтором (FTO). В качестве анода (или ППЭ) предлагается использовать FTO – как более подходящий материал, сохраняющий прозрачность в ИК-области на уровне 70 % (в ITO – 35 %) [8, 9]. Кроме того, материал FTO показывает незначительное ухудшение проводимости, что имеет большое значение при выборе материала ППЭ. Параметры моделирования для каждого случая были аналогичны.

Область симуляции составляла 4×1 мкм^2 в плоскости *x–y* (длина–ширина) и 0,5 мкм – в направлении координаты *z* (толщина). В качестве граничных условий области симуляции по координатам *x–y* выбраны идеально сочетающиеся слои (perfectly matched layers, PML). Количество слоев PML – 24. По координате *z* для верхнего слоя катода были установлены граничные условия металла (для отражения излучения обратно в область симуляции), а для нижнего слоя анода – стандартные PML. Границы PML поглощают почти все входящее излучение с минимальными отражениями, и, поскольку в реальной жизни свет выходит за пределы этих границ, было важно, чтобы излучение не отражалось обратно в область симуляции.

Границы симуляции по координате *z* устанавливались по середине слоев ITO и Au. Для моделирования люминесцирования в активном слое использовался точечный источник излучения (dipole source) [2], имеющий пик излучения на длине волны 1300 нм, что соответствует длине волны излучения ККТ PbS. Чтобы смоделировать угловое распределение волн, а также оценить изменение коэффициента пропускания, был задан диапазон для длин волн ИК-излучения 1,2–1,4 мкм. Чтобы создать эффект произвольного направления излучения (так как процессы рекомбинации с последующим направлением излучения носят случайный характер), точечный источник имел случайную ориентацию в пространстве активного слоя. Для упрощения вычисления и оптимизации использовался доступный официальный скрипт (sweep width and position) [3], суммирующий результаты для каждого положения точечного источника в пространстве активного слоя и направления его излучения. Шаг вычисления составил 0,25 нм, время симуляции 1200 фс. Для увеличения точности вычислений ультратонких слоев ЭБС и ДИС устанавливался индивидуальный шаг сетки 0,1 нм. Экран детектирования падающего излучения (пропускания) располагался по координате *z*.

Результаты исследований и их обсуждение

Анализ углового распределения E^2 контрольной модели (рис. 2, *a*) показывает, что распространение излучения происходит преимущественно при 60° за пределами границы устройства в дальнем поле. При замене слоя ППЭ на FTO (рис. 2, *b*) видно заметное усиление интенсивности

и расширения области выходящего за пределы устройства излучения. В этом случае распространение излучения происходит в пределах 70°. Однако стоит учитывать, что при задании диапазона длин волн излучения спектр имеет выраженный «хвост» на значении 1,4 мкм, соответственно, угловое распределение на данной длине волны имеет незатухающий характер.

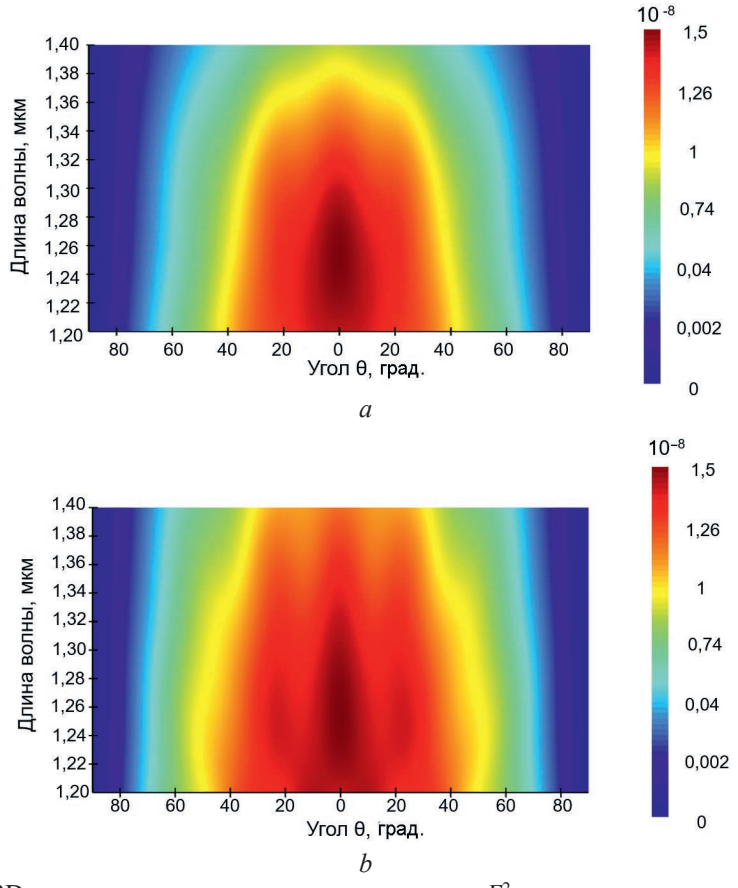


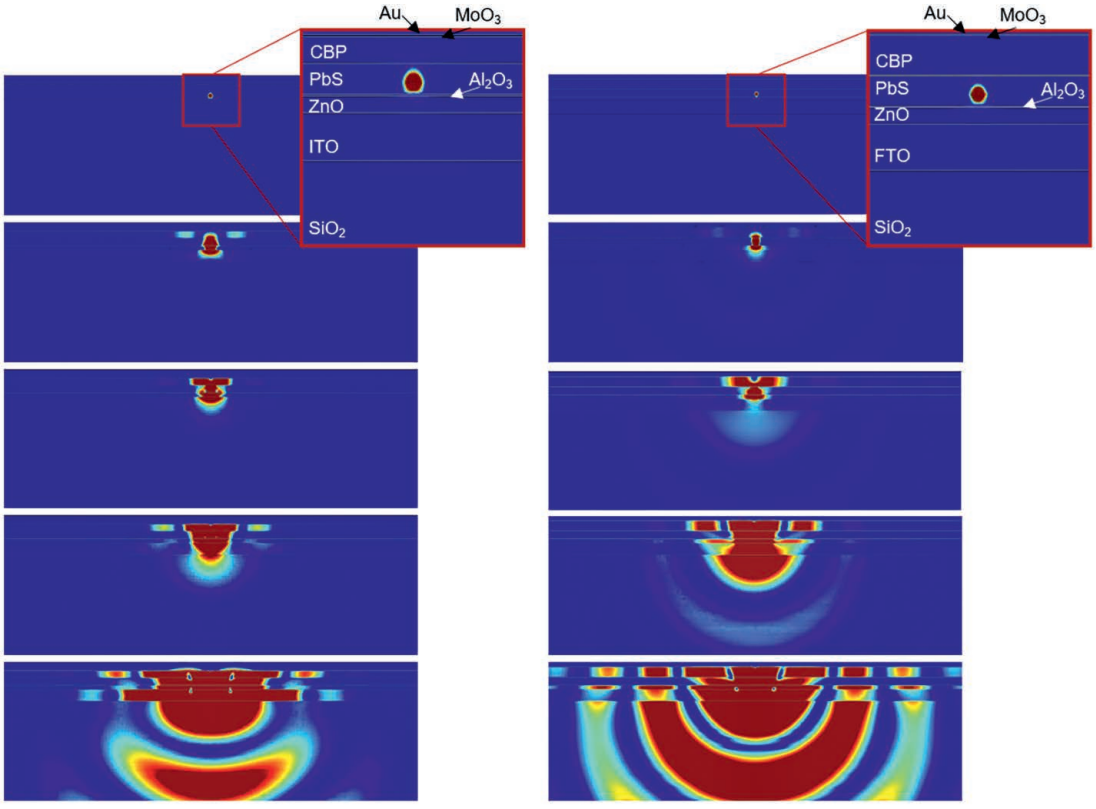
Рис. 2. 2D-представление углового распределения E^2 от длины волны излучения с использованием заданного коэффициента преломления за пределами границы устройства в дальнем поле для структуры со слоем: *a* – ITO; *b* – FTO

Fig. 2. 2D representation of the angular distribution of E^2 versus the radiation wavelength using a given refractive index beyond the device boundary in the far field for a structure with a layer of: *a* – ITO; *b* – FTO

По характеру распределения поля *E* (рис. 3) можно заметить, что ультратонкий слой оксида алюминия (ЭБС) выступает как барьер, в некоторой степени препятствующий прохождению излучения. Происходит частичная потеря излучения в активном слое, а также в ДТС и ЭТС. В конечном итоге, в результате отражения от золотого электрода путем множественного переотражения волны распространяются преимущественно в направлении ППЭ. В случае замены ППЭ на слой FTO (рис. 3, *b*) видно, что интенсивность излучения, проходящего в направлении ППЭ, выше по сравнению с контрольным устройством.

Согласно рис. 4, коэффициент пропускания значительно уменьшается после прохождения функциональных слоев на выходе из контрольного устройства. Это может быть связано с тем, что в ближней ИК-области происходит значительное поглощение излучения слоем ITO.

ITO по-прежнему занимает лидирующие позиции в качестве ППЭ ввиду низкого поверхностного сопротивления и высокой прозрачности в видимом диапазоне. Однако в ИК-области происходит заметное ухудшение пропускания [10]. В рассматриваемом случае значение пропускания на длине волны 1300 нм составляет порядка 35 %. Противоположная ситуация наблюдается при замене слоя ITO на FTO. Отмечено заметное улучшение пропускания в ИК-области с увеличением значений до 70 % на длине волны 1300 нм. Таким образом, полностью подтверждено предположение о целесообразности использования FTO в качестве ППЭ.



a *b*

Рис. 3. 2D-срез распределения поля *E* во временной области с 15 до 60 фс для структуры со слоем: *a* – ITO; *b* – FTO

Fig. 3. 2D slice of the E field distribution in the time domain from 15 to 60 fs for a structure with a layer: *a* – ITO; *b* – FTO

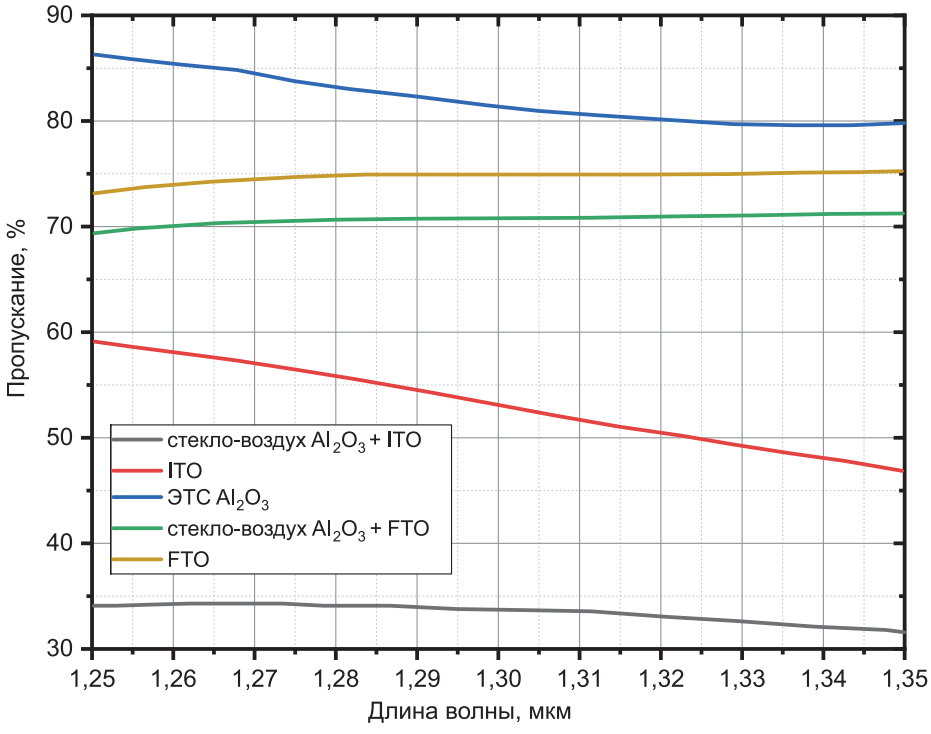


Рис. 4. Оптическое пропускание излучения в различных слоях сравниваемых устройств

Fig 4. Optical transmission of radiation in different layers of the compared devices

Заключение

1. В результате моделирования и первичной оценки параметров эффективности излучения света выявлены недостатки тонкопленочного ИК-светодиода на основе квантовых точек сульфида свинца, которые в значительной мере были устранены путем оптимизации состава функциональных слоев. Осуществлено моделирование двух сходных по конструкции ИК-светодиодов, отличающихся лишь материалом прозрачного проводящего электрода: контрольного устройства со слоем оксида индия-олова (ITO) в качестве прозрачного проводящего электрода и устройства с заменой ITO на слой оксида олова, легированного фтором (FTO).

2. В результате анализа полученных данных можно сделать вывод о целесообразности замены материала прозрачного электрода ITO на слой FTO, поскольку при этом повышается оптическое пропускание и улучшается распределение электромагнитных волн в структуре. Оптическое пропускание достигло 70 % на границе стекло–воздух устройства с заменой ППЭ на слой FTO, что на 35 % больше по сравнению с контрольным устройством со слоем ITO. Угловое распределение E^2 в устройстве с заменой материала прозрачного электрода на слой FTO показало значение 70° на длине волны 1300 нм, что на 10° больше, чем у контрольного устройства на той же длине волны.

3. Исследование проведено при поддержке Белорусского фонда фундаментальных исследований по теме гранта № Т23М-040 «Неорганические инфракрасные светодиоды на коллоидных квантовых точках PbS с улучшенными характеристиками».

Список литературы

1. Optical Tunneling to Improve Light Extraction in Quantum Dot and Perovskite Light-Emitting Diodes / G. Mei [et. al.] // IEEE Photonics Journal. 2020. Vol. 12, No 6. P. 1–14. http://dx.doi.org/10.1109/JPHOT.2020.3038275.
2. Photonics Design Theory Enhancing Light Extraction Efficiency in Quantum Dot Light Emitting Diodes / Diyar D. Othman [et al.] // Journal of Physics: Materials. 2022. Vol. 5, No 4. http://dx.doi.org/10.1088/2515-7639/ac9e77.
3. OLED (2D) [Electronic Resource] // Ansys Optics. Mode of access: https://optics.ansys.com/hc/en-us/articles/360042225934-OLED-2D. Date of access: 17.09.2024.
4. Bright Infra-Red Quantum Dot Light-Emitting Diodes Through Efficient Suppressing of Electrons / M. Marus [et al.] // Appl. Phys. Let. 2020. Vol. 116, No 19. http://dx.doi.org/10.1063/5.0005843.
5. Туровец, У. Е. Оптическое моделирование тонкопленочного ИК-светодиода на основе коллоидных квантовых точек / У. Е. Туровец // Компьютерное проектирование в электронике: сб. тр. Междунар. науч.-практ. конф., Минск, 28 нояб. 2024 г. Минск, 2024. С. 63–66.
6. Refractive Index [Electronic Resource]. Mode of access: https://refractiveindex.info. Date of access: 26.09.2024.
7. Filmmetrics [Electronic Resource]. Mode of access: https://www.filmetrics.com. Date of access: 29.09.2024.
8. Physical Investigations on Undoped and Fluorine Doped SnO_2 Nanofilms on Flexible Substrate Along with Wettability and Photocatalytic Activity Tests / S. Ben Ameur [et al.] // Materials Science in Semiconductor Processing. 2017. Vol. 61. P. 17–26. http://dx.doi.org/ 10.1016/j.mssp.2016.12.019.
9. Effect of Transparent Conductive Layers on the Functionality of Liquid Crystal Devices: Comparison of AZO, FTO and ITO / V. Marinova [et al.] // Optical Materials: X. 2024. Vol. 22. http://dx.doi.org/10.1016/j.omx.2024.100330.
10. Pasquarelli, R. M. Solution Processing of Transparent Conductors: From Flask to Film / R. M. Pasquarelli, D. S. Ginley, R. O'Hayre // Chemical Society Reviews. 2011. Vol. 40, No 11. P. 5406–5441. http://dx.doi.org/10.1039/c1cs15065k.

Поступила 03.03.2025 Принята в печать 09.04.2025

References

1. Mei G., Wu D., Ding S., Choy W. C. H., Wang K., Sun X. W. (2020) Optical Tunneling to Improve Light Extraction in Quantum Dot and Perovskite Light-Emitting Diodes. *IEEE Photonics Journal*. 12 (6), 1–14. http://dx.doi.org/10.1109/JPHOT.2020.3038275.
2. Othman D., Weinstein J. A, Lyu Q., Hou B. (2022) Photonics Design Theory Enhancing Light Extraction Efficiency in Quantum Dot Light Emitting Diodes. *Journal of Physics: Materials*. 5 (4). http://dx.doi.org/10.1088/2515-7639/ac9e77.

3. OLED (2D). *Ansys Optics*. Available: https://optics.ansys.com/hc/en-us/articles/360042225934-OLED-2D (Accessed 17 September 2024).
4. Marus M., Xia Y., Zhong H., Li D., Ding S., Turavets U., et al. (2020) Bright Infra-Red Quantum Dot Light-Emitting Diodes Through Efficient Suppressing of Electrons. *Appl. Phys. Let*. 116 (19). http://dx.doi.org/10.1063/5.0005843.
5. Turavets U. Y. (2024) Optical Modeling of Thin Film IR Quantum Dot Light Emitting Diode. *Electronic Design Automation: Conference materials, Minsk, 28 Nov.* 63–66 (in Russian).
6. *Refractive Index*. Available: https://refractiveindex.info (Accessed 26 September 2024).
7. *Filmmetrics*. Available: https://www.filmetrics.com (Accessed 29 September 2024).
8. Ben Ameur S., Barhoumi A., Bel Hadjltaief H., Mimouni R., Duponchel B., Leroye G., et al. (2017) Physical Investigations on Undoped and Fluorine Doped SnO_2 Nanofilms on Flexible Substrate Along with Wettability and Photocatalytic Activity Tests. *Materials Science in Semiconductor Processing*. 61, 17–26. http://dx.doi.org/10.1016/j.mssp.2016.12.019.
9. Marinova V., Petrov S., Petrova D., Napoleonov B., Chau N. H. M., Lan Y. P., et al. (2022) Effect of Transparent Conductive Layers on the Functionality of Liquid Crystal Devices: Comparison of AZO, FTO and ITO. *Optical Materials: X*. 22. http://dx.doi.org/10.1016/j.omx.2024.100330.
10. Pasquarelli R. M., Ginley D. S., O'Hayre R. (2011) Solution Processing of Transparent Conductors: From Flask to Film. *Chemical Society Reviews*. 40 (11), 5406–5441. http://dx.doi.org/10.1039/c1cs15065k.

Received: 3 March 2025 Accepted: 9 April 2025

Сведения об авторе

Туровец У. Е., мл. науч. сотр. науч.-исслед. лаб. интегрированных микро- и наносистем, Белорусский государственный университет информатики и радиоэлектроники

Адрес для корреспонденции

220013, Республика Беларусь,
Минск, ул. П. Бровки, 6
Белорусский государственный университет
информатики и радиоэлектроники
Тел.: +375 17 293-88-69
E-mail: u.turovets@bsuir.by
Туровец Ульяна Егоровна

Information about the author

Turavets U. Ya., Junior Researcher at the R&D Laboratory of Integrated Micro- and Nanosystems, Belarusian State University of Informatics and Radioelectronics

Address for correspondence

220013, Republic of Belarus,
Minsk, P. Brovki St., 6
Belarusian State University
of Informatics and Radioelectronics
Tel.: +375 17 293-88-69
E-mail: u.turovets@bsuir.by
Turavets Ulyana Yahorayna

CC BY

http://dx.doi.org/10.35596/1729-7648-2025-23-3-26-35

УДК 621.3.049.7

СИСТЕМНЫЕ МАТЕМАТИЧЕСКИЕ МАКРОМОДЕЛИ ЛОГИЧЕСКИХ МЭМС-ВЕНТИЛЕЙ

А. А. СОЛОВЬЕВ[1], Е. Ф. ПЕВЦОВ[1], В. А. КОЛЧУЖИН[2]

[1]МИРЭА – Российский технологический университет (Москва, Российская Федерация)
[2]Qorvo Munich GmbH (Мюнхен, Германия)

Аннотация. Актуальность исследования обусловлена стремительным развитием микроэлектромеханических систем (МЭМС), которые находят применение в качестве логических элементов благодаря низкому энергопотреблению, высокой надежности и возможности перепрограммирования в процессе работы. Однако существующие подходы к проектированию логических МЭМС-вентилей требуют разработки компактных и точных математических моделей для анализа их динамических характеристик и логических операций на системном уровне. В статье представлен метод синтеза параметрических математических моделей логических МЭМС-вентилей на основе кремниевых резонаторов с электростатическим возбуждением. Разработана библиотека конструкционных макромоделей, включающая трех-, пяти- и девятиконтактные МЭМС-резонаторы, которые позволяют проводить анализ статических, частотных и динамических характеристик. Проведены анализ эффекта схлопывания электродов, гармонический анализ и симуляции логических операций, таких как НЕ, И, ИЛИ-НЕ и исключающее ИЛИ. Разработанные модели могут быть использованы для проектирования аналого-цифровых и цифро-аналоговых преобразователей, триггеров и арифметико-логических устройств на основе МЭМС-технологий.

Ключевые слова: микроэлектромеханические системы, МЭМС-резонаторы, логические элементы, вентили, системное моделирование, язык описания аппаратуры.

Конфликт интересов. Авторы заявляют об отсутствии конфликта интересов.

Для цитирования. Соловьев, А. А. Системные математические макромодели логических МЭМС-вентилей / А. А. Соловьев, Е. Ф. Певцов, В. А. Колчужин // Доклады БГУИР. 2025. Т. 23, № 3. С. 26–35. http://dx.doi.org/10.35596/1729-7648-2025-23-3-26-35.

SYSTEM-LEVEL MATHEMATICAL MACRO-MODELS OF LOGIC MEMS GATES

ALEKSANDR A. SOLOVEV[1], EVGENY F. PEVTSOV[1], VLADIMIR A. KOLCHUZHIN[2]

[1]MIREA – Russian Technological University (Moscow, Russian Federation)
[2]Qorvo Munich GmbH (Munich, Germany)

Abstract. The relevance of the study is due to the rapid development of microelectromechanical systems (MEMS), which are used as logic elements due to low power consumption, high reliability and the ability to reprogram during operation. However, existing approaches to the design of MEMS logic gates require the development of compact and accurate mathematical models to analyze their dynamic characteristics and logical operations at the system level. The article presents a method for synthesizing parametric mathematical models of MEMS logic gates based on silicon resonators with electrostatic excitation. A library of design macromodels has been developed, including three-, five- and nine-pin MEMS resonators, which allow analyzing static, frequency and dynamic characteristics. The analysis of the electrode collapse effect, harmonic analysis and simulation of logical operations such as NOT, AND, OR-NOT and exclusive OR are carried out. The developed models can be used for designing analog-to-digital and digital-to-analog converters, triggers and arithmetic logic devices based on MEMS technologies.

Keywords: microelectromechanical systems, MEMS resonators, logic elements, gates, system modeling, hardware description language.

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

For citation. Solovev A. A., Pevtsov E. F., Kolchuzhin V. A. (2025) System-Level Mathematical Macro-Models of Logic MEMS Gates. *Doklady BGUIR*. 23 (3), 26–35. http://dx.doi.org/10.35596/1729-7648-2025-23-3-26-35 (in Russian).

Введение

Микроэлектромеханические системы (МЭМС) представляют собой одно из наиболее перспективных направлений современной микроэлектроники, которое находит применение в различных областях – от датчиков [1] и исполнительных устройств [2] до сложных логических схем [3]. В последние годы наблюдается значительный рост интереса к использованию МЭМС-резонаторов в качестве логических элементов, что обусловлено их низким энергопотреблением, высокой надежностью и возможностью перепрограммирования в процессе работы [4]. Однако, несмотря на очевидные преимущества, разработка эффективных логических МЭМС-вентилей требует решения ряда задач, связанных с созданием точных и компактных математических моделей, позволяющих анализировать их динамические характеристики и логические операции на системном уровне.

Предметом данного исследования являются системные математические модели логических МЭМС-вентилей на основе кремниевых резонаторов с электростатическим возбуждением. Актуальность исследования обусловлена необходимостью разработки новых подходов к проектированию логических элементов, которые могли бы конкурировать с традиционными КМОП-технологиями по энергоэффективности и функциональной гибкости. Гипотеза исследования заключается в том, что использование электростатического возбуждения и резонансных явлений в МЭМС-резонаторах позволяет реализовать логические операции с высокой точностью и низким энергопотреблением.

Степень изученности проблемы показывает, что, несмотря на значительное количество работ, посвященных МЭМС-резонаторам [5, 6], вопросы системного моделирования логических элементов на их основе остаются недостаточно исследованными. В частности, отсутствуют универсальные библиотеки макромоделей, которые могли бы быть использованы для проектирования сложных цифровых схем. В [7] рассмотрены вопросы синтеза и анализа пятиконтактного логического МЭМС-элемента. В данной статье представлен более комплексный подход, предложен метод синтеза библиотеки параметрических математических моделей логических МЭМС-вентилей на основе трех-, пяти- и девятиконтактных МЭМС-резонаторов.

Цель исследования – создание системных моделей, позволяющих анализировать статические, частотные и динамические характеристики МЭМС-резонаторов, а также реализовать на их основе логические операции. Для этого использован язык описания аппаратуры VHDL-AMS, что обеспечивает точное описание электромеханического поведения резонаторов. Проведены анализ эффекта схлопывания электродов, гармонический анализ и симуляция логических операций, таких как НЕ, И, ИЛИ-НЕ и исключающее ИЛИ.

Результаты исследования демонстрируют возможность реализации логических операций на основе МЭМС-резонаторов с электростатическим возбуждением, что открывает перспективы для создания энергоэффективных и перепрограммируемых цифровых схем. Разработанные модели могут быть использованы для проектирования аналого-цифровых и цифро-аналоговых преобразователей, триггеров и арифметико-логических устройств на основе МЭМС-технологий.

Трехконтактный МЭМС-резонатор

Трехконтактный МЭМС-резонатор представляет собой систему «пружина–масса–демпфер», состоящую из недеформируемой сейсмической массы, подвешенной на четырех подвесах (рис. 1). Собственная частота выбрана 439,49 кГц, сейсмическая масса – 0,019 мкг, зазор между сейсмической массой и боковыми электродами – 4 мкм.

Для создания системной модели трехконтактного кремниевого микрорезонатора [8] с электростатическим возбуждением был выбран язык описания аппаратных средств VHDL-AMS. Аналоговое поведение компонентов реализовано через величины, терминалы и уравнения. Меха-

ническое поведение резонатора описывается с помощью следующего дифференциального уравнения 2-го порядка:

$$m\ddot{u}(t) + d\dot{u}(t) + ku(t) = F_{exp}(t),$$

где d – коэффициент демпфирования; t – время; u – смещение сейсмической массы от положения равновесия; $F_{exp}(t)$ – внешняя сила, приводящая резонатор в движение.

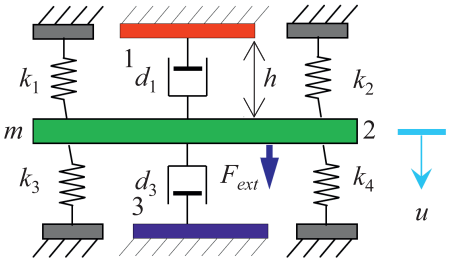


Рис. 1. Эквивалентная механическая схема трехконтактного резонатора
Fig. 1. Equivalent mechanical circuit of a three-terminal resonator

Согласно расчетам, коэффициент демпфирования $d = 2{,}57 \cdot 10^{-6}$ мкН·с/мкм для $Q = 20$. Резонатор возбуждается в нужном направлении электростатическими силами F_1, F_2, F_3, F_4 согласно следующему уравнению:

$$F_{elec}(t) = \frac{1}{2}\sum_{r=1,2,3,4}\frac{\partial C_{r5}}{\partial u}(V_r - V_5)^2,$$

где C_{r5} – емкость между подвижным 5 и боковым r электродами; V_r – электрический потенциал на r-электроде.

Согласно уравнению, электростатическая сила, приложенная к контактам, пропорциональна квадрату разницы потенциалов, а направление силы определяется знаком производной $\partial C_{r5}/\partial u$.

Для трехконтактной системы электрическая эквивалентная схема состоит из двух конденсаторов (рис. 2, *a*) без учета межэлектродных паразитных емкостей. Токи, протекающие в такой системе, можно представить в виде графа (рис. 2, *b*).

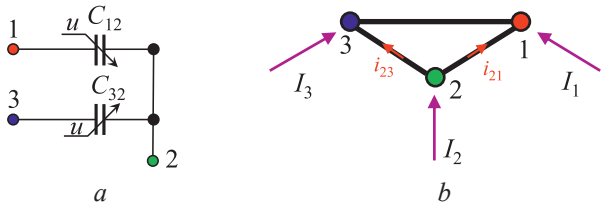


Рис. 2. Эквивалентная электрическая схема МЭМС-резонатора (*a*) и граф узловых токов (*b*)
Fig. 2. Equivalent electrical circuit of a MEMS resonator (*a*) and a graph of nodal currents (*b*)

Для нахождения емкостей между подвижным и боковыми электродами используется приближение плоскопараллельного конденсатора согласно выражениям:

$$C_{12}(u) = \xi\xi_0\frac{A}{h+u};\ C_{32}(u) = \xi\xi_0\frac{A}{h+u},$$

где h – зазор между электродами; u – смещение, ограниченное зазором, $|u| < h$.

Внешние токи, втекающие в узлы, исходя из графа узловых токов без учета межэлектродных паразитных емкостей (рис. 2, *b*) имеют следующий вид:

$$I_2 = -i_{21} - i_{23}.$$

За положительное направление принят втекающий в узел ток. Токи i_{12} и i_{32}, протекающие через конденсаторы C_{12} и C_{32}, пропорциональны количеству зарядов Q_{12} и Q_{32}, протекающих в единицу времени:

$$i_{r2} = \frac{\partial}{\partial t}Q_{r2} = \frac{\partial}{\partial t}\left(C_{r2}\left(V_r - V_2\right)\right).$$

Как видно из уравнения, изменение тока может быть вызвано изменением приложенных электрических потенциалов или перераспределением заряда из-за движения электрода.

В статье представлена концепция МЭМС-резонатора, функционирующего в режиме электромеханической обратной связи, где возбуждение механических колебаний и получение выходного сигнала осуществляется электрически. Механические колебания резонатора возбуждались электрическим способом с помощью синусоидального сигнала, подаваемого на управляющий электрод. Выходной сигнал получался путем детектирования тока на чувствительном электроде. Такой подход отличает устройство от традиционных датчиков и исполнительных устройств, где механическое возбуждение детектируется электрически, а электрическое возбуждение приводит к механическим колебаниям. Данная схема требует более сложной обработки сигналов и управления, что открывает новые перспективы для повышения точности и эффективности работы устройства. Это позволяет повысить чувствительность и реализовать новые функциональные возможности по сравнению с традиционными МЭМС-устройствами.

Функционирование разработанной модели проверено путем симуляции статического, частотного и динамического (во временной области) откликов структуры. Анализ эффекта схлопывания электродов состоял в вычислении смещения подвижного электрода под воздействием приложенного напряжения. Напряжение схлопывания – важный параметр, который ограничивает максимально возможные прикладываемые напряжения к структуре. Напряжение схлопывания между электродами 1–2 равняется 275,87 В согласно формуле

$$V_{PI} = \sqrt{\frac{8k_{eff}\,gap^3}{27\xi A}},$$

где $gap = h$ – начальный зазор между подвижным и неподвижными электродами; ξ – диэлектрическая проницаемость среды; A – площадь перекрытия между электродами (площадь срабатывания).

На рис. 3 представлены два случая:

1) электрод 1 – плавающий, электрод 3 имеет нулевой потенциал, линейно меняющееся напряжение подается на подвижный электрод 3;

2) электрод 1 имеет нулевой потенциал, электрод 3 – плавающий, линейно меняющееся напряжение подается на подвижный электрод.

Как видно из рис. 3, напряжение схлопывания составляет 276,1 В.

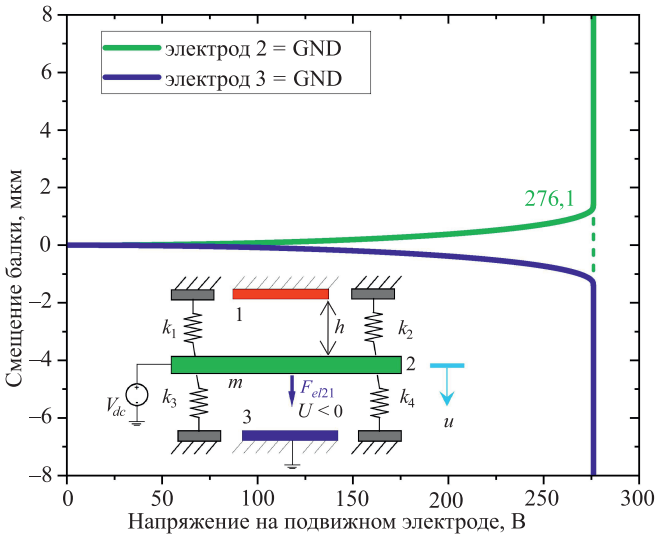


Рис. 3. Анализ эффекта схлопывания при смещении боковых электродов
Fig. 3. Analysis of the pull-in effect with lateral electrode displacement

Гармонический анализ заключается в симуляции реакции структуры на внешнее гармоническое возбуждение. Работа логического элемента основана на изменении резонансной частоты подвижного электрода с использованием эффекта электростатического смягчения жесткости резонатора. Для определения оптимальных величин напряжения смещения, подаваемого на подвижный

электрод, и амплитуды импульсных сигналов, подаваемых на боковые электроды, была проведена серия симуляций. Зависимости тока на чувствительном электроде 1 от частоты напряжения смещения на подвижном электроде 2 приведены на рис. 4. При увеличении напряжения на подвижном электроде электростатическая сила возрастает, что вызывает смещение резонансного пика влево (из-за электростатического смягчения жесткости резонатора) и увеличение амплитуды осцилляций.

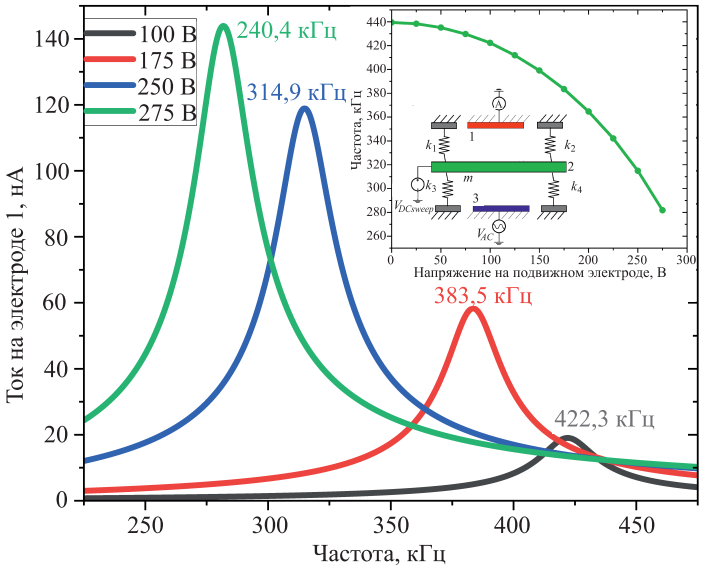


Рис. 4. Зависимость амплитудно-частотных характеристик от напряжения на подвижном электроде 2, $Q = 20$, $V_3 = 1$ В (AC)

Fig. 4. Amplitude-frequency response dependence on voltage applied to movable electrode 2, $Q = 20$, $V_3 = 1$ V (AC)

На каждый период импульсного сигнала на цифровом входе нужно как минимум 10 периодов синусоидального сигнала на приводе (электроде 4), чтобы элемент успел переключиться в другое состояние. Поэтому частота переключения логического элемента будет в 10 раз меньше, чем частота резонанса.

Реализация логических операций построена на принципе резонанса [6], когда резонансная частота подвижного электрода и частота подаваемого сигнала совпадают, и высокий выходной ток емкостным способом детектируется на чувствительном электроде 1. Для реализации логических операций на электрод 3, который используется как цифровой вход элемента, подается импульсный сигнал. Необходимо подключить внешние источники напряжения, структурные нагрузки (узловые силы) и задать параметры симуляции (размер шага по времени, время моделирования). На рис. 5 приведена схема подключения для выполнения логической операции НЕ.

Для реализации логических операций на подвижный электрод 2 подавалось смещение 250 В, на управляющий электрод 3, который используется как цифровой вход элемента, подавался импульсный сигнал, где напряжение цифровой «1» принято 225 В, а напряжение цифрового «0» равно 0 В. Напряжение цифровой «1» выбрано меньше напряжения смещения подвижного электрода, так как для снятия на чувствительном электроде 1 сигнала по току необходима разность потенциалов между 1–2 и 2–3 электродами.

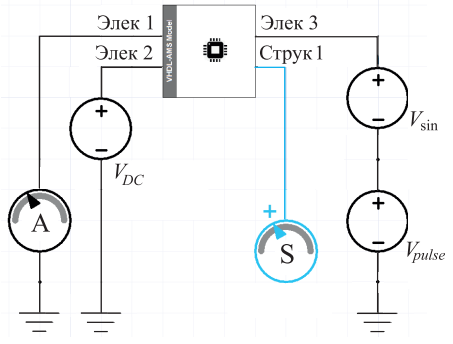


Рис. 5. Подключение микрорезонатора и внешних источников для выполнения логической операции НЕ

Fig. 5. Connecting a microresonator and external sources to perform a NOT logic operation

Если частота синусоидального сигнала возбуждения на приводном электроде 3 равна 314,89 кГц (рис. 4), только входной сигнал «0» приводит к совпадению между вынужденной резонансной частотой подвижного электрода и частотой сигнала возбуждения, что соответственно вызывает резонанс, и детектируется «высокий» выходной сигнал по току на чувствительном электроде 1. В случае, когда входной сигнал «1», – выходной сигнал будет слабым. Следовательно, резонатор в данном случае работает как логический вентиль НЕ (NOT). На рис. 6 представлена осциллограмма, демонстрирующая работу вентиля НЕ на основе разработанной модели резонатора.

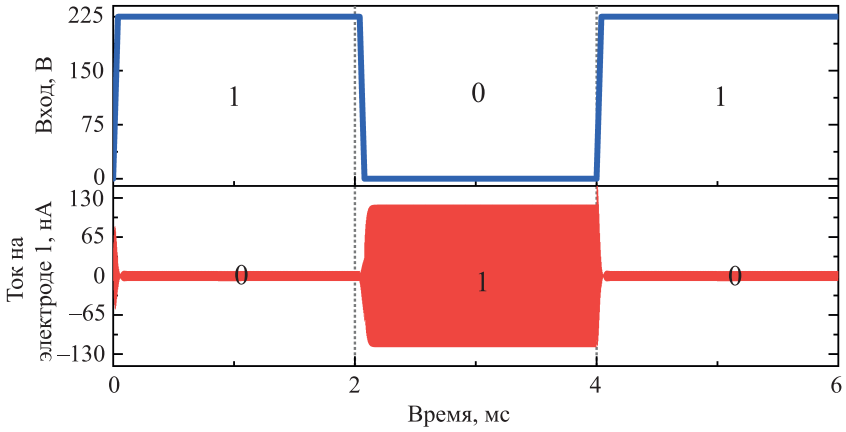


Рис. 6. Осциллограмма, демонстрирующая реализацию логической операции НЕ
Fig. 6. Oscilloscope trace demonstrating the implementation of the NOT logic operation

Девятиконтактный МЭМС-резонатор

Для реализации мультивходового логического вентиля необходимо большее количество боковых электродов. На рис. 7 изображен девятиконтактный МЭМС-резонатор, который представляет собой систему «пружина–масса–демпфер», состоящую из восьми боковых электродов – по четыре с каждой стороны от недеформируемой сейсмической массы, подвешенной на четырех подвесах.

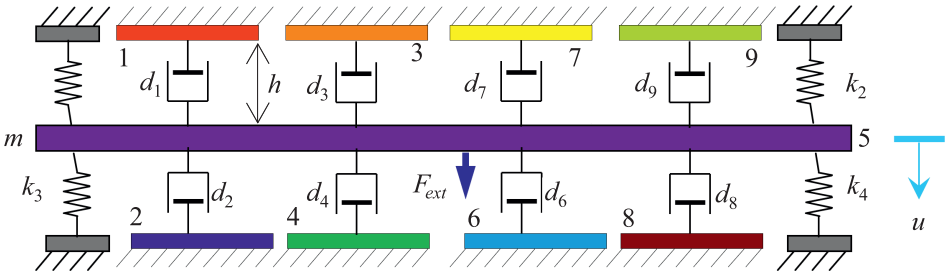


Рис. 7. Эквивалентная механическая схема девятиконтактного резонатора
Fig. 7. Equivalent mechanical circuit of a nine-terminal resonator

На рис. 8 показан анализ схлопывания, иллюстрирующего смещение подвижного электрода под воздействием меняющегося напряжения. Рассмотрены два случая:

1) электрод 1 имеет нулевой потенциал, а остальные электроды – плавающие, и линейно меняющееся напряжение подается на подвижный электрод;

2) электрод 2 имеет нулевой потенциал, а остальные электроды – плавающие, и линейно меняющееся напряжение подается на подвижный электрод.

На рис. 9 представлены четыре случая:

1) электрод 2 имеет нулевой потенциал, остальные электроды – плавающие, линейно меняющееся напряжение подается на подвижный электрод 5;

2) электроды 2 и 4 имеют нулевые потенциалы, остальные электроды – плавающие, линейно меняющееся напряжение подается на подвижный электрод;

3) электроды 2, 4 и 6 имеют нулевые потенциалы, остальные электроды – плавающие, линейно меняющееся напряжение подается на подвижный электрод;

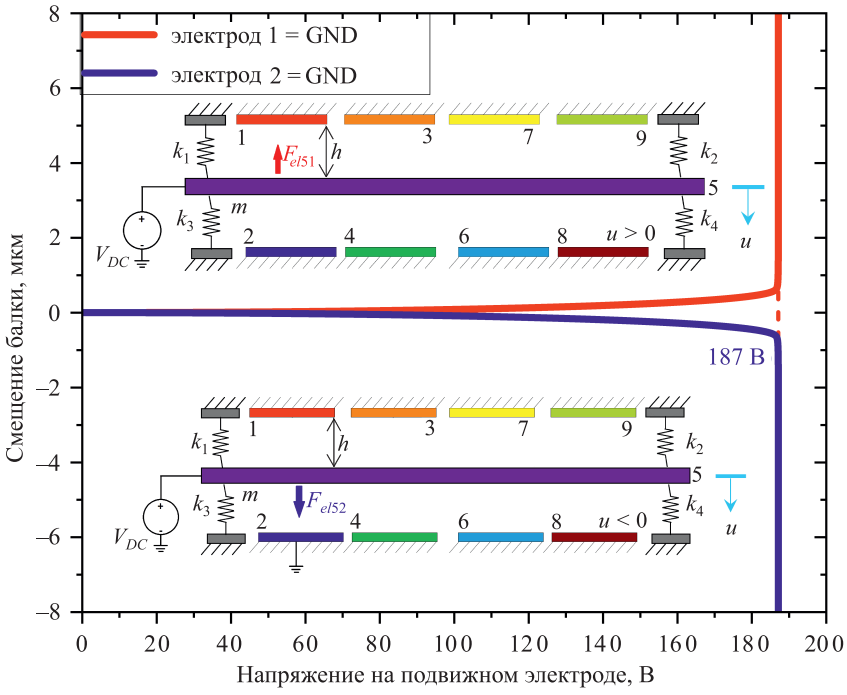


Рис. 8. Анализ эффекта схлопывания при смещении двух электродов с разных сторон балки
Fig. 8. Analysis of the pull-in effect due to the displacement of two electrodes on opposite sides of the beam

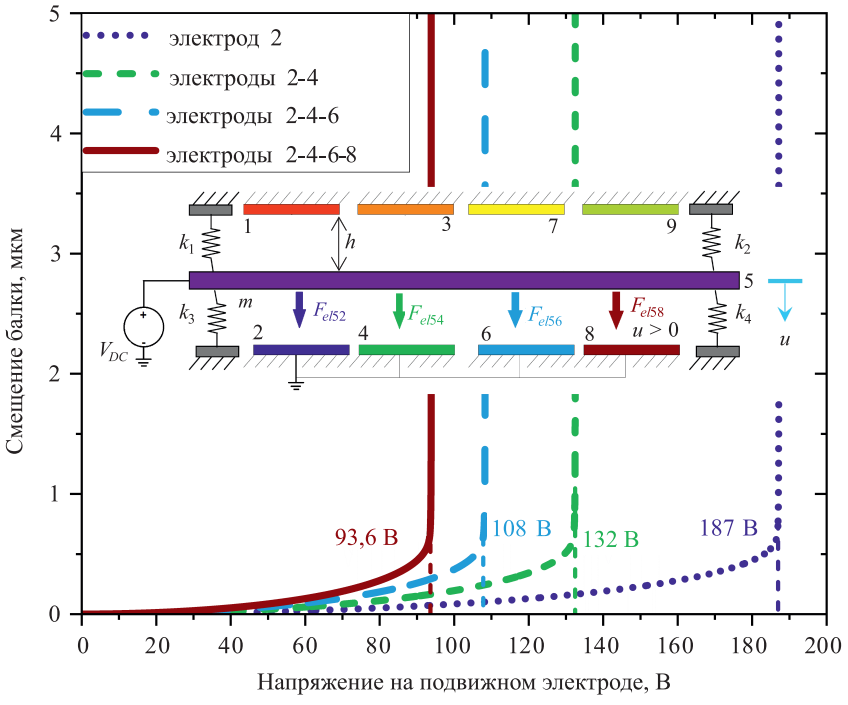


Рис. 9. Анализ эффекта схлопывания при смещении четырех электродов с одной стороны
Fig. 9. Analysis of the pull-in effect due to the displacement of four electrodes on one side

4) электроды 2, 4, 6 и 8 имеют нулевые потенциалы, остальные электроды – плавающие, линейно меняющееся напряжение подается на подвижный электрод.

Как видно из рис. 9, в первом случае, когда оба неподвижных электрода имеют нулевые потенциалы, напряжение схлопывания составляет 187 В, что меньше, чем во втором случае (132 В). Это объясняется удвоением площади подключенных электродов и, следовательно, увеличением электростатической силы между подвижным и боковыми электродами. В третьем и четвертом случаях площадь подключенных электродов больше, напряжение схлопывания становится меньше и составляет 108,0 и 93,6 В соответственно.

На рис. 10 представлены амплитудно-частотные характеристики девятиконтактного МЭМС-резонатора с собственной частотой 110 кГц, массой 22 мкг, при различных комбинациях на цифровых входах.

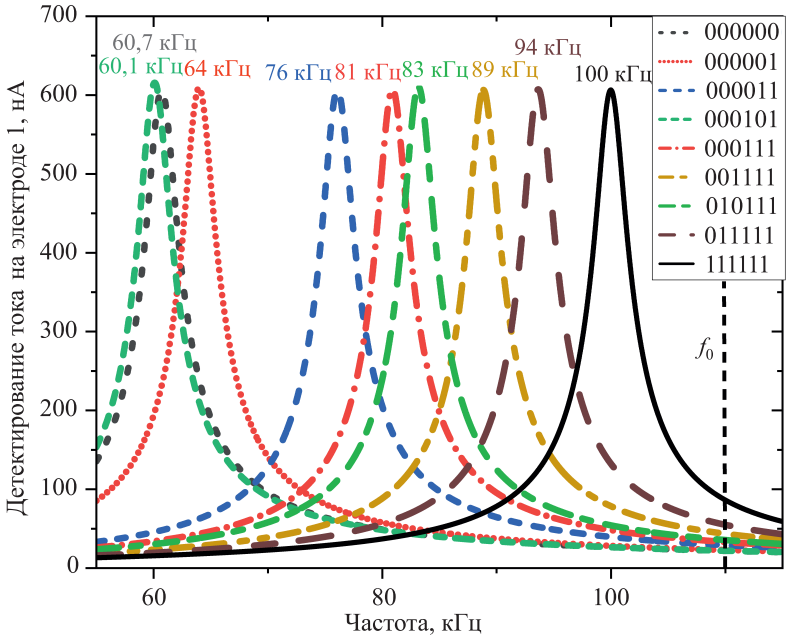


Рис. 10. Амплитудно-частотные характеристики девятиконтактного резонатора при различных входных комбинациях цифровых входов: «1» – 95 В, «0» – 0 В, V_4 = 1 В (AC), V_5 = 95 В (DC), Q = 40
Fig. 10. Amplitude-frequency characteristics of a nine-contact resonator with different input combinations of digital inputs: «1» – 95 V, «0» – 0 V, V_4 = 1 V (AC), V_5 = 95 V (DC), Q = 40

Данный резонатор позволяет реализовать, например, шестивходовый логический вентиль И, осциллограмма которого представлена на рис. 11.

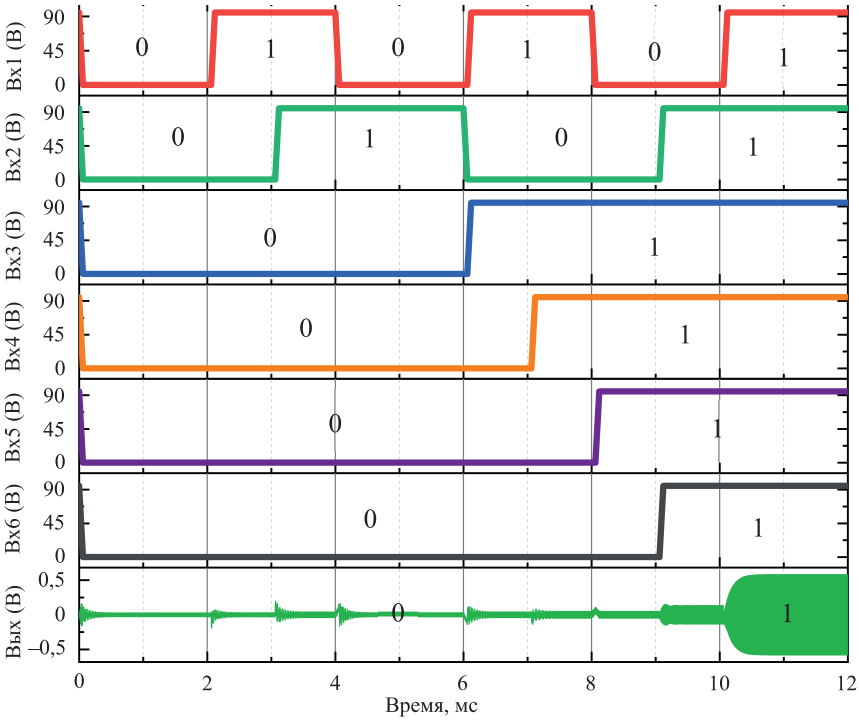


Рис. 11. Осциллограмма, демонстрирующая работу шестивходового логического вентиля И
Fig. 11. Oscillogram showing the operation of a six-input AND gate

Заключение

1. Разработаны системные математические модели логических МЭМС-вентилей на основе кремниевых резонаторов с электростатическим возбуждением. Для создания моделей использован язык описания аппаратуры VHDL-AMS, что позволило точно описать электромеханическое поведение резонаторов и провести анализ их статических, частотных и динамических характеристик. Создана библиотека конструкционных макромоделей, включающая трех-, пяти- и девятиконтактные МЭМС-резонаторы, которые могут быть использованы для реализации различных логических операций, таких как НЕ, И, ИЛИ-НЕ и исключающее ИЛИ.

2. Проведенные анализ эффекта схлопывания электродов и гармонический анализ подтвердили возможность управления резонансной частотой резонаторов с помощью электростатического возбуждения. Это открывает новые перспективы для создания энергоэффективных и перепрограммируемых логических элементов, которые могут конкурировать с традиционными КМОП-технологиями. Результаты симуляций демонстрируют, что разработанные модели позволяют точно прогнозировать поведение МЭМС-резонаторов в различных режимах работы, что является важным шагом в направлении их практического применения.

3. Практическая значимость исследований заключается в возможности использования разработанных моделей для проектирования сложных цифровых схем, таких как аналого-цифровые и цифро-аналоговые преобразователи, триггеры и арифметико-логические устройства на основе МЭМС-технологий. Это может способствовать дальнейшему развитию микроэлектромеханических систем в направлении создания более компактных, энергоэффективных и функционально гибких устройств.

4. Результаты исследования подтверждают перспективность использования МЭМС-резонаторов в качестве логических элементов и открывают новые возможности для их применения в современных микроэлектронных системах.

5. Работа выполнена при поддержке Министерства науки и высшего образования Российской Федерации (государственное задание для университетов № ФГФЗ-2023-0005) и с применением оборудования Центра коллективного пользования РТУ МИРЭА (соглашение от 01.09.2021 № 075–15-2021-689, уникальный идентификационный номер 2296.61321X0010).

Список литературы

1. Temperature Measurement System Comprising a Resonant MEMS Device: pat. USA US20110210801, IPC H03H9/2463 / Xavier Rottenberg, Roelof Jansen, Hendrikus Tilmans; Applicant and Patent Holder Interuniversitair Microelektronica Centrum vzw IMEC – US13/033,932. Claimed 24-02-2011; Published 01-09-2011.
2. Relays Do Not Leak: CMOS Does / H. Fariborzi [et al.] // Proceedings of the 50th Annual Design Automation Conference. 2013. No 127. P. 1–4.
3. Demonstration of Integrated Micro-Electro-Mechanical Relay Circuits for VLSI Applications / M. Spencer [et al.] // IEEE J. Solid-State Circuits. 2011. Vol. 46, No 1. P. 308–320.
4. Ahmed, S. Micro-Electromechanical Resonator-Based Logic and Interface Circuits for Low Power Applications / S. Ahmed // Ph. D. Thesis, KAUST Saudi Arabia. 2020.
5. Modeling and Simulation of a MEMS Resonator Based Reprogrammable Logic Gate Using Partial Electrodes / S. Ahmed [et al.] // Conference: Symposium on Design, Test, Integration and Packaging of MEMS/MOEMS. Paris, France, 2019. P. 1–5.
6. Hafiz, M. Microelectromechanical Reprogrammable Logic Device / M. Hafiz, L. Kosuru, M. Younis // Nature Communications. 2016. No 7. P. 1–9.
7. Соловьев, А. А. Системное моделирование мультиконтактного микроэлектромеханического логического элемента / А. А. Соловьев, Е. Ф. Певцов, В. А. Колчужин // Нано- и микросистемная техника. 2024. Т. 26, № 6. С. 260–267.
8. LMGT_MEMS_component_library [Electronic Resource]. Mode of access: https://github.com/Kolchuzhin/LMGT_MEMS_component_library/tree/master/resonator/transducer_e5. Date of access: 10.11.2024.

Поступила 08.01.2025 Принята в печать 15.04.2025

References

1. Xavier Rottenberg, Roelof Jansen, Hendrikus Tilmans (2011) Temperature Measurement System Comprising a Resonant MEMS Device. *Patent USA US20110210801, IPC H03H9/2463, Applicant and Patent Holder Interuniversitair Microelektronica Centrum vzw IMEC – US13/033,932*. Claimed 2011-02-24; Published 2011-09-01.
2. Fariborzi H., Chen F., Nathanael R., Chen I., Hutin L., Lee R., et al. (2013) Relays Do Not Leak: CMOS Does. *Proceedings of the 50th Annual Design Automation Conference*. 127, 1–4.
3. Spencer M., Chen F., Wang C., Nathanael R., Fariborzi H., Gupta A., et al. (2011) Demonstration of Integrated Micro-Electro-Mechanical Relay Circuits for VLSI Applications. *IEEE J. Solid-State Circuits*. 46 (1), 308–320.
4. Ahmed S. (2020) Micro-Electromechanical Resonator-Based Logic and Interface Circuits for Low Power Applications. *Ph. D. Thesis, KAUST Saudi Arabia*.
5. Ahmed S., Li R., Zou X., Al Hafiz M. A., Fariborzi H. (2019) Modeling and Simulation of a MEMS Resonator Based Reprogrammable Logic Gate Using Partial Electrodes. *Conference: Symposium on Design, Test, Integration and Packaging of MEMS/MOEMS*. Paris, France. 1–5.
6. Hafiz M., Kosuru L., Younis M. (2016) Microelectromechanical Reprogrammable Logic Device. *Nature Communications*. (7), 1–9.
7. Solovev A. A., Pevtsov E. F., Kolchuzhin V. A. (2024) System Modeling of a Multi-Contact Microelectromechanical Logic Element. *Nano- and Microsystems Technology*. 26 (6), 260–267 (in Russian).
8. LMGT_MEMS_component_library. Available: https://github.com/Kolchuzhin/ LMGT_MEMS_component_library/tree/master/resonator/transducer_e5 (Accessed 10 November 2024).

Received: 8 January 2025 Accepted: 15 April 2025

Вклад авторов / Authors' contribution

Авторы внесли равный вклад в написание статьи / The authors contributed equally to the writing of the article.

Сведения об авторах

Соловьев А. А., ст. преп. каф. наноэлектроники Института перспективных технологий и индустриального программирования, МИРЭА – Российский технологический университет (РТУ МИРЭА)

Певцов Е. Ф., канд. техн. наук, доц., дир. центра проектирования, РТУ МИРЭА

Колчужин В. А., канд. техн. наук, вед. инж.-констр. Qorvo Munich GmbH

Адрес для корреспонденции

119454, Российская Федерация,
Москва, просп. Вернадского, 78
МИРЭА – Российский
технологический университет
Tel.: +7 967 097-18-73
E-mail: solovev_aa@mirea.ru
Соловьев Александр Анатольевич

Information about the authors

Solovev A. A., Senior Lecturer at the Department of Nanoelectronics of the Institute of Advanced Technologies and Industrial Programming, MIREA – Russian Technological University (MIREA – RTU)

Pevtsov E. F., Cand. Sci. (Tech), Associate Professor, Director of the Design Center, MIREA – RTU

Kolchuzhin V. A., Cand. Sci. (Tech), Senior Desing Engineer, Qorvo Munich GmbH

Address for correspondence

119454, Russian Federation,
Moscow, Vernadsky Ave., 78
MIREA – Russian
Technological University
Tel.: +7 967 097-18-73
E-mail: solovev_aa@mirea.ru
Solovev Aleksandr Anatolevich

CC BY

http://dx.doi.org/10.35596/1729-7648-2025-23-3-36-45

УДК 621.396.218:614.89.086.5

АНАЛИЗ ЭЛЕКТРОМАГНИТНОГО ФОНА, СОЗДАВАЕМОГО СОЗВЕЗДИЯМИ НИЗКООРБИТАЛЬНЫХ СПУТНИКОВ СВЯЗИ, НА ОСНОВЕ ПРОГНОЗА ОБРАЗУЕМОЙ ИМИ ТЕРРИТОРИАЛЬНОЙ ПЛОТНОСТИ ТРАФИКА

В. И. МОРДАЧЕВ, Д. А. ЦИОНЕНКО

Белорусский государственный университет информатики и радиоэлектроники (Минск, Республика Беларусь)

Аннотация. Представлена методика, позволяющая анализировать характеристики электромагнитного фона у земной поверхности, создаваемого излучениями мегасозвездий низкоорбитальных спутников связи (НСС), на основе оценок средней территориальной плотности беспроводного трафика, создаваемого ими на земной поверхности, и доступных системных характеристик НСС и их группировок, таких как высота орбиты, характеристики диаграмм направленности антенн, ограничения на угол, под которым может быть направлен главный лепесток антенны НСС к поверхности Земли, и особенности сценариев обслуживания абонентских терминалов. Полученные оценки средних уровней электромагнитного фона, соответствующие диапазону реальных значений средней плотности беспроводного трафика, создаваемого созвездиями НСС на земной поверхности, существенно превышают уровни естественного электромагнитного фона. Это соответствует оценкам, полученным ранее с использованием данных о полной излучаемой мощности НСС и их количестве в мегасозвездии, и подтверждает адекватность представленной методики.

Ключевые слова: низкоорбитальный спутник, группировка, электромагнитное излучение, диаграмма направленности антенны, территориальная плотность трафика, электромагнитный фон.

Конфликт интересов. Авторы заявляют об отсутствии конфликта интересов.

Для цитирования. Мордачев, В. И. Анализ электромагнитного фона, создаваемого созвездиями низкоорбитальных спутников связи, на основе прогноза образуемой ими территориальной плотности трафика / В. И. Мордачев, Д. А. Ционенко // Доклады БГУИР. 2025. Т. 23, № 3. С. 36–45. http://dx.doi.org/10.35596/1729-7648-2025-23-3-36-45.

ANALYSIS OF THE ELECTROMAGNETIC BACKGROUND CREATED BY CONSTELLATIONS OF LOW-EARTH ORBIT SATELLITES, BASED ON THE PREDICTION OF THE CREATED AREA TRAFFIC CAPACITY

VLADIMIR I. MORDACHEV, DZMITRY A. TSYANENKA

Belarusian State University of Informatics and Radioelectronics (Minsk, Republic of Belarus)

Abstract. The paper presents a technique for analyzing the characteristics of the electromagnetic background near the earth's surface generated by emissions from megaconstellations of low-orbit communication satellites (LOCSS) based on estimates of the average territorial density of wireless traffic generated by them on the earth's surface and available system characteristics of the LOCSS and their constellations, such as orbit altitude, characteristics of antenna patterns, limitations on the angle at which the main lobe of the LOCSS antenna can be directed to the earth's surface, and features of subscriber terminal servicing scenarios. The obtained estimates of the average levels of the electromagnetic background, corresponding to the range of real values of the average density of wireless traffic generated by LOCSS constellations on the earth's surface, significantly exceed the levels of the natural electromagnetic background. This is consistent with the estimates obtained earlier using data on the total radiated power of LOCSS and their number in the megaconstellation, and confirms the adequacy of the presented technique.

Keywords: low-earth orbit satellite, constellation, electromagnetic radiation, antenna pattern, area traffic capacity, electromagnetic background.

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

For citation. Mordachev V. I., Tsyanenka D. A. (2025) Analysis of the Electromagnetic Background Created by Constellations of Low-Earth Orbit Satellites, Based on the Prediction Forecast of the Created Area Traffic Capacity. *Doklady BGUIR*. 23 (3), 36–45. http://dx.doi.org/10.35596/1729-7648-2025-23-3-36-45 (in Russian).

Введение

Создание в околоземном пространстве мегагруппировок низкоорбитальных спутников связи (НСС) (Starlink, OneWeb, Astra, Kuiper, GuoWang, «Сфера» и др.) увеличит на несколько порядков количество космических источников электромагнитного излучения (ЭМИ) диапазона СВЧ в направлении земной поверхности. Подобное «развитие событий» способно привести к недопустимому усложнению электромагнитной обстановки (ЭМО) для систем различных радиослужб, функционирующих в полосах частот этих группировок на земной поверхности и в воздушном пространстве на вторичной основе, а также к существенному росту интенсивности электромагнитного фона (ЭМФ), способному обострить проблемы электромагнитной безопасности населения и электромагнитной экологии среды обитания, особенно при интеграции этих систем спутниковой связи (ССС), обеспечивающих высокие скорости и относительно низкие задержки передачи информации наземным абонентским терминалам (АТ), с глобальными системами мобильной связи 5G/6G.

В [1, 2] разработана методика анализа интенсивности ЭМФ, создаваемого созвездиями НСС, на основе оценок количества космических аппаратов (КА) в группировке, полной излучаемой мощности КА, параметров ЭМИ в главном и боковых лепестках диаграммы направленности его антенны (ДНА), высоты орбиты КА и ограничений на угол, под которым могут быть направлены главные лепестки (ГЛ) излучения антенн КА к поверхности Земли при обслуживании АТ. Методика позволяет анализировать статистические и интегральные энергетические характеристики ЭМО, представляемой в виде ансамбля N радиочастотных электромагнитных полей (ЭМП), создаваемых группировками НСС (НГСС) в точке наблюдения (ТН) у земной поверхности. Недостатком методики, предложенной в [1, 2], является определенная трудность в сборе необходимых исходных данных, иногда относящихся к разряду конфиденциальных.

Цель исследований – разработка методики, позволяющей анализировать характеристики ЭМФ у земной поверхности на основе оценок средней территориальной плотности нисходящего трафика АТС (бит/с/м^2) (Area Traffic Capacity), создаваемого НГСС на земной поверхности, и доступных системных характеристик НГСС, таких как высота орбит КА, количество орбитальных оболочек, характеристики диаграммы направленности антенн КА и АТ и ограничения на угол места, при котором может быть реализован канал связи КА–АТ.

Методика анализа

При анализе характеристик ЭМФ у земной поверхности, создаваемых излучениями НГСС, воспользуемся моделями [1, 2] и подходом [3] к оценке уровня ЭМФ, создаваемого системами мобильной связи, на основе прогноза АТС.

Необходимая мощность полезного сигнала на входе радиоприемника абонентского терминала

1. В соответствии с теоремой Шеннона – Хартли потенциальная пропускная способность радиоканала V_P (бит/с) определяется следующими соотношениями:

$$V_P \approx \Delta F_R \log_2\left(1+SNIR_P\right);\ \upsilon_P = \frac{V_P}{\Delta F_R} \approx \log_2\left(1+SNIR_P\right);\ SNIR_P \approx \frac{p_0}{p_N + p_{INT}} = 2^{\upsilon_P} - 1, \qquad (1)$$

где υ_P – потенциальная спектральная эффективность, бит/с/Гц, передачи информации в радиоканале с шириной полосы частот ΔF_R, Гц,; $SNIR_P$ – отношение «сигнал/(шум+помеха)» (Signal-to-Noise-plus-Interference Ratio); $p_0 \approx$ const, $p_N = kT_0K_N \approx$ const и $p_{INT} \approx$ const, Вт/Гц – спектральные плотности мощности радиосигнала, собственного шума приемника и помех в полосе ΔF_R соответственно; k – постоянная Больцмана, 1,38 · 10^{-23} Дж/К; K_N – коэффициент шума радиоприемника, ед.; T_0 – температура окружающей среды, К (T_0 = 290 К).

2. Реальная скорость V_R передачи информации в канале связи в m раз меньше потенциальной пропускной способности радиоканала V_P; во столько же раз реальная спектральная эффективность υ_P передачи информации в канале меньше потенциальной υ_P. Обеспечение скорости передачи данных, равной V_P при постоянном ΔF_R, требует значительного увеличения $SNIR_R$ в реальном канале по сравнению с потенциальным уровнем $SNIR_P$. В этом случае:

$$V_R = \frac{\upsilon_P}{m} = \frac{\log_2\left(1+SNIR_P\right)}{m} = \frac{\log_2\left(1+SNIR_R\right)^{1/m}}{m} \approx \frac{\log_2 SNIR_R}{m^2};\ SNIR_R \approx SNIR_P^m. \quad (2)$$

Параметр m в (2), отражающий соотношение потенциальных и реальных характеристик радиоканала, позволяет учесть вклад технологии MIMO [4] в повышение спектральной эффективности передачи информации по радиоканалу и в некоторых случаях может быть меньше единицы. Но в радиоканалах сотовой (мобильной) связи без использования технологии MIMO $m \approx 2$–10 [5], и достигаемое увеличение спектральной эффективности в этих радиоканалах за счет технологии MIMO в 2–8 раз фактически позволяет лишь скомпенсировать неидеальность процессов модуляции/демодуляции и кодирования-декодирования. Поэтому интерес представляет выполнение оценок для $m = 1$ в предположении, что скорость передачи данных V_R в радиоканалах ССС близка к потенциальной V_P в определении (1): $V_R \approx V_P$.

3. Минимально необходимая мощность полезного сигнала P_0 в радиоканале (реальная чувствительность радиоприема), при которой обеспечивается пропускная способность V_P радиоканала с тепловым шумом мощностью $P_N \approx p_N\,\Delta F_R$ и помехой мощностью $P_{INT} \approx p_{INT}\,\Delta F_R$, определяется следующим соотношением (пренебрегая различиями влияния внутрисетевой помехи и теплового шума на пропускную способность радиоканала):

$$P_0 = p_{N\Sigma}\Delta F_R\left(2^{\upsilon_P}-1\right) = p_{N\Sigma}\Delta F_R SNIR_P,\ p_{N\Sigma} = \left(K_{CC}+1\right)p_N,\ p_0/p_{N\Sigma} = SNIR_P, \quad (3)$$

где $p_{N\Sigma}$ – суммарная спектральная плотность мощности собственного шума и помехи в радиоканале; K_{CC} – коэффициент, характеризующий превышение уровнем помехи уровня теплового шума, $K_{CC} \approx P_{INT}/P_N$.

Величина K_{CC} определяется качеством частотно-пространственного планирования ССС, функционированием в соседних радиоканалах других ССС, а также наличием функционирующих в тех же полосах частот систем других радиослужб на вторичной основе, и может принимать значения в широких пределах от 0 (внутрисетевая помеха отсутствует) до 10 и более при низком качестве обеспечения внутрисистемной и межсистемной электромагнитной совместимости (ЭМС).

Энергия на бит при приеме информации абонентским терминалом

При минимально необходимой мощности полезного сигнала P_0 на входе радиоприемника АТ минимально необходимая энергия сигнала для приема одного бита информации со скоростью V_P, бит/с, должна быть не менее следующего значения, Дж/бит:

$$E_{br} = \frac{P_0}{V_P} = \frac{\left(K_{CC}+1\right)p_N\left(2^{\upsilon_P}-1\right)}{\upsilon_P} = \frac{\left(K_{CC}+1\right)p_N SNIR_P}{\log_2\left(1+SNIR_P\right)}. \quad (4)$$

Поскольку энергия, необходимая для приема АТ одного бита информации, может быть рассчитана через плотность потока мощности (ППМ) излучения КА у земной поверхности и эквивалентную площадь приемной антенны АТ, можно определить значение ППМ и соответствующую мощность излучения в ГЛ антенны КА, которая необходима для передачи одного бита информации обслуживаемому АТ, находящемуся на поверхности Земли. Для этого используем следующую идеализированную двухуровневую модель ДНА АТ [1, 2], в которой ГЛ имеет коническую форму шириной $\Delta\varphi_{\text{АТ}}$, а соотношение мощностей в режиме приема по главному ($P_{M\,\text{АТ}}$) и боковым ($P_{S\,\text{АТ}}$) лепесткам ДНА равно $C_{P\,\text{АТ}}$:

$$G_{M\,\text{АТ}} = \frac{C_{P\,\text{АТ}}}{\left(1+C_{P\,\text{АТ}}\right)\sin^2\left(0{,}25\Delta\varphi_{\text{АТ}}\right)};\ G_{S\,\text{АТ}} = \frac{1}{\left(1+C_{P\,\text{АТ}}\right)\cos^2\left(0{,}25\Delta\varphi_{\text{АТ}}\right)};\ C_{P\,\text{АТ}} = P_{M\,\text{АТ}}/P_{S\,\text{АТ}}, \quad (5)$$

где $G_{M\,\text{АТ}}$ – коэффициент усиления (КУ) антенны АТ по ГЛ (КПД антенны $\eta = 1$ в рамках рассматриваемой модели); $G_{S\,\text{АТ}}$ – КУ антенны АТ по уровню боковых лепестков (БЛ), принимаемый

постоянным во всех направлениях вне ГЛ, в этой модели при относительно малой ширине ГЛ ($\Delta\varphi_{\text{АТ}} \le 20°$) уровень БЛ мало зависит от $\Delta\varphi_{\text{АТ}}$: $G_{S\,AT} \approx 1/(1+C_{P\,\text{АТ}})$.

Эффективная площадь $S_{e\,\text{АТ}}$ антенны наземного АТ при приеме сигнала КА главным коническим лепестком шириной $\Delta\varphi_{\text{АТ}}$ определяется по формуле

$$S_{e\,\text{АТ}} = \frac{G_{M\,\text{АТ}}\lambda^2}{4\pi} = \frac{C_{P\,\text{АТ}}\lambda^2}{4\pi\left(1+C_{P\,\text{АТ}}\right)\sin^2(0{,}25\Delta\varphi_{\text{АТ}})}. \quad (6)$$

Таким образом, ППМ Z_E, Вт/м^2, ЭМИ, создаваемая у земной поверхности в точке размещения обслуживаемого АТ, при $m = 1$ может быть определена следующим образом:

$$Z_E = \frac{P_0}{S_{e\,\text{АТ}}} = \frac{4\pi P_0}{G_{M\,\text{АТ}}\lambda^2} = \frac{4\pi\Delta F_R\left(K_{CC}+1\right)p_N\left(2^{\upsilon_P}-1\right)}{G_{M\,\text{АТ}}\lambda^2} = \frac{4\pi\Delta F_R\left(K_{CC}+1\right)p_N SNIR_p}{G_{M\,\text{АТ}}\lambda^2}. \quad (7)$$

Оценки излучаемой энергии на бит информации на основе учета потерь при РРВ

Используя для представления характеристик излучения КА двухуровневую модель (5) с параметрами $P_{M\,SC}$, $P_{S\,SC}$, $C_{P\,SC}$ (мощности излучения по Гл и БЛ, а также их отношение соответственно), $G_{M\,SC}$, $G_{S\,SC}$ (уровни ГЛ и БЛ соответственно), определим требуемую излучаемую энергию на бит передаваемой информации с учетом затухания сигнала на радиолинии КА–АТ. Для того чтобы сигнал на входе приемника имел энергетику (4), необходимо выполнение следующих условий:

– радиосигналы должны излучаться передатчиками КА с мощностью, компенсирующей базовые потери при распространении радиоволн (РРВ) от антенны передатчика КА к антенне приемника АТ [6]; применительно к рассматриваемому случаю ограничимся учетом потерь L_{bf} при РРВ в свободном пространстве (поскольку радиолинии КА–АТ диапазона СВЧ и КВЧ являются радиолиниями прямой видимости, лишь в малой части захватывают плотные слои атмосферы и в большинстве случаев реализуются на частотах, на которых затуханием сигналов в атмосфере можно пренебречь, в том числе в окнах прозрачности нижней части миллиметрового диапазона);

– мощность излучения может быть уменьшена при использовании направленных антенн КА и АТ, поскольку потери L при передаче сигнала по радиолинии определяются известным соотношением [6]

$$L = L_{bf}\big/\left(G_T G_R\right), \quad (8)$$

где $G_T = G_{M\,SC}$, $G_R = G_{M\,\text{АТ}}$ – коэффициенты усиления передающих антенн КА и приемных антенн АТ соответственно.

Затухание при РРВ с длиной волны λ в свободном пространстве на расстояние R и потери при передаче сигнала по радиолинии определяются по формулам [7]:

$$L_{bf} = \left(\frac{4\pi R}{\lambda}\right)^2; \quad L = \frac{L_{bf}}{G_{M\,SC}G_{M\,\text{АТ}}} = \frac{16\pi^2 R^2}{\lambda^2 G_{M\,SC}G_{M\,\text{АТ}}}. \quad (9)$$

Излучаемая энергия на бит передаваемой информации должна быть обеспечена на уровне, не ниже $E_{bt} = E_{br}L$

$$E_{bt}\left(R\right) = \frac{E_{br}L_{bf}}{G_{M\,SC}G_{M\,\text{АТ}}} = \frac{16\pi^2\left(K_{CC}+1\right)p_N\left(2^{\upsilon_P}-1\right)R^2}{\lambda^2 G_{M\,SC}G_{M\,\text{АТ}}\upsilon_P} = \frac{16\pi^2\left(K_{CC}+1\right)p_N SNIR_P R^2}{\lambda^2 G_{M\,SC}G_{M\,\text{АТ}}\log_2\left(1+SNIR_P\right)}. \quad (10)$$

Если скорость передачи информации по радиоканалу КА–АТ с полосой частот ΔF_R соответствует его потенциально пропускной способности (1), то мощность $P_{M\,SC}(R)$, излучаемая в ГЛ ДНА КА, и полная излучаемая мощность (total radiated power) КА $P_{TR\,SC}(R)$, ($P_{TR\,SC} = P_{M\,SC} + P_{S\,SC}$) могут быть определены следующим образом в зависимости от расстояния R между КА и АТ:

$$\left.\begin{aligned} &P_{M\,SC}\left(R\right) = E_{bt}\left(R\right)V_P = YR^2, \quad P_{TR\,SC}\left(R\right) = P_{M\,SC}\left(R\right)\frac{1+C_{P\,SC}}{C_{P\,SC}}; \\ &Y = \frac{16\pi^2\left(K_{CC}+1\right)p_N\left(2^{\upsilon_P}-1\right)\Delta F_R}{\lambda^2 G_{M\,SC}G_{M\,\text{АТ}}} = \frac{16\pi^2\left(K_{CC}+1\right)p_N SNIR_P\Delta F_R}{\lambda^2 G_{M\,SC}G_{M\,\text{АТ}}}. \end{aligned}\right\} \quad (11)$$

Для получения зависимости полной излучаемой мощности $P_{TR\ SC}(V_R)$ в реальном радиоканале КА–АТ от его требуемой реальной пропускной способности V_R при заданных параметрах, входящих в (11), модифицируем соотношение (10) с использованием (1), (2) следующим образом:

$$E_{bt}(R) = \frac{16\pi^2 (K_{CC}+1) p_N \left(2^{m\upsilon_R}-1\right) R^2}{\lambda^2 G_{M\ SC} G_{M\ \mathrm{AT}} m\upsilon_R} = \frac{16\pi^2 (K_{CC}+1) p_N m\left[(1+SNIR_R)^{1/m}-1\right] R^2}{\lambda^2 G_{M\ SC} G_{M\ \mathrm{AT}} \log_2(1+SNIR_R)}. \quad (12)$$

В приведенных соотношениях также требуется учесть, что расстояние R между КА и АТ определяется не только высотой H_S орбиты КА и углом места θ, под которым его ГЛ направлен к земной поверхности, но и кривизной земной поверхности – ее сферичностью с радиусом R_E (рис. 1). Соотношение, связывающее величины K, R_E, H_S, записывается в виде:

$$R = \sqrt{R_E^2 \sin^2\theta + H_S^2 + 2R_E H_S} - R_E \sin\theta. \quad (13)$$

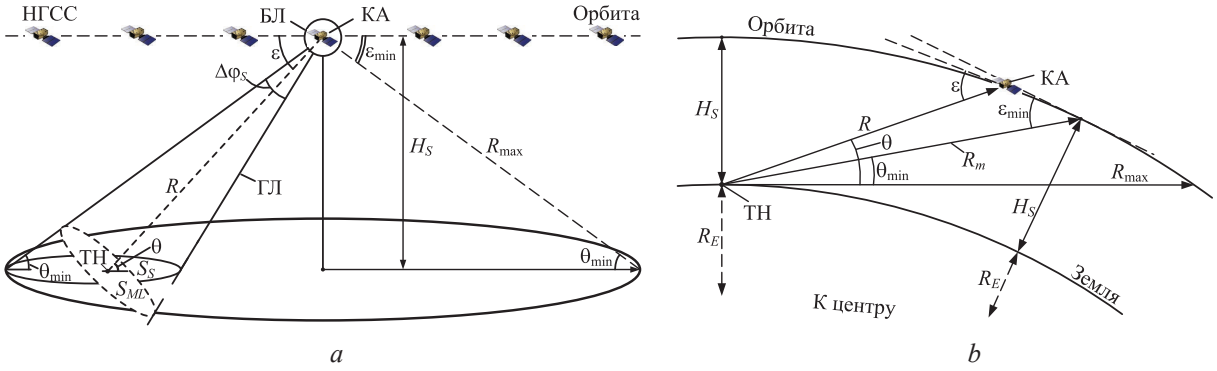


Рис. 1. Взаимное расположение КА-источника электромагнитного излучения и наземной точки наблюдения: *a* – формирование «пятна» главного лепестка на земной поверхности; *b* – учет сферичности земной поверхности

Fig. 1. Relative position of the low-orbit communication satellites of electromagnetic radiation and the ground observation point: *a* – formation of the "spot" of the main lobe on the earth's surface; *b* – taking into account the sphericity of the earth's surface

Средняя территориальная плотность мобильного трафика и уровень ЭМФ

В зависимости от сценария функционирования НГСС с круговой орбитой постоянной высоты H_S расстояние R между передающей антенной КА и ТН на земной поверхности, попадающей в «пятно облучения» ГЛ ДНА КА (рис. 1), может быть принято либо фиксированным, либо изменяющимся в определенных пределах.

Сценарий 1. Если КА использует узкий конический луч шириной $\Delta\varphi_{SC}$ с постоянным углом наклона ε по отношению к касательной к орбите или с фиксированным углом места θ направления прихода сигнала НГСС к наземным АТ (рис. 1), то расстояние от КА-источника ЭМИ до ТН в пятне облучения на земной поверхности практически постоянно: $R \approx$ const. Таким образом, при этом сценарии необходимая полная излучаемая мощность КА $P_{TR\ CS}(R)$ может быть определена непосредственно с использованием (11).

Сценарий 2. Если КА и АТ используют узкие конические лучи шириной $\Delta\varphi_{SC}$ и $\Delta\varphi_{\mathrm{AT}}$ с переменными углами наклона ε по отношению к касательной к орбите или с переменными углами места θ направления прихода сигнала НГСС к наземным АТ таким образом, чтобы в диапазоне углов места $\theta \in [\theta_{min}, 90°]$, $\theta_{min} \geq 0°$, ГЛ антенн КА и АТ были направлены друг на друга, то интервал значений расстояния между КА и АТ $R \in [H_S, R_m \leq R_{max}]$ (рис. 1). При равновероятном расположении КА в различных точках орбиты расстояние (13) также случайно с плотностью распределения вероятности (п.р.в.) $w(R)$, имеющей следующий вид [1, 2]:

$$w(R) = 2R\Big/\left(R_m^2 - H_S^2\right);\ H_S \leq R \leq R_m \leq R_{max};\ R_{max} = \sqrt{2R_E H_S + H_S^2}. \quad (14)$$

При анализе данного сценария необходимо учитывать следующие варианты.

• В случае наличия в системе регулировки мощности $P_{M\,LS}$ ($P_{M\,SC}(R) \approx P_{TR\,SC}(R)$) постоянной величиной является необходимая мощности P_0 полезного сигнала (3), определяющая заданную пропускную способность радиоканала КА–АТ. Случайные величины – необходимая полная излучаемая мощность КА $P_{TR\,SC}(R)$ и необходимый уровень излучаемой энергии на бит передаваемой информации $E_{bt}(R)$. П.р.в. $w(E_{bt})$ и матожидание $\langle E_{bt}\rangle$ могут быть определены по известным методикам [1, 2]:

$$w(E_{bt}) = \frac{1}{E_{bt\max} - E_{bt\min}} = \frac{1}{X\left(R_m^2 - H_S^2\right)};\ \langle E_{bt}\rangle = \frac{E_{bt\max} + E_{bt\min}}{2} = \frac{X\left(R_m^2 + H_S^2\right)}{2}; \tag{15}$$

$$X = \frac{16\pi^2\left(K_{CC}+1\right)p_N\left(2^{m\upsilon_R}-1\right)}{\lambda^2 G_{M\,SC}G_{M\,AT}m\upsilon_R} = \frac{16\pi^2\left(K_{CC}+1\right)p_N m\left[\left(1+SNIR_R\right)^{1/m}-1\right]}{\lambda^2 G_{M\,SC}G_{M\,AT}\log_2\left(1+SNIR_R\right)}. \tag{16}$$

• Средняя территориальная плотность мобильного трафика АТС$_A$, создаваемая на земной поверхности излучениями всех КА, принадлежащих ССС, которая состоит из N_{SC} КА, по нисходящим радиоканалам КА–АТ с емкостью каждого канала V_{CH}, бит/с, определяется по формуле:

$$\text{АТС}_A = V_{CH}N_{SC}\Big/\left(4\pi R_E^2\right). \tag{17}$$

Результаты анализа

Расчетные зависимости АТС$_A$(C_{CH}) для различного числа КА N_S приведены на рис. 2, *a*.

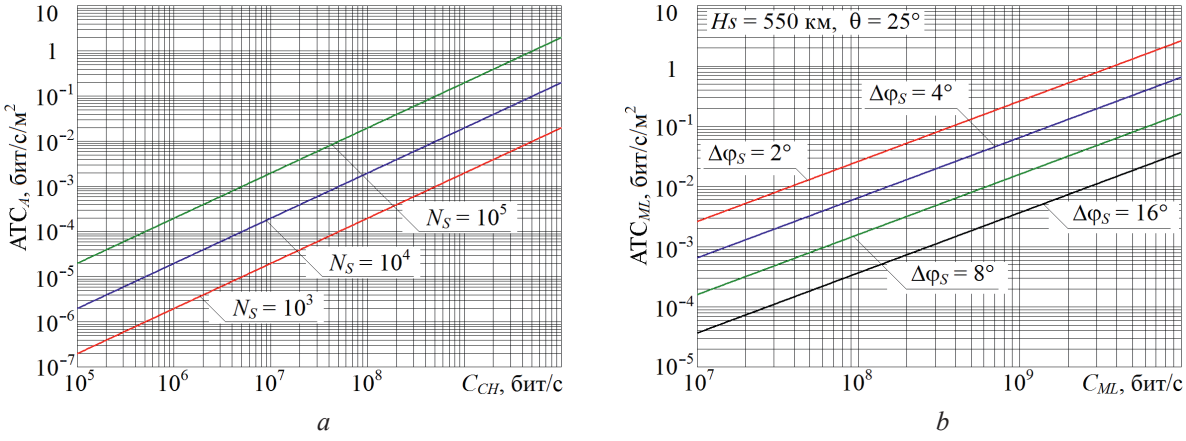


Рис. 2. Расчетные зависимости средней территориальной плотности беспроводного трафика от скорости передачи данных по радиоканалу НСС→АТ (C_{CH}) и в ГЛ НСС (C_{ML}): *a* – зависимости АТС$_A$(C_{CH}) при различном числе НСС в их мегасозвездии; *b* – зависимости АТС$_A$(C_{ML}) в «пятне» ГЛ ЭМИ НСС на земной поверхности при θ = 25°, H_S = 550 км и различной ширине ГЛ Δφ$_S$

Fig. 2. Calculated dependencies of the average area traffic capacity on the data transfer rate in the downlink radio channel (C_{CH}) and in the LEOS main lobe (C_{ML}): *a* – АТС$_A$(C_{CH}) dependencies for different LEOS numbers in their megaconstellation; *b* – АТС$_A$(C_{ML}) dependencies in the LEOS main lobe "spot" on the earth's surface for θ = 25°, H_S = 550 km and different main lobe width Δφ$_S$

Территориальная плотность мобильного трафика АТС$_{ML}$, создаваемая излучением единичного КА в пятне ГЛ на земной поверхности, равна отношению скорости передачи данных по нисходящему радиоканалу КА–АТ $V_{ML} = V_{CH}$ (бит/с) к площади пятна ГЛ на земной поверхности [1, 2]:

$$\text{АТС}_{ML} = \frac{V_{CH}}{S_S} \approx \frac{V_{CH}\,2\text{tg}\left(\varepsilon+0{,}5\Delta\varphi_S\right)\text{tg}\left(\varepsilon-0{,}5\Delta\varphi_S\right)\sin\left(\varepsilon\right)}{\pi H_S^2\left[\text{tg}\left(\varepsilon+0{,}5\Delta\varphi_S\right)-\text{tg}\left(\varepsilon-0{,}5\Delta\varphi_S\right)\right]\text{tg}\left(0{,}5\Delta\varphi_S\right)}. \tag{18}$$

Расчетные зависимости АТС$_{ML}$(C_{ML}) при различных значениях ширины Δφ$_S$ приведены на рис. 2, *b*. Эти зависимости свидетельствуют о том, что средняя территориальная плотность беспроводного трафика, которую может создавать НГСС у земной поверхности, на 6–8 порядков меньше уровня, декларируемого в [8] для систем мобильной связи 5G, и на 8–10 – деклари-

руемого в [9] для перспективных систем 6G. Если известна средняя территориальная плотность мобильного трафика АТС$_A$ (бит/с/м^2), то в данном сценарии, когда вся энергия, излучаемая КА в ГЛ, достигает земной поверхности, создаваемая этими излучениями ППМ, которая определяет ЭМФ от космического компонента НГСС у земной поверхности, вычисляется по формуле:

$$Z_{\text{АТС}} = \langle E_{bt} \rangle \text{АТС}_A = \frac{8\pi^2 \left(K_{CC}+1\right) p_N \left(2^{m\upsilon_R}-1\right)\left(R_m^2+H_S^2\right)}{\lambda^2 G_{M\,SC} G_{M\,\text{КА}} m\upsilon_R} \text{АТС}_A =$$

$$= \frac{8\pi^2 \left(K_{CC}+1\right) p_N m\left[\left(1+SNIR_R\right)^{1/m}-1\right]\left(R_m^2+H_S^2\right)}{\lambda^2 G_{M\,SC} G_{M\,\text{КА}} \log_2\left(1+SNIR_R\right)} \text{АТС}_A . \quad (19)$$

На рис. 3 приведены расчетные зависимости ЭМФ $Z_{\text{АТС}}$(АТС$_A$) (19) для λ = 2 см (15 ГГц), K_{CC} = 2, m = 2, $\theta_{\min}$ = 0° и реальных диапазонов изменения $G_{M\,SC}$, $G_{M\,\text{АТ}}$, υ_R, $SNIR_R$. Использована двойная оцифровка шкал: оцифровка черным шрифтом соответствует диапазону значений АТС$_A$, декларируемых в [8, 9] для систем 4G/5G/6G, а оцифровка цветным шрифтом – диапазону значений АТС$_A$, соответствующих данным рис. 2. Особенность этих зависимостей состоит в том, что средняя плотность мобильного трафика считается постоянной по всей земной поверхности вне зависимости от реального неравномерного «очагового» распределения потребителей информационных услуг по территории. Именно этим можно объяснить то, что при гипотетических значениях АТС, соответствующих декларациям 4G (10^5 бит/с/м^2), средняя интенсивность создаваемого ЭМФ ненамного ниже предельно допустимого уровня (ПДУ) 10 мкВт/см^2; при АТС = 10^7 бит/с/м^2, декларируемой для 5G, средняя интенсивность ЭМФ близка к ПДУ только при очень высокой направленности антенн КА и АТ ($G_T = G_R$ = 40 дБ), а при АТС = 10^9 бит/с/м^2, декларируемой для 6G, расчетная средняя интенсивность создаваемого ЭМФ близка к уровням, соответствующим «тепловым» рекомендациям ICNIRP, и даже существенно превышающим данные уровни при $G_T = G_R$ < 40 дБ. Это означает, что средние уровни АТС, декларируемые для систем 4G/5G/6G, как и расчетные средние уровни ЭМФ, иллюстрируемые графиками на рис. 3, являются пиковыми, характерными для наземных участков малой площади (хотспотов), оборудованных терминалами ССС, на которые «фокусируются» потоки данных с КА, и где радиочастотный ЭМФ создается излучениями множества точек доступа (пико-БС) хотспотов. Расчетные же уровни $Z_{\text{АТС}}$ на рис. 3, соответствующие диапазону реальных значений средней АТС$_A$ [10^{-4}–10^2 бит/с/м^2] (цветной оцифровке шкал по горизонтали и вертикали), на 2–3 порядка превышают уровни [10] естественного ЭМФ в диапазоне 10–20 ГГц и близки к ранее выполненным оценкам интенсивности создаваемого ЭМФ с использованием данных о полной излучаемой мощности $P_{TR\,SC}$ КА и количестве КА в НГСС, что может служить подтверждением адекватности модели (19).

При отсутствии в системе регулировки $P_{M\,SC}$, $P_{TR\,SC}$ фиксирована и обеспечивает постоянство излучаемой энергии на бит передаваемой информации в радиоканале КА–АТ: E_{bt} = const. Случайными величинами являются значения ППМ, создаваемой КА у земной поверхности в точке размещения обслуживаемого АТ, мощность полезного сигнала P_0 и $SNIR$ на входе приемника АТ и пропускная способность радиоканала V_P. При радиоприеме АТ энергия E_{br} одного бита информации зависит от случайной величины R с п.р.в. П.р.в. $w(E_{br})$ и матожидание $\langle E_{br} \rangle$ могут быть определены с использованием (10) по методикам [1, 2]:

$$w\left(E_{br}\right) = \frac{E_{br\min} E_{br\max}}{\left(E_{br\max}-E_{rb\min}\right)E_{br}^2}; \left\langle E_{br} \right\rangle = \frac{E_{br\min} E_{br\max}}{E_{br\max}-E_{br\min}} \ln \frac{E_{br\max}}{E_{br\min}}; \; E_{br\min} \le E_{br} \le E_{br\max};$$

$$E_{br\min} = \frac{E_{bt} G_T G_R \lambda^2}{16\pi^2 R_m^2}; \; E_{br\max} = \frac{E_{bt} G_T G_R \lambda^2}{16\pi^2 H_S^2}, \quad (20)$$

где R_m – определяется по (13) при $\theta = \theta_{\min}$.

Используя (1), (2) и (4), значения $E_{br\min}$, $E_{br\max}$ можно связать с минимальными и максимальными значениями отношений $SNIR_{P\min}$, $SNIR_{R\min}$, $SNIR_{P\max}$, $SNIR_{R\max}$ и с соответствующими значениями спектральной эффективности $(\upsilon_P)_{\min}$, $(\upsilon_R)_{\min}$, $(\upsilon_P)_{\max}$, $(\upsilon_R)_{\max}$:

$$\left.\begin{aligned}
E_{br\,\min} &= \frac{(K_{CC}+1)\,p_N SNIR_{P\min}}{\log_2(1+SNIR_{P\min})} = \frac{(K_{CC}+1)\,p_N m\left(\left(1+SNIR_{R\min}\right)^{1/m}-1\right)}{\log_2(1+SNIR_{R\min})} = \\
&= \frac{(K_{CC}+1)\,p_N\left(2^{(\upsilon_P)_{\min}}-1\right)}{(\upsilon_P)_{\min}} = \frac{(K_{CC}+1)\,p_N\left(2^{m(\upsilon_R)_{\min}}-1\right)}{m\cdot(\upsilon_R)_{\min}}, \quad (\upsilon_P)_{\min} = m\cdot(\upsilon_R)_{\min}; \\
E_{br\,\max} &= \frac{(K_{CC}+1)\,p_N SNIR_{P\max}}{\log_2(1+SNIR_{P\max})} = \frac{(K_{CC}+1)\,p_N m\left(\left(1+SNIR_{R\max}\right)^{1/m}-1\right)}{\log_2(1+SNIR_{R\max})} = \\
&= \frac{(K_{CC}+1)\,p_N\left(2^{(\upsilon_P)_{\max}}-1\right)}{(\upsilon_P)_{\max}} = \frac{(K_{CC}+1)\,p_N\left(2^{m(\upsilon_R)_{\max}}-1\right)}{m\cdot(\upsilon_R)_{\max}}, \quad (\upsilon_P)_{\max} = m\cdot(\upsilon_R)_{\max},
\end{aligned}\right\} \quad (21)$$

где

$$(\upsilon_P)_{\min} \approx \log_2(1+SNIR_{P\min}),\ (\upsilon_P)_{\max} \approx \log_2(1+SNIR_{P\max}),\ SNIR_{P\max}/SNIR_{P\min} = R_m^2/H_S^2. \quad (22)$$

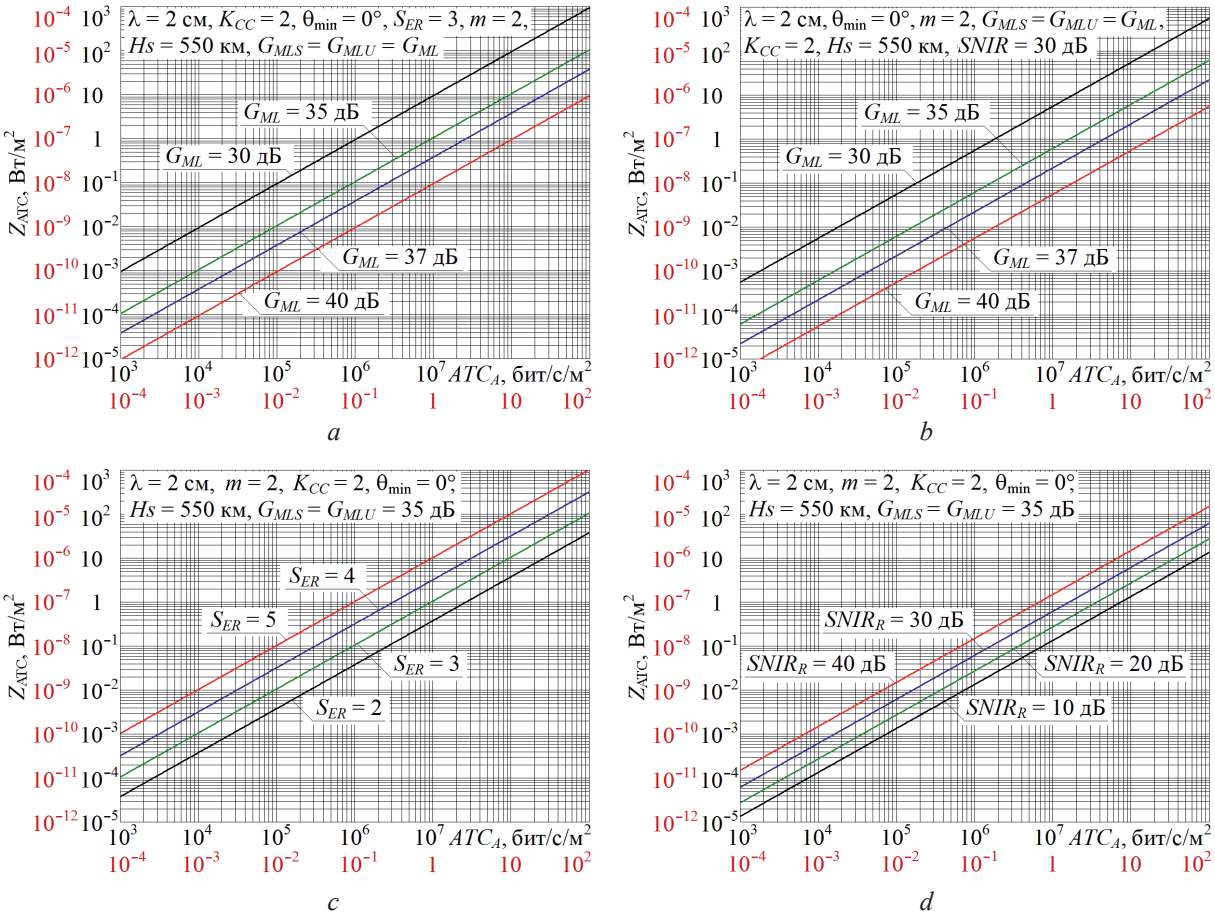


Рис. 3. Расчетные зависимости Z_{ATC}(ATC) для принятых ранее типовых характеристик радиоканала: *a* – при S_{ER} = 3; *b* – при $SNIR_R$ = 30 дБ; *c* – при различной спектральной эффективности радиоканала; *d* – при различном отношении $SNIR_R$ в радиоканале

Fig. 3. Calculated Z_{ATC}(ATC) dependencies for the previously adopted radio channel typical characteristics: *a* – for S_{ER} = 3; *b* – for $SNIR_R$ = 30 дБ; *c* – for different radio channel spectral efficiency; *d* – for different $SNIR_R$ ratios in radio channel

Используя (20) и (4), можно оценить значение ЭМФ у земной поверхности, создаваемого как в пятне ГЛ отдельного КА, так и всем НГСС, а также изменения в $SNIR_R$, υ_R и Z_{ATC} при изменении расстояния между АТ и КА из-за его перемещения по орбите. Однако очевидно, что в связи с относительно малым динамическим диапазоном изменений расстояния R между КА и ТН на земной поверхности эти оценки будут мало отличаться от оценок, приведенных на рис. 3.

Средняя мощность принимаемого сигнала $<P_0>$ равна произведению средней энергии (20) одного бита этого сигнала $<E_{br}>$ на скорость V_P приема информации: $<P_0> = V_P<E_{br}>$. Среднее значение ППМ $<Z_{UEA}>$, создаваемое антеннами КА в пятне ГЛ, которое обеспечивает прием информации АТ с антенной эффективной площадью $S_{e\,\text{АТ}}$, будет равно $<Z_{UEA}> = = <P_0>/S_{e\,\text{АТ}}$. При этом мощность $P_{r\,SC}$, излучаемая антенной КА в ГЛ, может быть определена через площадь S_{ML} поверхности, стягивающей телесный угол ГЛ (рис. 1, *а*), как $P_{r\,SC} = <Z_{UEA}>S_{ML}$.

Заключение

1. Представлены основные положения методики, позволяющей анализировать характеристики электромагнитной обстановки у земной поверхности на основе оценок средней территориальной плотности нисходящего трафика, создаваемого мегасозвездиями низкоорбитальных спутников связи (НСС), и их доступных системных характеристик. К исходным данным относятся высота орбит космических аппаратов, их количество в группировках НСС, диаграммы направленности антенн космических аппаратов и абонентских терминалов и ограничения на угол места, при котором допускается обслуживание абонентских терминалов, а также особенности сценариев их обслуживания.

2. Расчетные средние уровни электромагнитного фона, соответствующие диапазону реальных значений средней плотности беспроводного трафика, создаваемого созвездиями НСС на земной поверхности, на 2–3 порядка превышают уровни естественного электромагнитного фона [10] в полосе частот 10–20 ГГц. Это соответствует оценкам в [1, 2], которые получены с использованием данных о полной излучаемой мощности космических аппаратов и их количестве в группировках НСС, что может служить подтверждением адекватности разработанной методики оценки электромагнитного фона.

3. Прием/передача больших объемов данных по радиоканалам систем спутниковой связи для индивидуальных потребителей вместо использования ими наземной инфраструктуры сетей мобильной связи 4G/5G могут обеспечить значительное локальное снижение интенсивности радиочастотного электромагнитного фона у поверхности Земли до уровней, на несколько порядков ниже установленных ПДУ для населения. Поэтому интеграция спутниковых систем связи с использованием группировок НСС и наземных систем 4G/5G/6G может стать перспективным путем обеспечения электромагнитной экологичности и безопасности этих систем для населения. В частности, интеграция систем спутниковой связи с наземными системами может быть основана на создании хот-спотов и локальных кластеров, оборудованных одним абонентским терминалом систем спутниковой связи, в которых реализован «проводной», волоконно-оптический или низкоэнергетический беспроводной вариант организации доступа пользователей к каналам приема/передачи информации систем спутниковой связи и мобильной связи 4G/5G/6G.

4. Исследованные расчетные уровни электромагнитного фона, создаваемого у земной поверхности излучениями множества НСС, являются только одним из слагаемых интенсивности этого фона, создаваемых излучениями систем спутниковой связи, использующими НСС. Следует ожидать, что в местах с высокой плотностью населения вклад излучений множества абонентских терминалов этих систем и наземных шлюзовых станций в интенсивность наземного антропогенного электромагнитного фона в диапазоне СВЧ по меньшей мере сопоставим со вкладом излучений космического сегмента этих систем спутниковой связи. Данный вопрос требует отдельного изучения.

Список литературы / References

1. Mordachev V., Tsyanenka D., Svistunou A. (2024) Characteristics of Electromagnetic Environment Created by Communication Low Earth Orbit Satellite Systems Near the Earth's Surface. *Proc. of the Int. Symp. "EMC Europe 2024", Bruges, Belgium, Sept. 2–5.* 1178–1183.
2. Mordachev V. I., Svistunou A. S., Tsyanenka D. A. (2024) Prediction the Level of the Electromagnetic Background Created by Constellations of Satellites Near the Earth's Surface Using Registration Data. *Doklady BGUIR*. 22 (5), 43–52. http://dx.doi.org/10.35596/1729-7648-2024-22-5-43-52 (in Russian).
3. Mordachev V. I. (2019) Estimation of Intensity of Electromagnetic Background, Created by Wireless Systems of Public Information Services, on the Base of Forecast of Traffic Terrestrial Density. *Doklady BGUIR*. (2), 39–49. https://doklady.bsuir.by/jour/article/view/1066 (in Russian).
4. Sibille A., Oestges C., Zanella A. (2011) MIMO from Theory to Implementation. *Academic Press*.

5. Tikhvinskiy V., Terentiev S., Visochin V. (2014) *LTE/LTE Advanced Mobile Networks: 4G Technologies, Applications Architecture*. Moscow, Media Publ. (in Russian).
6. The Concept of Transmission Loss for Radio Links. *Rec. ITU-R 341-7.* 2019.
7. Calculation of Free-Space Attenuation. *Rec. ITU-R 525-4.* 2019.
8. IMT Vision – Framework and Overall Objectives of the Future Development of IMT for 2020 and Beyond. *Rec. ITU-R. M.2083.* 2015.
9. Zhang Z., et al. (2019) 6G Wireless Networks: Vision, Requirements, Architecture, and Key Technologies. *IEEE VT Magazine.* 14 (3), 28–41.
10. Radio Noise. *Rec. ITU-R P.372–16*. 2022.

Поступила 10.11.2024
Received: 10 November 2024

Принята в печать 06.01.2025
Accepted: 15 January 2025

Вклад авторов / Authors' contribution

Авторы внесли равный вклад в написание статьи / The authors contributed equally to the writing of the article.

Сведения об авторах

Мордачев В. И., канд. техн. наук, доц., зав. науч.-исслед. лаб. «Электромагнитная совместимость радиоэлектронных средств» (НИЛ 1.7), Белорусский государственный университет информатики и радиоэлектроники

Ционенко Д. А., канд. физ.-мат. наук, доц., ст. науч. сотр. НИЛ 1.7, Белорусский государственный университет информатики и радиоэлектроники

Information about the authors

Mordachev V. I., Cand. Sci. (Tech.), Associate Professor, Head of the Research Laboratory "Electromagnetic Compatibility of Radio-Electronic Equipment" (Lab 1.7), Belarusian State University of Informatics and Radioelectronics

Tsyanenka D. A., Cand. Sci. (Phys. and Math.), Associate Professor, Senior Researcher at the Lab 1.7, Belarusian State University of Informatics and Radioelectronics

Адрес для корреспонденции

220013, Республика Беларусь,
Минск, ул. П. Бровки, 6
Белорусский государственный университет
информатики и радиоэлектроники
Тел.: +375 17 293-84-38
E-mail: mordachev@bsuir.by
Мордачев Владимир Иванович

Address for correspondence

220013, Republic of Belarus,
Minsk, P. Brovki St., 6
Belarusian State University
of Informatics and Radioelectronics
Tel.: +375 17 293-84-38
E-mail: mordachev@bsuir.by,
Mordachev Vladimir Ivanovich

CC BY

http://dx.doi.org/10.35596/1729-7648-2025-23-3-46-53

УДК 004.89

ЭКСПРЕСС-АНАЛИЗ СТРУКТУРНЫХ И ЭЛЕКТРОННЫХ СВОЙСТВ НАНОМАТЕРИАЛОВ МЕТОДАМИ БОЛЬШИХ ДАННЫХ, БОЛЬШИХ ЯЗЫКОВЫХ МОДЕЛЕЙ И ГЕНЕРАТИВНОГО ИСКУССТВЕННОГО ИНТЕЛЛЕКТА

Н. А. ШИМАНСКИЙ[1,2], А. В. БАГЛОВ[2,3], Х. Б. МУСАЕВ[4], О. Н. РУЗИМУРАДОВ[5], Л. С. ХОРОШКО[2,3]

[1]Компания «АндерсенБел» (Минск, Республика Беларусь)
[2]Белорусский государственный университет (Минск, Республика Беларусь)
[3]Белорусский государственный университет информатики и радиоэлектроники (Минск, Республика Беларусь)
[4]Институт фундаментальных и прикладных исследований при Национальном исследовательском университете «ТИИИМСХ» (Ташкент, Республика Узбекистан)
[5]Туринский политехнический университет в городе Ташкенте (Ташкент, Республика Узбекистан)

Аннотация. Рассмотрена возможность использования больших языковых моделей, генеративного машинного обучения и методов работы с большими данными для предиктивного анализа электронных свойств наноструктур на основе полупроводниковых материалов, не ограничивая при этом общность данного подхода для иных кристаллических материалов. Описано концептуальное решение в виде кроссплатформенного программного приложения, имеющего функцию генерации расширенного поиска, для применения указанного подхода в работе с массивами научных данных. Результаты предварительного тестирования показывают положительный результат использования разработанного решения для предсказания полупроводниковых свойств (ширина запрещенной зоны, энергия Ферми) эталонного материала. Обсуждаются перспективы развития и внедрения методов больших данных, больших языковых моделей и генеративного искусственного интеллекта в рамках тенденций современного материаловедения.

Ключевые слова: машинное обучение, нейронные сети, большие языковые модели, большие данные, наноматериалы, предсказательный анализ.

Конфликт интересов. Авторы заявляют об отсутствии конфликта интересов.

Для цитирования. Экспресс-анализ структурных и электронных свойств наноматериалов методами больших данных, больших языковых моделей и генеративного искусственного интеллекта / Н. А. Шиманский [и др.] // Доклады БГУИР. 2025. Т. 23, № 3. С. 46–53. http://dx.doi.org/10.35596/1729-7648-2025-23-3-46-53.

EXPRESS ANALYSIS OF THE STRUCTURAL AND ELECTRONIC PROPERTIES OF NANOMATERIALS USING BIG DATA, LARGE LANGUAGE MODELS AND GENERATIVE ARTIFICIAL INTELLIGENCE

NIKITA A. SHYMANSKI[1,2], ALEKSEY V. BAGLOV[2,3], KHUSNIDDIN B. MUSAEV[4], OLIM N. RUZIMURADOV[5], LIUDMILA S. KHOROSHKO[2,3]

[1]AndersenBel company HTP (Minsk, Republic of Belarus)
[2]Belarusian State University (Minsk, Republic of Belarus)
[3]Belarusian State University of Informatics and Radioelectronics (Minsk, Republic of Belarus)
[4]Institute of Basic and Applied Research at the National Research University Tashkent Institute of Engineers of Irrigation and Agricultural Mechanization (Tashkent, Republic of Uzbekistan)
[5]Turin Polytechnic University in Tashkent (Tashkent, Republic of Uzbekistan)

Abstract. The possibility of using large language models, generative machine learning and big data methods for predictive analysis of the electronic properties of nanostructures based on semiconductor materials is considered, without limiting the generality of this approach for other crystalline materials. The conceptual solution in the form of a cross-platform software application with an advanced search generation capability for applying this approach to working with scientific data arrays is described. Preliminary test results showed a positive result of using the developed solution to predict semiconductor properties (bandgap width, Fermi energy) of a reference material. Prospects for the development and implementation of big data methods, large language models, and generative artificial intelligence in the context of modern trends in materials science are discussed.

Keywords: machine learning, neural networks, large language models, big data, nanomaterials, predictive analysis.

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

For citation. Shymanski N. A., Baglov A. V., Musaev Kh. B., Ruzimuradov O. N., Khoroshko L. S. (2025) Express Analysis of the Structural and Electronic Properties of Nanomaterials Using Big Data, Large Language Models and Generative Artificial Intelligence. *Doklady BGUIR*. 23 (3), 46–53. http://dx.doi.org/10.35596/1729-7648-2025-23-3-46-53 (in Russian).

Введение

Современное материаловедение – междисциплинарная область, требующая от исследователей знаний в теоретической и прикладной физике и химии, биологии и инженерии. В настоящее время тенденция в материаловедении характеризуется дуализмом между потребностью в глобальном сотрудничестве и необходимостью защиты уникальных разработок и интеллектуальной собственности, что, в свою очередь, обеспечивает сохранение передовых позиций государства на международной арене. Глобализация, характерная для процессов научных исследований в последние десятилетия, способствовала укреплению позиций мощнейших мировых исследовательских коллабораций и научных центров (например, Объединенный институт ядерных исследований в г. Дубне, РФ, и др.), в которых совместное получение новых знаний и использование результатов являются основами существования проектов и устойчивого развития всей исследовательской инфраструктуры. Разрабатывая стратегии для балансирования этих конкурирующих интересов, исследователи, учреждения и правительства могут способствовать прогрессу в материаловедении, обеспечивая при этом защиту конфиденциальной информации и интеллектуальной собственности. Использование современных информационных технологий может стать дополнительным средством обеспечения эффективного взаимодействия между учеными и исследователями всех стран, в том числе для обмена актуальными результатами, сопоставления и верификации новых данных, облегчения поисковой исследовательской работы и повышения ее эффективности, обеспечивая при этом требуемую степень конфиденциальности.

Успешное развитие теоретического и прикладного разделов современного материаловедения сегодня практически невозможно без использования вычислительных средств и методов компьютерной автоматизации. Интересующие свойства исследуемых материалов могут быть не только проанализированы с минимальными временными затратами, но и предположены с высокой степенью достоверности с применением современных программных пакетов, которые позволяют, например, предсказывать структуры низкоразмерных материалов: наночастиц, полимеров, по-

верхностей, межзеренных границ и 2D-кристаллов [1–4]. Современные нейросетевые инструменты и программные интерфейсы для них (машинное обучение – Machine Learning, ML, дискриминационный искусственный интеллект – Discriminative Artificial Intelligence, AI, генеративный искусственный интеллект – Generative AI), большие данные и продвинутые аналитические системы (Big Data и Advanced Analytics) могут быть привлечены для оптимизации подобных задач, сокращения времени обработки больших массивов данных, автоматизации анализа полученных результатов и т. д. Также возможно получение сгенерированного ответа в структурированном виде на запрос пользователя, сформулированный естественным языком. В области материаловедения такие подходы предоставляют исследователям возможности генерации новых экспериментальных моделей наноструктур с предиктивным анализом их электронных свойств фактически в режиме реального времени, в отличие от привычных «ручных» режимов моделирования.

В статье рассмотрен пример решения специальной исследовательской задачи – моделирование и анализ структуры и свойств выбранного полупроводникового материала. Данный подход может быть использован также для изучения свойств объемных и сверхтонких наноструктур из практически любых материалов, характеризующихся наличием упорядоченной структуры.

Анализ и постановка задачи

Эффективность вычислительных экспериментов для исследования и определения структур и электронных свойств наноматериалов в значительной мере обусловлена качеством программной реализации и используемыми вычислительными моделями компьютерного моделирования. Хорошо зарекомендовали себя такие общедоступные программные комплексы, как VESTA (Visualisation for Electronic and Structural Analysis) и OpenMX (Open source package for Material eXplorer) [5–7]. Первый используется для 3D-моделирования пространственных структур и визуализации их электронных свойств, второй – для определения электронных свойств (зонная структура, плотность электронных состояний и т. д.). на основе теорий функционала плотности, нормосохраняющих псевдопотенциалов и псевдоатомных локализованных базисных функций. Электронные свойства исследуемых материалов, полученные в результате моделирования в программных пакетах типа OpenMX, позволяют проанализировать перспективы и возможности их использования в полупроводниковой электронике и фотонике, что особенно актуально в случае затрудненной или невозможной реализации физического эксперимента, а также при поиске новых перспективных материалов и их комбинаций. Практическая трудность реализации данного подхода обусловлена широким разнообразием параметров полупроводниковых структур, характеризующих их морфологические свойства (такие как симметрия кристаллической ячейки, взаимное расположение кристаллографических плоскостей и пространственная ориентация интерфейса в случае двумерных материалов, деформация и взаимодействие слоев для слоистых структур и др.). На этапе формирования входного файла, задающего параметры вычислений, необходимо произвести корректное описание всех этих параметров, что осуществляется преимущественно в ручном режиме и требует значительных временных затрат. При этом процесс моделирования в различных программных пакетах в совокупности с подготовкой исходных данных в зависимости от величины и сложности исследуемой системы может занимать достаточно продолжительное время, а прогнозирование успешности результата является ответственностью самого исследователя и не всегда соответствует ожиданиям.

Одним из вариантов решения данных вопросов может стать комплексный подход, сочетающий в себе возможности оптимизации и ускорения процессов моделирования и анализа наноструктур с применением инструментов Big Data & Machine Learning. Данный подход использовался в [8, 9] для оценки и анализа результатов исследований свойств и структуры кристаллических материалов методом дифракции рентгеновских лучей. Для предварительного прогнозирования свойств наноструктур было применено предиктивное машинное обучение, которое включало в себя создание и обучение нейронной сети с помощью эталонных образцов дифрактограмм – результатов, полученных для известных нанокристаллических материалов с описанными свойствами. Для исследования структурных свойств кристаллических материалов обученная нейросеть может предложить вероятностное определение их характеристик с некоторой степенью точности, зависящей от количества циклов обучения и детальности обучающего материала. Однако реализация аналогичного подхода для прогнозирования электронных свойств будет более сложной

и трудоемкой ввиду наличия сложных зависимостей данных свойств от структурной конфигурации, особенно при понижении размерности. Для двумерных материалов и сверхтонких (до 10 монослоев) наноструктур реализацию такого подхода усложняет большое количество возможных форм-факторов (включая нанотрубки, наношнуры и т. п.), обладающих уникальными электронными свойствами.

Учет всех перечисленных особенностей с использованием универсального алгоритма является весьма трудоемкой и не всегда оправданной по затратам задачей, а вероятностный (предиктивный) прогноз свойств с применением нейросети в результате может быть снижен вплоть до недостоверного результата [6]. Таким образом, требуется концептуальное изменение подхода к описанию электронных свойств наноматериалов с применением современных методов автоматизации и анализа данных.

Описание решения

Принципиально новый подход к применению нейронных сетей и машинного обучения для решения задач в физике наноструктур и наноматериаловедении предполагает использование больших языковых моделей (Large Language Models, LLM) [4]. Для интеграции LLM в такого рода узкоспециальные области предлагается использовать специальную методологию «обогащения» контекста для расширения области знаний нейронной сети, называемую генерацией дополненного поиска (Retrieval-Augmented Generation, RAG). Данная методология представляет собой оптимизацию выходных данных LLM, в результате которой она ссылается на дополненную (так называемую обогащенную) базу знаний, таким образом выходя за пределы своих источников обучающих данных перед генерацией ответа (рис. 1). Без необходимости переобучения модель RAG дополнительно расширяет возможности LLM, используя определенные домены или внутреннюю базу данных организации. В настоящее время примерами языковых моделей с поддержкой RAG являются популярные у широкого круга пользователей Meta Llama 3, Anthropic Claude, Amazon Titan и др.

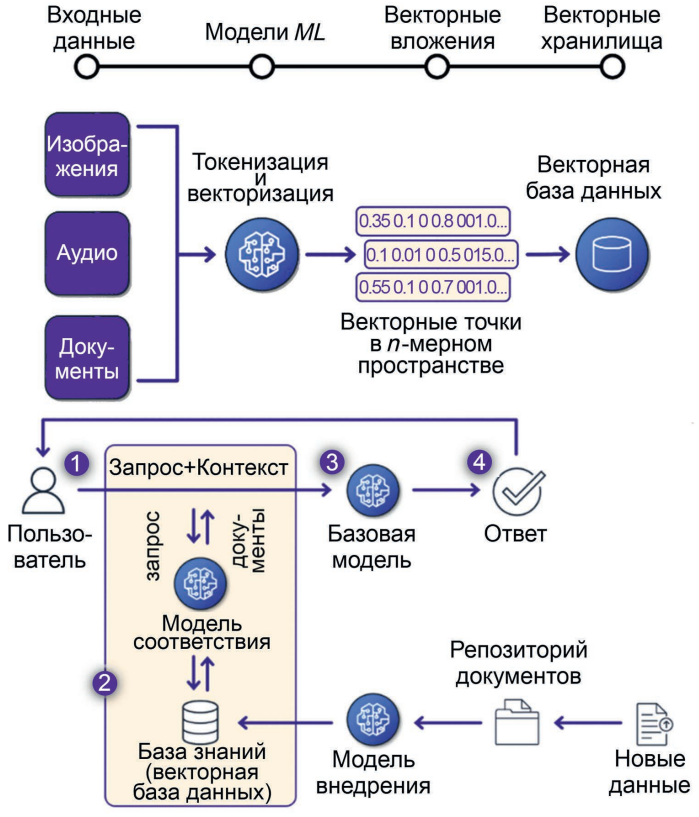


Рис. 1. Схема процессов привлечения векторных баз данных
для обогащения ответов языковой модели при пользовательском запросе в LLM
Fig. 1. Scheme for incorporating vector databases to enrich language model responses to user queries in LLM

Описанное RAG-обогащение в рамках решаемой задачи происходит в несколько этапов. На начальном этапе релевантная информация накапливается в текстовом и бинарном формате, имеющем произвольную организацию и структуру, но содержащем определенные количественные описания свойств наноструктур и их качественную интерпретацию. В качестве источников информации могут быть использованы текущие отчеты о научных исследованиях, рабочие записи и сводные таблицы, публикации в рецензируемых научных изданиях, теоретические исследования фундаментальных свойств материалов и т. д.; допустимы форматы файлов .txt, .doc(x), .xls(x), .pdf и др. RAG позволяет при обучении нейросети работать также и с графическими материалами: графиками зависимостей, изображениями структур, результатами визуализации расчетных и экспериментальных данных и другими в виде сканируемых и распознаваемых изображений. Отсутствие строгих требований к оглавлению и структурированию содержания материала, а также ограничений по размерам обрабатываемых файлов является важным преимуществом использования предлагаемой методики. Накопленная информация (контент) трансформируется в бинарный вид и интегрируется с языковой моделью посредством подключения специализированной векторной базы данных (например, LanceDB, AWS OpenSearch). Обобщенно архитектурная диаграмма описанной методики в виде программного интерфейса показана на рис. 2.

Рис. 2. Архитектурная диаграмма в виде программного интерфейса с пользовательским вводом
Fig. 2. Architectural diagram in the form of a software interface with user input

При получении пользовательского запроса генерация ответа пользователю происходит с учетом специализированного контекста из векторной базы данных после сверки с RAG-контекстом, при этом релевантный поиск по большим объемам векторных данных (десятки и сотни терабайт) осуществляется языковой моделью за несколько секунд, затем пользователь получает ответ, состоящий из наиболее вероятных словосочетаний и предложений. В отличие от работы с традиционными нейронными сетями, реализация описанного подхода сокращает вычислительные и временные затраты на использование искусственного интеллекта, поскольку не требует дополнительных тренировок и переобучений языковой модели.

Реализация описываемого подхода предполагается в виде кроссплатформенного программного решения, которое находится в стадии разработки и накопления векторных баз данных. Ряд предварительных экспериментов в виде диалогов (prompting) с искусственным интеллектом подтверждает перспективность развития данной концепции [10]. Используемый для взаимодействия RAG чат-бот в ходе ряда предварительных тестов смог определить характеристики зонной структуры (ширину запрещенной зоны, энергию Ферми) для интересующей наноструктуры – широкозонного сегнетоэлектрика титаната бария ($BaTiO_3$). Чат-бот определил эти значения, используя обогащенный контекст и созданную заранее векторную базу данных, и сгенерировал корректный ответ на пользовательский запрос. Тестируемая языковая модель смогла провести статистическое сопоставление и сгенерировать верный ответ, не обладая запрашиваемыми параметрами в явном виде. Однако следует отметить, что на данном этапе языковая модель не смогла

корректно обработать запросы прогнозного характера по инжинирингу запрещенной зоны путем легирования, модификации кристаллической решетки или привнесения дефектов: чат-бот вместо этого предлагал в ответах соединения другого химического состава и структуры, но с искомыми параметрами запрещенной зоны. Обнаруженная особенность лишний раз подтверждает важность наличия открытых результатов исследований для глобализации научной сети и полноценного внедрения современных технологий в наукоемкие сферы по всему миру, что в рамках контекста описанных исследований подразумевает создание и своевременную актуализацию материаловедческих баз данных достаточной полноты.

Заключение

1. Как продемонстрировано в статье, применение нейронных сетей (Generative AI, LLM) в качестве больших языковых моделей для исследования и анализа свойств наноструктур является перспективным, особенно для повышения уровня автоматизации ряда исследовательских задач при значительном сокращении временных и трудовых затраты на обработку. Предложенный подход предусматривает постоянное развитие и обучение LLM-модели в рамках специализированного научного контекста (в рассматриваемом случае – в области наноматериаловедения), что будет содействовать по мере накопления заданного контекста увеличению точности прогнозного анализа и способности искусственного интеллект не только корректно производить поиск по имеющимся данным, но и предлагать новые наноструктуры с требуемыми и предсказанными свойствами.

2. Использование генеративного Generative AI и LLM в наноматериаловедении может ускорить нахождение и проектирование новых материалов с требуемыми и уникальными электронными и структурными свойствами. Применяя виртуальные платформы для совместной работы, шифрование данных и контроль доступа, модели и программное обеспечение с открытым исходным кодом, а также соглашения о совместных исследованиях, исследователи могут сбалансировать потребность в глобальном сотрудничестве и необходимость защиты уникальных разработок и интеллектуальной собственности.

3. Исследования поддержаны в рамках заданий № 2.14.3 (№ ГР 20212445) и № 2.25 (№ ГР 20240603) подпрограммы «Наноструктура» ГПНИ «Материаловедение, новые материалы и технологии» и проекта Т23УЗБ-111 Белорусского республиканского фонда фундаментальных исследований (№ ГР 20240142).

Список литературы

1. Niu, H. N. Simple and Accurate Model of Fracture Toughness of Solids / H. N. Niu, S. Niu, A. R. Oganov // J. Appl. Phys. 2019. Vol. 125. https://doi.org/10.1063/1.5066311.
2. First-Principles Prediction of Ferroelectric Janus Si_2XY (X/Y = S/Se/Te, $X \neq Y$) Monolayers with Negative Poisson's Ratios / Y. Zhu [et al.] // Phys. Chem. Chem. Phys. 2024. Vol. 26. P. 4555–4563. https://doi.org/10.1039/D3CP05107B.
3. Graph Theory and Graph Neural Network Assisted High-Throughput Crystal Structure Prediction and Screening for Energy Conversion and Storage / J. Ojih [et al.] // J. Mater. Chem. A. 2024. Vol. 12. P. 8502–8515. https://doi.org/10.1039/D3TA06190F.
4. Price, S. L. Computed Crystal Energy Landscapes for Understanding and Predicting Organic Crystal Structures and Polymorphism / S. L. Price // Acc. Chem. Res. 2009. Vol. 42, Iss. 1. P. 117–126. https://doi.org/10.1021/ar800147t.
5. Ozaki, T. Variationally Optimized Atomic Orbitals for Large-Scale Electronic Structures / T. Ozaki // Phys. Rev. B. 2003. Vol. 67. https://doi.org/10.1103/physrevb.67.155108.
6. Ozaki, T. Numerical Atomic Basis Orbitals from H to Kr / T. Ozaki, H. Kino // Phys. Rev. B: Condens. Matter Mater. Phys. 2004. Vol. 69. https://doi.org/10.1103/PhysRevB.69.195113.
7. Ozaki, T. Efficient Projector Expansion for the Ab Initio LCAO Method / T. Ozaki, H. Kino // Phys. Rev. B. 2005. Vol. 72. https://doi.org/10.1103/PhysRevB.72.045121.
8. Шиманский, Н. А. Автоматизация обработки результатов исследования структуры и свойств наноматериалов / Н. А. Шиманский, А. В. Баглов, Л. С. Хорошко // BIG DATA и анализ высокого уровня: сб. науч. ст. IX Междунар. науч.-практ. конф., Минск, 17–18 мая 2023 г.: в 2 ч. Ч. 1. Минск: Белор. гос. ун-т информ. и радиоэлек., 2023. С. 296–300.

9. Шиманский, Н. А. Автоматизация обработки результатов исследования структуры материалов / Н. А. Шиманский, А. В. Баглов, Л. С. Хорошко // Information Tehnologies and Systems 2023 (ITS 2023): матер. Междунар. науч. конф., Минск, 22 нояб. 2023 г. Минск: Белор. гос. ун-т информ. и радиоэлек., 2023. С. 207–208.
10. Шиманский, Н. А. Экспресс-анализ структурных и электронных свойств наноматериалов методами Big Data, Large Language Models & Generative AI / Н. А. Шиманский, А. В. Баглов, Л. С. Хорошко // Компьютерное проектирование в электронике: сб. тр. Междунар. науч.-практ. конф, Минск, 28 нояб. 2024 г. Минск: Белор. гос. ун-т информ. и радиоэлек., 2024. С. 100–103.

Поступила 04.02.2025 Принята в печать 05.03.2025

References

1. Niu H. N., Niu S., Oganov A. R. (2019) Simple and Accurate Model of Fracture Toughness of Solids. *J. Appl. Phys*. 125. https://doi.org/10.1063/1.5066311.
2. Zhu Y., Qu Z., Zhang J., Wang X., Jiang S., Xu Z., et al. (2024) Dai First-Principles Prediction of Ferroelectric Janus Si_2XY (X/Y = S/Se/Te, $X \neq Y$) Monolayers with Negative Poisson's Ratios. *Phys. Chem. Chem. Phys*. 26, 4555–4563. https://doi.org/10.1039/D3CP05107B.
3. Ojih J., Al-Fahdi M., Yao Y., Hu J., Ming Hu (2024) Graph Theory and Graph Neural Network Assisted High-Throughput Crystal Structure Prediction and Screening for Energy Conversion and Storage. *J. Mater. Chem. A.* 12, 8502–8515. https://doi.org/10.1039/D3TA06190F.
4. Price S. L. (2009) Computed Crystal Energy Landscapes for Understanding and Predicting Organic Crystal Structures and Polymorphism. *Acc. Chem. Res*. 42 (1), 117–126. https://doi.org/10.1021/ar800147t.
5. Ozaki T. (2003) Variationally Optimized Atomic Orbitals for Large-Scale Electronic Structures. *Phys. Rev. B.* 67. https://doi.org/10.1103/physrevb.67.155108.
6. Ozaki T., Kino H. (2004) Numerical Atomic Basis Orbitals from H to Kr. *Phys. Rev. B: Condens. Matter Mater. Phys.* 69. https://doi.org/10.1103/PhysRevB.69.195113.
7. Ozaki T., Kino H. (2005) Efficient Projector Expansion for the Ab Initio LCAO Method. *Phys. Rev. B.* 72. https://doi.org/10.1103/PhysRevB.72.045121.
8. Shymanski N. A., Baglov A. V., Khoroshko L. S. (2023) Automation of Processing the Results of Nanomaterials Structure and Properties Studying. *BIG DATA and Advanced Analytics: Proceed. of IX Int. Sci. & Pract. Conf., Minsk, May 17–18, in 2 p. P. 1*. Minsk, Belarusian State University of Informatics and Radioelectronics. 296–300 (in Russian).
9. Shymanski N. A., Baglov A. V., Khoroshko L. S. (2023) Automate the Processing of the Results of Studying the Material Structure. *Information Tehnologies and Systems 2023 (ITS 2023): Proceed. of Int. Sci. Conf., Minsk, Nov. 22*. Minsk, Belarusian State University of Informatics and Radioelectronics. 207–208 (in Russian).
10. Shymanski N. A., Baglov A. V., Khoroshko L. S. (2024) Express Analysis of the Structural and Electronic Properties of Nanomaterials Using Big Data, Large Language Models & Generative AI. *Electronic Design Automation: Proceed. of Int. Sci.&Pract. Conf., Minsk, Nov. 28*. Minsk, Belarusian State University of Informatics and Radioelectronics. 100–103 (in Russian).

Received: 4 February 2025 Accepted: 5 March 2025

Вклад авторов

Шиманский Н. А. осуществил концептуализацию идеи, разработал схемы процессов, участвовал в разработке программного кода, тестировании решения, подготовке рукописи статьи.

Баглов А. В. осуществил валидацию результатов, подготовку массивов входных данных, участвовал в разработке программного кода.

Мусаев Х. Б. и Рузимурадов О. Н. предоставили наборы экспериментальных данных для формирования массивов входных данных, участвовали в обсуждении результатов.

Хорошко Л. С. провела подготовку массивов входных данных, валидацию результатов, предварительное тестирование решения, подготовку рукописи, отвечала за общую организацию работы над проектом.

Authors' contribution

Shymanski N. A. conceptualized the idea, developed process diagrams, participated in the development of the program code, testing the solution, and preparation of the manuscript.

Baglov A. V. validated the results, prepared the input data arrays, and participated in the development of the program code.

Musaev Kh. B. and Ruzimuradov O. N. provided sets of experimental data for the formation of input data arrays and participated in the discussion of the results.

Khoroshko L. S. prepared the input data arrays, validated the results, pre-tested the solution, prepared the manuscript, and was responsible for the overall organization of the project.

Сведения об авторах

Шиманский Н. А., системный архитектор компании «АндерсенБел»

Баглов А. В., магистр техн. наук, ст. науч. сотр. науч.-исслед. лаб. энергоэффективных материалов и технологий, Белорусский государственный университет (БГУ); науч. сотр. Центра наноэлектроники и новых материалов, Белорусский государственный университет информатики и радиоэлектроники (БГУИР)

Мусаев Х. Б., д-р философии по химии, докторант, Институт фундаментальных и прикладных исследований при Национальном исследовательском университете «ТИИИМСХ»

Рузимурадов О. Н., д-р хим. наук, проф. каф. естественно-математических наук, Туринский политехнический университет в городе Ташкенте

Хорошко Л. С., канд. физ.-мат. наук, доц., вед. науч. сотр. науч.-исслед. лаб. энергоэффективных материалов и технологий, БГУ; вед. науч. сотр. Центра наноэлектроники и новых материалов, БГУИР

Адрес для корреспонденции

220030, Республика Беларусь,
Минск, просп. Независимости, 4
Белорусский государственный университет
Тел.: +375 17 209-55-34
E-mail: l.s.khoroshko@gmail.com
Хорошко Людмила Сергеевна

Information about the authors

Shymanski N. A., Solutions Architect at AndersenBel Ltd

Baglov A. V., M. Sci (Eng.), Senior Researcher at the R&D Laboratory of Energy Efficient Materials and Technologies, Belarusian State University (BSU); Researcher at the Center of Nanoelectronics and Advanced Materials, Belarusian State University of Informatics and Radioelectronics (BSUIR)

Musaev Kh. B., Ph. D. (Chem.), Doctoral Student, Institute of Basic and Applied Research at the National Research University Tashkent Institute

Ruzimuradov O. N., Dr. Sci. (Chem.), Professor at the Department of Natural-Mathematical Science, Turin Polytechnic University in Tashkent

Khoroshko L. S., Cand. Sci. (Phys. and Math.), Associate Professor, Leading Researcher at the Laboratory of Energy Efficient Materials and Technologies, BSU; Leading Researcher at the Center of Nanoelectronics and Advanced Materials, BSUIR

Address for correspondence

220030, Republic of Belarus,
Minsk, Nezavisimosti Ave., 4
Belarusian State University
Tel.: +375 17 209-55-34
E-mail: l.s.khoroshko@gmail.com
Khoroshko Liudmila Sergeevna

CC BY

http://dx.doi.org/10.35596/1729-7648-2025-23-3-54-61

УДК 681.3

МЕТОД АВТОМАТИЧЕСКОЙ ГЕНЕРАЦИИ ВОПРОСОВ И ОТВЕТОВ ДЛЯ СИСТЕМ ТЕСТИРОВАНИЯ ЗНАНИЙ

С. А. МИГАЛЕВИЧ, Ю. О. ГЕРМАН, О. В. ГЕРМАН

Белорусский государственный университет информатики и радиоэлектроники (Минск, Республика Беларусь)

Аннотация. Рассмотрена техника автоматической генерации вопросов и ответов для тестирования знаний. Предложенный подход построен на разработанном методе определения и использования кластеров ключевых слов для генерации по ним вопросов и ответов на базе глобальной или локальной языковой модели, что определяет теоретико-прикладную новизну. Кластеры строятся на основе корпуса текстовых документов, относящихся к изучаемой области (учебная литература, методические пособия, электронные ресурсы). Представленное в статье техническое решение является логически законченным и может служить основой для практических разработок.

Ключевые слова: информационные технологии, обучение, тестирование знаний, кластерный анализ.

Конфликт интересов. Авторы заявляют об отсутствии конфликта интересов.

Для цитирования. Мигалевич, С. А. Метод автоматической генерации вопросов и ответов для систем тестирования знаний / С. А. Мигалевич, Ю. О. Герман, О. В. Герман // Доклады БГУИР. 2025. Т. 23, № 3. С. 54–61. http://dx.doi.org/10.35596/1729-7648-2025-23-3-54-61.

METHOD OF AUTOMATIC GENERATION OF QUESTIONS AND ANSWERS FOR KNOWLEDGE TESTING SYSTEMS

SERGEI A. MIGALEVICH, JULIA O. GERMAN. OLEG V. GERMAN

Belarusian State University of Informatics and Radioelectronics (Minsk, Republic of Belarus)

Abstract. The technique of automatic generation of questions and answers for knowledge testing is considered. The proposed approach is based on the developed method of defining and using clusters of keywords to generate questions and answers based on a global or local language model, which constitutes theoretical and applied novelty. Clusters are built on the basis of a corpus of text documents related to the studied area (textbooks, methodological manuals, electronic resources). The technical solution presented in the article is logically complete and can serve as a basis for practical developments.

Keywords: information technologies, education, knowledge testing, cluster analysis.

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

For citation. Migalevich S. A., German Ju. O., German O. V. (2025) Method of Automatic Generation of Questions and Answers for Knowledge Testing Systems // *Doklady BGUIR*. 23 (3), 54–61. http://dx.doi.org/10.35596/1729-7648-2025-23-3-54-61 (in Russian).

Введение

Известные подходы к решению задачи генерации вопросов в автоматизированной системе тестирования знаний обучаемых [1, 2] используют генеративные нейросети, интеллектуальные интерактивные текстовые аналитические системы типа глобальной языковой модели ChatGPT, системы генерации вопросов на основе подготовленных шаблонов (templates), динамический отбор вопросов из базы данных аналогичных систем, механизмы грамматического разбора и др.

Проблемным местом является генерация вопросов к конкретному текстовому материалу, например электронному ресурсу, учебнику, конспекту лекций, авторской монографии, книге и т. п. Вопросы генерируют либо к конкретным предложениям, параграфам, выдержкам, темам, либо ко всему документу.

В статье рассмотрена новая техника генерации вопросов в автоматизированной системе тестирования знаний обучаемых, а именно – генерация вопросов ко всему документу. Подход основан на выделении ключевых слов в документе и разбиении их на кластеры, причем допускается пересечение кластеров. Изложены общий подход и метод построения кластеров ключевых слов, используемых для формирования вопросов и получения ответов на них, приемлемые с точки зрения вычислительной сложности для практических приложений. Для каждого кластера можно построить несколько вопросов, используя локальную (кратко рассмотренную в статье) или глобальную языковую модель типа ChatGPT, и получить ответы на базе этой модели. Глобальные языковые модели представляют мощнейший семантический аппарат, но не свободный от недостатков. Так, на вопрос к ChatGPT: «Объясни способ построения матрицы покрытия для системы дизъюнктов задачи выполнимости по книге О. В. Германа. Введение в теорию экспертных систем и обработку знаний» [3] система выдала неполный (по существу, неверный) ответ. Проиллюстрируем это на следующем примере. Предположим, что имеются дизъюнкты:

$$D1: A \vee \neg B$$
$$D2: \neg A \vee C$$
$$D3: B \vee \neg C$$
$$D4: \neg A \vee \neg B \vee C.$$

Матрица покрытия по ChatGPT выглядит согласно рис. 1.

Литералы	**A**	**B**	**C**	**¬A**	**¬B**	**¬C**
D1	1	1	0	0	1	0
D2	1	0	1	1	0	0
D3	0	1	0	0	0	1
D4	1	1	1	1	1	0

Рис. 1. Матрица покрытия по ChatGPT
Fig. 1. ChatGPT Coverage Matrix

Правильный ответ с учетом данных, приведенных в [3, с. 108], должен быть таким, как изображено на рис. 2.

Литералы	**A**	**B**	**C**	**¬A**	**¬B**	**¬C**
D1	1	1	0	0	1	0
D2	1	0	1	1	0	0
D3	0	1	0	0	0	1
D4	1	1	1	1	1	0
D5	1	0	0	1	0	0
D6	0	1	0	0	1	0
D7	0	0	1	0	0	1

Рис. 2. Матрица покрытия по [3]
Fig. 2. Coverage matrix according to [3]

Дополнительные строки D5–D7 соответствуют тавтологическим дизъюнктам, добавление которых определяет размер минимального покрытия при выполнимости (невыполнимости) системы дизъюнктов, что необходимо для исключения одновременного появления контрольной пары литер в минимальном покрытии. Нарушение этого требования может давать неверный ответ.

Использование кластеров ключевых слов для генерации вопросов

Ключевые слова помогают сравнительно просто выделить основные идеи и темы, содержащиеся в тексте, и составляют основу многих хорошо зарекомендовавших себя методов. Выделе-

ние ключевых слов – это задача экспертов, либо ее можно решить с помощью средств TextMining, например, реализованных в Python NLTK [4] (т. е. быть полностью автоматизированной). Каждый тестовый вопрос строится на основе нескольких (например, 3–5) выбранных взаимосвязанных ключевых слов. Взаимная связь данных (в том числе ключевых слов) передается обычно через коэффициенты корреляции/ковариации (другие техники могут базироваться на методе типа ANOVA [5], на регрессионном анализе [6] и т. д.).

Для решения задачи отбора ключевых слов в список ключевых слов вопроса строили матрицу информационно-смысловой близости слов (МИСБС). Теоретически числа в ячейках МИСБС могут представлять коэффициенты корреляции между словами, которые можно получить по некоторому корпусу текстов по проблематике экзамена (дисциплины, теста). Это обеспечило бы объективность оценки информационно-смысловой связности слов. Вместе с тем есть проблема, когда коэффициенты корреляции сравнительно одинаковы по всей матрице корреляции, в силу чего определение кластерной структуры сопряжено с известной неопределенностью. Поэтому автоматический расчет коэффициентов корреляции можно заменить заполнением МИСБС экспертом с использованием оценочной шкалы Харрингтона [7]. Шкала Харрингтона – это дискретная оценочная шкала, состоящая из пяти интервалов единичного отрезка со следующими оценочными интервалами для степени близости слов: очень высокая (0,8–1,0), высокая (0,63–0,8), средняя (0,37–0,63), низкая (0,37–0). Имея такую оценочную МИСБС, можно построить кластеры с ограничением на число слов в каждом кластере и число кластеров. Кроме того, с помощью оценочной МИСБС можно построить кластеры с ограничением на число слов в каждом из них. Задача имеет две особенности:

М1: размер кластера ограничен тремя-пятью ключевыми словами, при этом каждый кластер соответствует одному или более тестовому вопросу, который будет сгенерирован для этого кластера;

М2: разные кластеры могут содержать общие ключевые слова (но не совпадать).

Достаточно полное описание известных методов кластеризации дано в [8, с. 20], где кластеры объектов формируются таким образом, чтобы объекты внутри кластера имели высокое сходство и были похожи друг на друга, но весьма не похожи на объекты в других кластерах. В рассматриваемом в статье случае это определение требует существенного уточнения, связанного с более общей трактовкой кластера:

1) «сходство» объектов (ключевых слов) внутри кластера определяется высокими значениями коэффициентов парной корреляции;

2) допустимо, что объекты, попадающие в разные кластеры, также могут иметь высокие коэффициенты попарной корреляции в силу замечания М2.

Важно, чтобы в кластер входили связанные ключевые слова, и при этом одно и то же ключевое слово может попасть более чем в один кластер. Под кластером в данной статье понимается общепринятое понятие объединения однородных элементов, связанных единством темы, рассматриваемых как самостоятельная единица.

Теоретически многие задачи разбиения объектов на кластеры относятся к категории NP-трудных [3, 8]. Задача является эффективно решаемой, поскольку число всех кластеров, содержащих $t = 3$–5 ключевых слов в множестве с $N \geq t$ слов, определяется формулой для числа сочетаний из N по t: $N(N-1) \cdot \ldots \cdot (N-t+1)/t!$ Для фиксированного t это полином $O(N^t)$. Однако при больших N перебор приводит к существенным потерям времени, а при $t \to N/2$ временные затраты становятся экспоненциально сложными.

В связи с этим предлагается приближенный алгоритм отыскания кластеров ключевых слов с практически подтвержденной эффективностью, учитывающий М2. Идея алгоритма основана на последовательном исключении из МИСБС слов с минимальной (суммарной/средней) связностью с остальными словами. Подобный подход аналогичен использованному в [9] методу для отыскания максимальной по размеру нулевой подматрицы 0,1-матрицы, т. е. в описанной реализации нулевые элементы соответствовали бы сильно коррелированным парам ключевых слов, а единичные – некоррелированным или слабо коррелированным, причем 0,1-матрица связей получается из МИСБС заменой каждого числа в ячейке на 0, если это число больше среднего по МИСБС, и на 1 – если не больше.

Алгоритм построения кластеров

В качестве исходных данных задаются множество размером N всех ключевых слов и МИСБС размером $N \times N$. На выходе получаются кластеры Cl_1, Cl_2, …, Cl_z, где z определяется устанавливаемым пользователем числом генерируемых вопросов по теме документа. Реальное число может оказаться меньше, поскольку одинаковые кластеры k исключены. Построение каждого очередного кластера начинается с выбора его центроида a_i. Для слова «центроид a_i» выполняется следующий алгоритм:

шаг 1: удаляются строка и столбец, соответствующие выбранному слову из МИСБС;

шаг 2: подсчитываются суммы элементов в невычеркнутых строках. Выбирается строка (и одноименный столбец) для вычеркивания с вероятностью, обратно пропорциональной сумме элементов в строке (столбце);

шаг 3: если число оставшихся невычеркнутыми строк (столбцов) больше заданного числа слов в кластере, то происходит возврат на шаг 1. Иначе оставшиеся невычеркнутыми строки МИСБС определяют ключевые слова, образующие найденный кластер. Построение очередного кластера завершено.

Рассмотрим пример по теме объектно-ориентированного программирования. Пусть некоторый вариант МИСБС имеет вид, представленный в табл. 1, где a_1 – объект, a_2 – класс, a_3 – метод, a_4 – тип, a_5 – оператор, a_6 – параметр.

Таблица 1. Пример матрицы информационно-смысловой близости слов
Table 1. An example of a matrix of information-semantic proximity of words

Слова	a_1	a_2	a_3	a_4	a_5	a_6
a_1	1	0,8	0,3	0,4	0,7	0,3
a_2	0,8	1	0,6	0	0,5	0,2
a_3	0,3	0,6	1	0,1	0,6	0,6
a_4	0,4	0	0,1	1	0,8	0
a_5	0,7	0,5	0,6	0,8	1	0,7
a_6	0,3	0,2	0,6	0	0,7	1

После первого шага выполняем редукцию матрицы информационно-смысловой близости слов (табл. 2), заключающуюся во включении в кластер строки a_1 (в качестве центроида) и в вычеркивании ее (и столбца a_1) из матрицы.

Таблица 2. Редукция матрицы информационно-смысловой близости слов после первого шага
Table 2. Reduction of the matrix of information-semantic proximity of words after the first step

Слова	a_2	a_3	a_4	a_5	a_6
a_2	1	0,6	0	0,5	0,2
a_3	0,6	1	0,1	0,6	0,6
a_4	0	0,1	1	0,8	0
a_5	0,5	0,6	0,8	1	0,7
a_6	0,2	0,6	0	0,7	1

Подсчитаем суммы элементов в строках и получим следующие значения: строка a_2 – 2,3, a_3 – 2,9, a_4 – 1,9, a_5 – 3,6, a_6 – 2,5. Определим вероятности для строк, обратно пропорциональные этим суммам с помощью Python. Затем случайно выбираем строку по этим вероятностям. Скриншот выбора элемента для удаления по вероятности приведен на рис. 3.

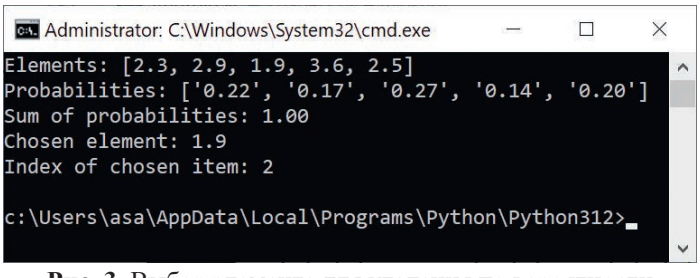


Рис. 3. Выбор элемента для удаления по вероятности
Fig. 3. Selecting an element to remove by probability

На втором шаге удаляем выбранный элемент 1,9 (a_4) из списка, пересчитываем суммы оставшихся строк в матрице и рассчитываем вероятности (табл. 3).

Таблица 3. Редукция матрицы информационно-смысловой близости слов после второго шага
Table 3. Reduction of the matrix of information-semantic proximity of words after the second step

Слова	a_2	a_3	a_5	a_6
a_2	1	0,6	0,5	0,2
a_3	0,6	1	0,6	0,6
a_5	0,5	0,6	1	0,7
a_6	0,2	0,6	0,7	1

Действуя по аналогии далее, получаем в итоге первый кластер из неудаленных элементов (a_1, a_3, a_5) с центроидом a_1. Поскольку элементы кластера выбирались по вероятностям, очередной прогон процедуры может дать другой кластер с тем же центроидом a_1. Поставим задачу получить 2–4 различных кластера с центроидом a_1. Для того чтобы обеспечить неповторение кластеров с данным центроидом, после формирования первого кластера удалим из МИСБС второй выбранный элемент в кластере (a_1, a_3, a_5), т. е. a_3. Повторим процедуру без a_3. Так, при повторном прогоне получили кластер (a_1, a_4, a_2). Снова удалим второй элемент, т. е. a_4, и отыщем очередной кластер – (a_1, a_2, a_5). Ограничиваемся формированием 2–4 кластеров. Каждый кластер используется для генерации связанного с ним тестового вопроса. Кластер из ключевых слов, сформированный за один прогон программы, показан на рис. 4.

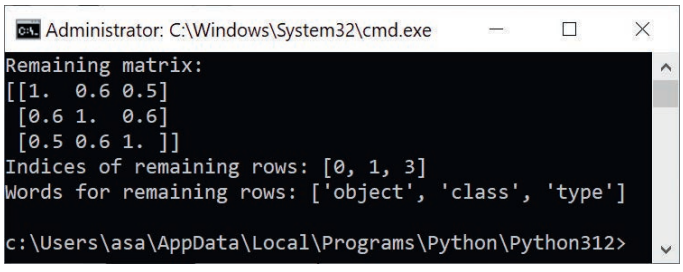


Рис. 4. Кластер из ключевых слов, сформированный за один прогон программы
Fig. 4. A cluster of keywords formed in one run of the program

При повторном прогоне программы определен новый кластер, включающий слова object, type, operator.

Оценка вычислительной сложности и определение числа ключевых слов в тексте

Для оценки вычислительной сложности описанного метода использовали программу на языке Python. В экспериментах отыскивали кластеры из 3–5 ключевых слов, варьируя общее число ключевых слов N в тексте. Фиксировали время, затрачиваемое компьютером на выполнение программы. Результаты приведены в табл. 4.

Таблица 4. Оценка вычислительной сложности для k, равного 3, 4, 5
Table. 4. Estimated computational complexity for k equal to 3, 4, 5

N	Время, с, для		
	$k = 3$	$k = 4$	$k = 5$
10	0,24	0,31	0,31
20	0,62	0,62	0,62
50	2,15	2,19	2,17
200	16,7	17,3	17,4

Выполняли эксперименты для оценки вычислительной сложности порождения различных кластеров, содержащих $k = 3$–5 ключевых слова из N (метод простого перебора), с последующим отбором четырех кластеров с максимальной суммой парных коэффициентов корреляции. Результаты представлены в табл. 5.

Таблица 5. Оценка вычислительной сложности переборного метода для k, равного 3, 4, 5
Table. 5. Estimation of computational complexity of the enumeration method for k equal to 3, 4, 5

N	Время, с, для		
	$k = 3$	$k = 4$	$k = 5$
10	0,15	0,15	0,31
20	0,50	1,37	3,7
50	5,8	79,9	1143,7
200	1238,2	Остановлено после 30 мин счета	

Приведенные в табл. 4, 5 экспериментальные данные определяют преимущество предложенного подхода. Задачу определения ключевых слов можно связать с задачей о минимальном покрытии 0,1-матрицы множеством строк. Воспользуемся оценкой [10], согласно которой число строк μ в покрытии π 0,1-матрицы **B** с R строками и C столбцами, которое отыскивается «жадным» алгоритмом, оценивается как:

$$\mu \approx 1 - [\ln(Cpq^{0,5})]/\ln q, \quad (1)$$

где p – плотность единичных элементов в матрице **B**; q – плотность нулевых элементов в матрице **B**, $q = 1 - p$.

Плотность p вычисляется по формуле

$$p = E/RC, \quad (2)$$

где E – число «1» в матрице **B**.

Покрытие π состоит из строк матрицы **B**, причем хотя бы одна строка из π содержит «1» в каждом столбце матрицы **B**. «Жадный» алгоритм на каждом шаге включает в покрытие строку с максимальным числом «1» и удаляет покрываемые ею столбцы из **B**, пока таковые есть.

Воспользуемся следующими оценками. Среднее число слов в предложении английского текста составляет 15–20 (данные ChatGPT). Текст из 10 000 слов будет включать порядка 500 предложений. Каждое предложение соответствует столбцу 0,1-матрицы **B** ($C = 500$), каждое слово – строке матрицы **B**. В среднем в англоязычном тексте употребляется 1000–3000 различных слов w_k. Возьмем половину этого диапазона – 1500 ($R = 1500$). Итак, при таких «среднеупотребительных» данных число единиц в матрице **B** составит $E = 10\ 000$ (по общему числу слов в тексте). Единица в строке i и столбце j означает, что слово w_i вошло в предложение (столбец) C_j. Размер матрицы **B** составит $CR = 500 \cdot 1500 = 750\ 000$. Отсюда плотность единиц $p = 10\ 000/750\ 000 = 1/75$. Соответственно, плотность «0» $q = 1 - p = 74/75$. Из (1) вычислим, что $\mu \approx 141$. Таким образом, при $\mu \approx 141$ каждое предложение будет содержать как минимум одно ключевое слово. Увеличив число ключевых слов, мы, очевидно, повысим качество поиска ответов. Для практических целей можно использовать приближенную оценку числа ключевых слов, выбираемых из текста с общим числом слов Z как $\mathrm{const}Z^{0,5}$, где коэффициент const для текстов в несколько десятков тысяч слов близок к 2–3.

Генерация тестовых вопросов и получение ответов на них

Используя полученные ключевые слова, мы обратились к ChatGPT для генерации вопросов с этими словами и ответов на них. Получили следующие варианты вопросов, выданные ChatGPT:

вопрос *a*: How do objects, types, and classes relate? (Как соотносятся объекты, типы и классы?). (Explain the relationship between objects, types, and classes in object-oriented programming. How does a class define the type of an object?);

вопрос *b*: Can you have multiple objects of the same type? (Можете ли вы располагать множеством объектов одного и того же типа?). (Discuss how multiple objects can be created from a single class and how they share the same type. What implications does this have for memory usage and data management?).

На приведенные вопросы в ChatGPT получены следующие ответы:

ответ на вопрос *a*: In essence, classes define types, and objects are instances of these types, allowing for organized and modular programming (По сути, классы определяют типы, а объекты являются экземплярами этих типов, обеспечивая организованное модульное программирование);

ответ на вопрос *b*: Yes, you can have multiple objects of the same type in object-oriented programming (У вас может быть несколько объектов одного и того же типа в объектно-ориентированном программировании).

Глобальная языковая модель типа ChatGPT вполне приемлема для генерации вопросов к общим темам программирования, физики, истории и т. п. Но и в этом случае определение ключевых слов для конкретного учебного документа отражает его специфику (например, ограниченное по затрагиваемым темам содержание, для которого нужно сформулировать тестовые вопросы). Для узкоспециализированной тематики более точна локальная модель.

В качестве примера использования локальной языковой модели можно взять [11]. Работа такой модели основана на определении позиций ключевых слов вопроса в оригинальном тексте и локализации смысловых блоков (обычно параграфов) с этими ключевыми словами. Неключевые слова вопроса используются для уточнения оценки степени соответствия каждого из найденных смысловых блоков тексту вопроса. Формулировки вопросов могут строиться на базе общих шаблонов типа «Укажите связь между k_1, k_2, …, k_z» (где k_1, k_2, …, k_z – ключевые слова); «Дайте определение k_1», «В каком контексте используют k_1, k_2, …, k_z» и др. Более интересен и практически целесообразен способ генерации вопросов непосредственно к смысловым блокам, предварительно определяемым по ключевым словам. При этом можно обратиться к ChatGPT для генерации примеров вопросов к данному текстовому блоку и воспользоваться этим блоком в качестве ответа. Кроме того, ChatGPT дает примеры неправильных ответов, что позволяет использовать тесты для контроля знаний.

Заключение

1. Реализованный подход может быть в значительной степени автоматизирован, допуская участие эксперта для оценки качества и содержательности построенных тестов и ответов на них.

2. Предложен метод определения кластеров ключевых слов для формирования на их основе тестовых вопросов и получения ответов из используемых языковых моделей.

3. Приведены теоретические оценки для числа ключевых слов в тексте и получены результаты практической апробации, свидетельствующие об эффективности всего подхода.

Список литературы

1. A Review on Question Generation from Natural Language Text / Ch. Ruqing [et al.] // ACM Trans. on Information Systems. 2021. Vol. 40, No 1.
2. Brusilowsky, P. Web-Based Testing for Distance Education / P. Brusilowsky, P. Miller // WebNet’99, World Conference of the WWW and Internet, AACE, Honolulu, 1999. P. 149–154.
3. Герман, О. В. Введение в теорию экспертных систем и обработку знаний / О. В. Герман // Минск: ДизайнПро, 1995.
4. Bird, S. Natural Language Processing with Python / S. Bird, E. Klein, E. Loper // O’Reilly Media Inc. USA, 2007.
5. Montgomery, D. C. Design and Analysis of Experiments / D. C. Montgomery // John Wiley&Soms Inc. USA, 2013.
6. Neter, J. Applied Linear Regression Models / J. Neter, W. Wasserman, M. H. Kutner // R. D. Irwin Inc. USA, 1983.
7. Blei, D. M. Latent Dirichlet Allocation / D. M. Blei, A. Y. Ng, M. I. Jordan // Journal of Machine Learning Research. 2003. No 3. P. 993–1022.
8. Han, J. Data Mining. Concepts and Techniques / J. Han, M. Kamber, J. Pei // Elsevier. 2012.
9. Герман, О. В. Алгоритм булевой оптимизации на (0,1)-матрицах / О. В. Герман, Е. И. Германович, В. Г. Найденко // Журнал вычислительной математики и математической физики. 1992. Т. 32, № 7. С. 1114–1125.
10. Герман, О. В. Разрешающий принцип для задачи о минимальном покрытии 0,1-матрицы / О. В. Герман // Кибернетика и системный анализ. 1996. № 1. С. 135–145.
11. Герман, Ю. О. Метод извлечения информации из резюме / Ю. О. Герман, О. В. Герман, С. Наср // Труды БГТУ. Серия 3. Физико-математические науки и информатика. 2019. № 1. С. 64–68.

Поступила 08.01.2025 Принята в печать 22.05.2025

References

1. Ruqing Ch., Jiafeng G., Lu Ch., Xueqi Cheng., Yixing Fan. (2021) A Review on Question Generation from Natural Language Text. *ACM Trans. on Information Systems.* 40 (1).
2. Brusilowsky P., Miller P. (1999) Web-Based Testing for Distance Education. *WebNet'99, World Conference of the WWW and Internet, AACE, Honolulu*. 149–154.
3. German O. V. (1995) *Introduction to Expert Systems Theory and Knowledge Processing*. Minsk, DesignPro Publ. (in Russian).
4. Bird S., Klein E., Loper E. (2007) Natural Language Processing with Python. *O'Reilly Media Inc*. USA.
5. Montgomery D. C. (2013) Design and Analysis of Experiments. *John Wiley&Soms Inc*. USA.
6. Neter J., Wasserman W., Kutner M. H. (1983) Applied Linear Regression Models. *R.D. Irwin Inc*. USA.
7. Blei D. M., Ng A. Y., Jordan M. I. (2003) Latent Dirichlet Allocation. *Journal of Machine Learning Research*. (3), 993–1022.
8. Han J., Kamber M., Pei J. (2012) Data Mining. Concepts and Techniques. *Elsevier*.
9. German O. V., Germanovich E. I., Naidenko V. G. (1992) Boolean Optimization Algorithm on (0,1)-Matrices. *Journal of Computational Mathematics and Mathematical Physics*. 32 (7), 1114–1125 (in Russian).
10. German O. V. (1996) Decision Principle for the Minimum Covering Problem of a 0,1-Matrix. *Cybernetics and Systems Analysis*. (1), 135–145 (in Russian).
11. German Yu. O., German O. V., Nasr S. (2019) Information Extraction Method from a Resume (CV). *Works BSTU. Series 3. Physics and Mathematics Science and Computer Science*. (1), 64–68.

Received: 8 January 2025 Accepted: 22 May 2025

Вклад авторов

Мигалевич С. А. определил способ генерации кластеров ключевых слов и формирования текстов вопросов к кластерам.

Герман Ю. О. осуществила взаимодействие системы генерации вопросов с ChatGPT для формирования вопросов и получения ответов.

Герман О. В. определил общую концепцию работы и основные задачи.

Authors' contribution

Migalevich S. A. defined the method to generate clusters of keywords and creating questions associated to clusters.

German Ju. O. realized the interaction of the proposed system and ChatGPT to form questions and getting answers.

German O. V. defined the general concept of the work and the related basic tasks.

Сведения об авторах

Мигалевич С. А., нач. центра информатизации и инновационных разработок, соискатель каф. программного обеспечения информационных технологий, Белорусский государственный университет информатики и радиоэлектроники (БГУИР)

Герман Ю. О., канд. техн. наук, доц. каф. информационных технологий автоматизированных систем, БГУИР

Герман О. В., канд. техн. наук, доц. каф. информационных технологий автоматизированных систем, БГУИР

Адрес для корреспонденции

220013, Республика Беларусь,
Минск, ул. П. Бровки, 6
Белорусский государственный университет
информатики и радиоэлектроники
Тел.: +375 17 293-23-20
E-mail: migalevich@bsuir.by
Мигалевич Сергей Александрович

Information about the authors

Migalevich S. A., Head of the Center of Informatization and Innovations, Applicant at the Department of Software Provision of Information Technologies, Belarusian State University of Informatics and Radioelectronics (BSUIR)

German Ju. O., Cand. Sci. (Tech.), Associate Professor at the Information Technologies in Automatized System Department, BSUIR

German O. V., Cand. Sci. (Tech.), Associate Professor at the Information Technologies in Automatized System Department, BSUIR

Address for correspondence

220013, Republic of Belarus,
Minsk, P. Brovki St., 6
Belarusian State University
of Informatics and Radioelectronics
Tel.: +375 17 293-23-20
E-mail: migalevich@bsuir.by
Migalevich Sergei Alexandrovich

CC BY

http://dx.doi.org/10.35596/1729-7648-2025-23-3-62-69

УДК 004.051

A COMPARATIVE ANALYSIS OF QoS, GoS, AND QoE METRICS IN NETWORK PERFORMANCE EVALUATION

VAGIF A. GASIMOV, LEYLA R. ZEYNALLI-HUSEYNZADE, GULBICHA A. ZAHIDOVA

Baku Engineering University (Baku, Republic of Azerbaijan)

Abstract. The concepts of QoE (quality of experience), QoS (quality of service), and GoS (goal of service) are commonly discussed in network performance and user satisfaction studies. Although QoE has become a popular topic in research, the boundaries between QoS and QoE are often blurred, making it difficult to clearly define them. This paper examines the differences and relationships between these terms with regard to their practical applications. QoS is a subjective metric that reflects how users perceive a service. It is influenced by personal preferences and various environmental factors. GoS measures the probability of a successful connection or call under certain conditions. The results showed that implementing QoS features such as traffic prioritization can positively affect both GoS and QoE by reducing packet loss and improving service reliability. It is shown how network performance management using QoS tools can improve user experience and overall service reliability, providing a clearer understanding of how these concepts interact in practice.

Keywords: quality of service, grade of service, quality of experience, network performance metrics.

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

For citation. Gasimov V. A., Zeynalli-Huseynzade L. R., Zahidova G. A. (2025) A Comparative Analysis of QoS, GoS, and QoE Metrics in Network Performance Evaluation. *Doklady BGUIR*. 23 (3), 62–69. http://dx.doi.org/10.35596/1729-7648-2025-23-3-62-69.

СРАВНИТЕЛЬНЫЙ АНАЛИЗ ПОКАЗАТЕЛЕЙ QoS, GoS И QoE ПРИ ОЦЕНКЕ ПРОИЗВОДИТЕЛЬНОСТИ СЕТИ

В. А. КАСУМОВ, Л. Р. ЗЕЙНАЛЛЫ-ГУСЕЙНЗАДЕ, Г. А. ЗАХИДОВА

Бакинский инженерный университет (Хырдалан, Азербайджанская Республика)

Аннотация. Концепции QoE (качество взаимодействия), QoS (качество обслуживания) и GoS (уровень обслуживания) обычно обсуждаются в исследованиях производительности сети и удовлетворенности пользователей. Несмотря на то, что QoE стало популярной темой в исследованиях, границы между QoS и QoE часто размыты, что затрудняет их четкое определение. В статье рассмотрены различия и связи между этими терминами с учетом их практического применения. Качество обслуживания – это субъективный показатель, отражающий то, как пользователи воспринимают услугу. На него влияют личные предпочтения и различные факторы окружающей среды. GoS измеряет вероятность успешного соединения или вызова при определенных условиях. Результаты показали, что внедрение функций QoS, таких как приоритизация трафика, может положительно повлиять как на GoS, так и на QoE, уменьшая потерю пакетов и повышая надежность обслуживания. Показано, как управление производительностью сети с помощью инструментов QoS может повысить удобство работы пользователей и общую надежность услуг, обеспечивая более четкое понимание того, как эти концепции взаимодействуют на практике.

Ключевые слова: качество обслуживания, уровень сервиса, качество взаимодействия, показатели производительности сети.

Конфликт интересов. Авторы заявляют об отсутствии конфликта интересов.

Для цитирования. Касумов, В. А. Сравнительный анализ показателей QoS, GoS и QoE при оценке производительности сети / В. А. Касумов, Л. Р. Зейналлы-Гусейнзаде, Г. А. Захидова // Доклады БГУИР. 2025. Т. 23, № 3. С. 62–69. http://dx.doi.org/10.35596/1729-7648-2025-23-3-62-69.

Introductıon

In the context of the network media ecosystem, ensuring quality requires a holistic approach that combines technical specifications, user perspectives, and system considerations. However, in most digital multimedia entertainment programs, the overall quality of experience (QoE) accepted by the end user/viewer is considered. QoE encompasses how good the video looks, how well the audio is rendered, how well the combined audio-visual content is accepted, and how well a specific audio-visual service's mutual relationship functions. When referring to video content, quality of service (QoS) is on par with any other service. This includes coverage, continuity of service functionality, provided reliability, device compatibility, and ensuring the quality of signals and features that enable usage. Many agree that digital communication is one of the driving forces for the development of businesses and societies, and it is normal for everyone to expect the best from the telecommunication services they use in their daily digital interactions.

In today's competitive market reality, the customer perspective will definitely demand rejection if a service cannot deliver the highest quality. On the other end of the line, service providers may not be fully informed about the tools that can be used to meet their customers' needs or address their dissatisfaction. To bridge this gap, the concept QoS was introduced as an initiative to enter users' bots and ultimately provide better services and solutions. However, the abundance of overlapping understandings and their confusing abbreviations in telecommunication services does not facilitate the understanding of the importance of QoE. While concepts such as UX (user experience), QoS, GoS (grade of service) or QoE are useful in determining both the quality and quantity parameters of the interaction between the user and the system or service, there has been a recent consensus among experts that the focus should be on QoE [1]. To assess customer satisfaction, QoE has been widely adopted in various industry parameters related to consumer-related aspects. However, since its inception, the Paradigm of QoE has mainly been applied in the telecommunications field, where research tools are the sharpest and measurements are the most accurate.

QoS is a term referring to the technology used to manage information traffic over the network. QoS aims to control and manage network resources to reduce interference such as packet loss and delay. QoS also determines limitations and priorities for different categories of data moving between IP networks, such as bandwidth traffic on the network. It is expected that today's businesses will provide secure and reliable services to end-users with minimal interruptions. In recent years, applications such as audio, video, file sharing, and data streaming have increasingly become part of our daily lives, making QoS more important than ever. The volume of application usage, the increasing number of connected devices to the network, and the significant increase in the importance of social media lead to frequent network congestion. Excessive network loading can cause performance differences. As a result, IT departments are filled with information about video conferences being interrupted, poor audio quality, delays, and even dropped phone calls, which significantly impact daily productivity in the workplace [2]. QoS manages network resources by prioritizing certain types of data communication to determine the transmission needs of multiple data.

Quality of experience and measurement

In the realm of telecommunications, QoE encompasses a broader and more comprehensive range compared to the concept of QoS, which focuses on measuring, improving, and ensuring software and hardware features. To put it simply, QoE refers to the measurement of overall satisfaction with a service from the user's perspective. However, the challenging part comes next. True QoE doesn't aim to measure the objective parameters of system performance (as in QoS). Instead, it aims to encompass the subjective experiences of the service user, including all complexities and factors related to human interactions, such as physical, temporary, social, and even economic aspects. But is it possible to accurately measure someone's subjective feelings about something and derive useful insights from aggregated data? The answer is yes, but with some "buts". Measuring QoE in telecommunications services is quite complex and cannot be successfully achieved without proper tools for data collection and analysis [3]. Therefore, the nature of determining QoE in a specific situation does not solely depend on the quantity or scope of what is measured and converted into metrics. It rather involves understanding which service parameters are the main drivers of user satisfaction and measuring them from a perspective as close to the user's perception as possible. Over the years, awareness of how QoE contributes to increasing user satisfaction and, consequently, user loyalty has been growing among telecommunications operators.

As a result, solutions for monitoring the QoE parameters, especially in telecommunications services, have emerged. However, deriving reliable and consistent results based on these parameters is quite challenging because QoE aims to encompass every factor related to users' perception of service quality, including people, systems, and contextual aspects. Thus, QoE is highly reliant on subjective evaluations. Nevertheless, QoE is determined by using statistical analysis of objective measures through Customer Satisfaction Surveys [4]. QoE evaluations can be subjective or objective. Subjective evaluations are determined by inquiring about quality from human evaluators (though not just based on a single user's opinion). Objective evaluations assess quality based on values obtained through the measurement of assessed systems. Subjective evaluations typically require human evaluators who assess each experience using a five-point scale. The average (or "mean") of the ratings is then considered as the mean opinion score. Evaluations can be conducted in a laboratory or in the field, where they have the following advantages and disadvantages:

• in the laboratory, subjective evaluations can simulate certain conditions of normal usage to some extent, but they never fully reflect normal usage. However, they can be carefully controlled so that evaluators can focus their attention on specific aspects of interest. In particular, evaluated systems can be modified to learn how the evaluators' experience scores change over time. Typically, there are several evaluators (and variations in systems) for people to spend time in the laboratory, due to the costs and efforts involved;

• in the field, subjective judgments can be made under certain conditions of normal usage. For example, services with limited coverage can be tested when the coverage area is available. However, they cannot be performed under all conditions of normal usage. For example, residential multimedia streaming services may not be tested during peak hours at homes. If opinions are gathered through crowdsourcing to assess their own experiences where and when they want, there can be multiple raters, but they may not represent users as a whole.

The most complete possible subjective evaluations of systems do not take into account people's feelings and environments at places and times other than those of the evaluators. Objective evaluations do not require human evaluators, which is why they are often cheaper and faster than subjective evaluations (Fig. 1). They involve measurements of the evaluated systems that can be combined with mathematical models to derive values that often provide mean score estimates. Their accuracy depends on the realism of the models in determining the quality accepted by the average of human evaluators. Typically, models are based on perception theories, empirical studies of influencing factors, and subjective evaluations. They do not consider all factors related to human and context effects on systems. These models are suitable for adaptive streaming and Over-The-Top and managed content providers (ITU-T 2020).

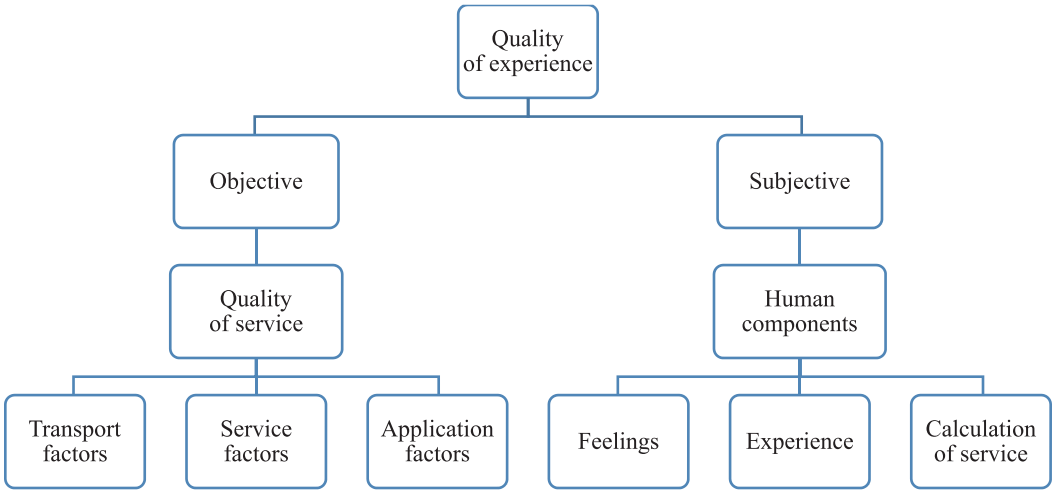


Fig. 1. Types of quality of experience evaluation

Objective evaluations can be conducted in a laboratory, in the field, or in a network environment. They have the following advantages and disadvantages:

• objective evaluations in the laboratory have the same advantages and disadvantages as subjective evaluations in the laboratory;

• objective evaluations in the field can be more reliable than subjective evaluations because they do not require human evaluators. They can be performed at any location and time with appropriate evaluation tools. Installation may require users to have equipment and software installed at their homes or terminals or to have time to travel with equipment or software provided by testers. While achieving this is still challenging, it should be easier than compelling users to conduct regular tests;

• objective evaluations in network environments resemble objective evaluations in the laboratory as they are still not performed under conditions of normal usage. The installation of evaluation tools may be unnecessary, and if equipment and software are already integrated into network nodes or support systems, measurements can cover existing values;

• objective evaluations can be performed without measuring the evaluated systems. Instead, values based on previous knowledge are accepted and combined in a model to provide mean score estimates. This technique is useful for planning networks. Most work on developing QoE standards is based on laboratory evaluations, which can result in products consisting of hardware and software to perform objective evaluations that correspond to subjective mean score evaluations.

Quality of service

Service quality describes and measures the overall performance or output quality of a network service. It can apply to various services such as telephone companies providing voice services, computer networks, or even cloud computing networks. Network performance is measured based on the level of satisfaction experienced by end-users, rather than the guarantees provided by service providers. QoS can be quantified using parameters such as packet loss, bit rate, transmission delay, throughput capacity, availability, and more. However, in the case of packet-switched telecommunication networks like telephone and computer networks, QoS is used to describe traffic management, prioritization, and resource allocation, unlike the actual quality of voice or video streams. Many organizations offer delay-sensitive services like real-time video and voice communication, and they employ QoS to meet the traffic requirements of these types of data. Service class (CoS) can be associated with service quality, but these two concepts are not exactly the same. CoS applies a less detailed approach to traffic management compared to QoS. However, many people use the terms interchangeably, although they represent different concepts. QoS is primarily used by network systems that control data traffic for resource-constrained systems. These services include internet protocol television, online gaming, video conferencing, voice over IP (VoIP) over the internet, video and audio streaming, and demand-based video [5]. Service quality also applies to IoT industries, businesses, and individual end-users. Various QoS tools available possess similar functionalities regardless of the manager. These functionalities include classification, queuing, policing, shaping, weighted random early detection, fragmentation, and compression. QoS can be associated with ensuring the smooth operation of emergency assistance under congested traffic conditions. However, to understand the concept adequately, it is necessary to examine the parameters related to service quality.

The picture below shows the results of the two speed tests, illustrating significant differences in download and upload speeds. Download speed increased from 8.84 to 59.55 Mbps with QoS, highlighting more efficient bandwidth utilization. Upload speed improved from 29.93 to 61.68 Mbps, demonstrating how QoS prioritizes upstream traffic. Latency (Ping) remained stable at 5 ms in both tests, indicating consistent network response times. The figure visually represents the difference in download and upload speeds between the two scenarios:

1) with QoS – both download and upload speeds show a significant increase, ensuring better performance for critical tasks;

2) without QoS – bandwidth allocation is inefficient, leading to lower overall speeds, especially for high-demand activities.

These results clearly show that enabling QoS significantly improves network performance by managing bandwidth more effectively. This ensures that critical services, such as video conferencing or streaming, receive the necessary resources for smooth operation.

Grade of service

It is a measure used to evaluate the performance and availability of a telecommunication network or service. GoS indicates the probability or percentage of successfully completing a call or connection

within certain specified criteria. For example, if a GoS of 0.01 is specified, it means that only 1 out of 100 attempted calls may be blocked or fail to connect due to resource limitations. Conversely, a GoS of 0.001 indicates a higher level of service, with only 1 out of 1000 calls being blocked. The specific criteria used to define GoSmay vary depending on the type of service and the industry. In telecommunication networks, GoSis often expressed in terms of call blocking probability, which measures the likelihood of a call being blocked or rejected due to insufficient network resources. A lower call blocking probability indicates a better GoS.

GoS plays a crucial role in network planning, dimensioning, and capacity management. It helps service providers determine the required network resources, such as bandwidth, channels, or connections, to meet the desired level of service [6]. By analyzing the GoS requirements and network performance, providers can make informed decisions on network upgrades, expansions, or optimizations (Fig. 2).

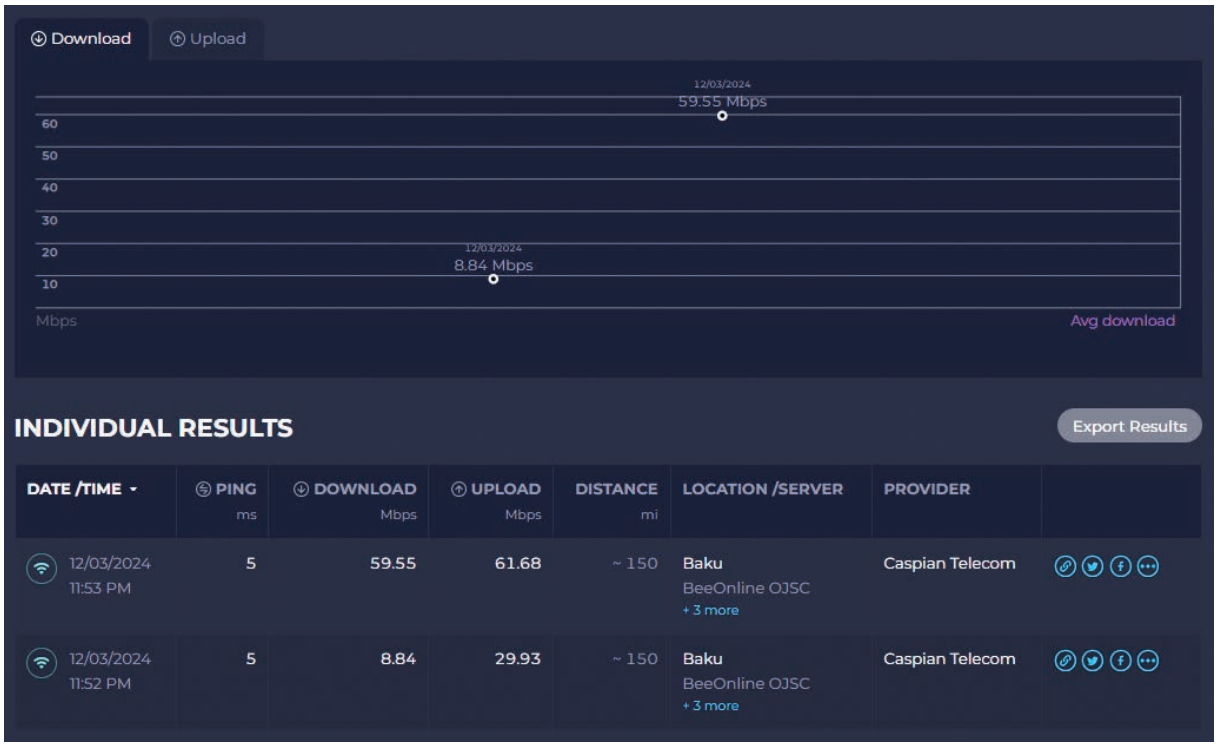


Fig. 2. Comparison of network performance with and without quality of service

Overall, GoS provides a quantitative measure of the service availability and success rates in telecommunication networks. It helps ensure that the network can deliver the desired level of service and meet the expectations of users or customers. GoS metrics are measurements used to assess the performance and quality of a telecommunication system. These metrics help evaluate the level of service provided to users and identify areas for improvement [7]. Here are some common GoS metrics:

• call blocking probability: this metric represents the probability that a call or connection request will be denied or blocked due to insufficient resources in the system. It is typically expressed as a percentage or decimal value, with lower values indicating better service quality;

• call completion rate: this metric measures the percentage of calls or connections that are successfully established and completed without being dropped or disconnected. A higher call completion rate indicates a more reliable and robust system;

• call setup time: this metric measures the time it takes to establish a call or connection from the moment the request is initiated. It includes the time for signaling, routing, and resource allocation. A shorter call setup time indicates faster and more efficient call establishment;

• network congestion: this metric assesses the level of congestion or traffic load on the network. It measures factors such as the number of active calls, data throughput, or utilization of network resources. Monitoring network congestion helps identify periods of high demand and potential bottlenecks;

• QoS parameters: GoS metrics may also include various QoS parameters that measure the quality and performance of the communication service. These parameters can include metrics such as packet

loss, delay, jitter, throughput, and availability. They provide insights into the reliability and effectiveness of the network in delivering data or voice services;

• service level agreements (SLAs): SLAs are contractual agreements between service providers and customers that define the expected levels of service quality and performance. GoS metrics can be used to monitor and measure the compliance with SLAs, ensuring that the agreed-upon service levels are being met.

These metrics are typically monitored and analyzed regularly to evaluate the performance of the telecommunication system, identify areas for improvement, and make informed decisions regarding network optimization, capacity planning, and resource allocation.

Assessment of TCP errors with metrics

In this practical analysis, we evaluated the effect of enabling QoS features on a Wi-Fi router to enhance network performance, with a focus on reducing TCP errors and improving both QoE and GoS. The experiment involved capturing network traffic, identifying errors, and assessing the impact of QoS on real-time application performance under typical home network conditions.

The experiment followed these key steps:

• enabling QoS on the router: QoS features were configured to prioritize latency-sensitive traffic (e. g., video streaming, VoIP) while assigning lower priority to background downloads. Rules were set based on application type and specific devices connected to the network;

• traffic monitoring: network traffic was captured using Wireshark on a connected device while various activities, such as streaming, browsing, and downloading, were performed simultaneously;

• error identification: Wireshark filters were applied to isolate TCP errors, focusing on retransmissions, duplicate acknowledgments, and lost packets;

• visualization: TCP error trends were visualized using Wireshark's I/O Graph, which displayed the frequency of errors over time in 1-second intervals.

To conduct the analysis, all QoS settings on the router were enabled, prioritizing high-bandwidth and real-time applications such as video streaming and VoIP traffic. Devices generating significant background traffic were deprioritized to allow optimal allocation of network resources. During the tests, network traffic was captured using Wireshark while activities such as streaming, video calls, and concurrent downloads were performed to simulate real-world conditions. TCP errors, including retransmissions, duplicate acknowledgments, and lost packets, were isolated using specific filters in Wireshark. The captured data was visualized through the I/O graph tool, which displayed error trends over time. Metrics relevant to QoE and GoS, such as error frequency, packet retransmissions, and service stability, were evaluated to understand the impact of QoS (Fig. 3).

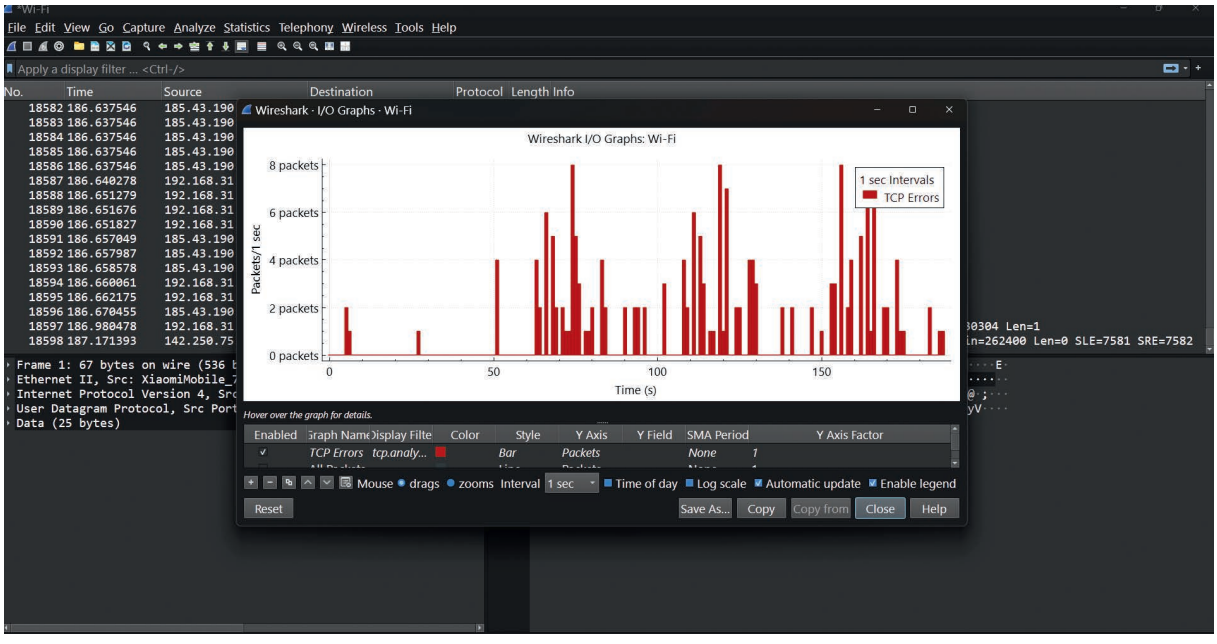


Fig. 3. I/O graph from Wireshark, with the data focusing on TCP errors over time

The results demonstrated a significant reduction in the frequency and duration of TCP error spikes after enabling QoS. The I/O graph revealed that error bursts during high traffic demand were shorter and less frequent compared to scenarios without QoS. High-priority applications, such as video calls and streaming services, maintained consistent performance with minimal interruptions. Background downloads experienced minor delays but did not disrupt overall network functionality. With fewer retransmissions and reduced packet loss, video streaming was smoother, and VoIP calls achieved high audio and video quality.

Enabling QoS directly improved the QoE by minimizing disruptions and maintaining consistent service quality during high-demand scenarios. In terms of GoS, QoS ensured reliable service delivery by managing network congestion, reducing packet loss, and prioritizing critical traffic. This demonstrated a clear enhancement in service reliability metrics. The findings highlight the practical value of QoS in improving QoE and GoS in modern networks. By enabling QoS features, network administrators can optimize traffic flow, mitigate the impact of congestion, and ensure high-quality performance for real-time applications. This approach is particularly relevant in environments with bandwidth limitations or competing traffic demands

Conclusion

1. Quality of service (QoS), quality of experience (QoE), and grade of service (GoS) are three related concepts in telecommunications and network management, each focusing on different aspects of service quality:

• QoS:

– refers to measurable parameters that define the performance and reliability of a network or service;

– primarily considers technical factors such as bandwidth, delay, packet loss, jitter, throughput, and availability;

– is defined and implemented by network administrators or service providers to meet specific performance criteria;

– is objective and can be managed using network engineering techniques like traffic shaping, prioritization, and resource allocation;

• QoE:

– focuses on the subjective evaluation of the overall user experience when using a service or application;

– includes factors such as perceived quality of video/audio, responsiveness, ease of use, interactivity, and user satisfaction;

– is influenced by user expectations, personal preferences, and contextual factors;

– is typically assessed through user feedback, surveys, and subjective tests that reflect the user's perception of service quality;

• GoS:

– relates to the level of service provided by a telecommunication network, particularly in terms of the probability of call blocking or dropping;

– is commonly used in voice and real-time communication systems to assess the system's ability to handle traffic load while maintaining service quality;

– relies on the network's capacity to manage expected traffic and ensure desired service levels through proper resource allocation and planning.

2. While these concepts are distinct, they are interconnected and can influence each other. Effective QoS mechanisms directly impact network performance, which, in turn, affects GoS and QoE. A well-designed QoS framework can improve both GoS and QoE by ensuring optimal network performance. Similarly, managing GoS helps reduce call blocking and dropping rates, which can enhance user satisfaction (QoE).

3. The successful implementation of these metrics requires technical expertise, data analysis, and a user-centric approach. Network administrators and engineers must consider the specific requirements of applications and users to optimize QoS, GoS, and QoE. Continuous monitoring and adaptation are essential as technology evolves and user expectations change.

4. This study emphasizes the practical configurations and analysis involved in improving QoS, QoE, and GoS. Implementing QoS features on a Wi-Fi router and analyzing network traffic with tools like

Wireshark demonstrated that QoS mechanisms are crucial for enhancing network performance and user satisfaction.

5. The assessment showed that enabling QoS reduced TCP errors, minimized packet loss, and ensured consistent performance for high-priority applications such as video streaming and VoIP. The I/O graph analysis revealed a decrease in error spikes during high traffic, confirming the effectiveness of traffic prioritization. Additionally, real-time network monitoring and error analysis provided actionable insights into improving QoE and GoS.

6. These findings highlight the importance of proactive network management in environments with diverse traffic demands. By prioritizing critical applications and balancing resource allocation, QoS improves both user experience and network reliability. The results indicate that QoS is not just a theoretical concept but a practical solution for addressing real-world network challenges.

7. Study underscores the importance of applying theoretical concepts in practical scenarios to achieve measurable improvements in network performance. Future research could explore advanced QoS algorithms or evaluate performance under varying network conditions to further refine and optimize service quality.

References

1. Hayder S., Shah N., Awan A. A. (2015) Comparative Analysis of QoS, QoE, and GoS in Multimedia Networks. *IEEE International Conference on Cyber Security of Smart Cities, Industrial Control System and Communications (SSIC)*. Riyadh, Saudi Arabia. 1–6.
2. Rahim A., Zain A. N. M., Rahman A. F. A., Nor M. F. A. M., Salleh R. (2013) A Comparative Analysis of QoS, GoS, and QoE Metrics in Multimedia Communication. *International Conference on Informatics and Creative Multimedia (ICICM)*. Kuala Lumpur, Malaysia. 89–94.
3. Hossain S., Matin M. A., Ahmed M. A. (2019) Comparative Analysis of QoS, QoE and GoS Metrics in Telecommunication Network. *5th International Conference on Electrical Engineering and Information & Communication Technology (ICEEICT)*. Dhaka, Bangladesh. 1–6.
4. Prajapati A. M. V., Solanki A. M., Jain S. L. B. (2019) Comparative Analysis of QoS, GoS and QoE Parameters in WiMAX Network. *International Conference on Power Electronics, Smart Grid and Renewable Energy (PESGRE)*. Ahmedabad, India. 1–6.
5. Choi J. M., Kong S. G., Kwak K. S. (2013) Comparative Analysis of QoS and QoE Metrics in the Context of IPTV. *International Conference on Information Networking (ICOIN)*. Bangkok, Thailand. 317–322.
6. Akhtar M. S., Islam M. Z., Ali A. B. M. S. (2019) Comparative Analysis of QoS, QoE and GoS in Wireless Multimedia Communication. *International Conference on Electrical, Computer and Communication Engineering (ECCE).* Cox's Bazar, Bangladesh. 1–6.
7. Cano M., Hernandez J. A., Alcober J., Fernández A., Soto M. (2016) A Comparative Analysis of QoS, QoE, and GoS Metrics for Multimedia Services in Wireless Networks. *IEEE International Conference on Communications (ICC).* Kuala Lumpur, Malaysia. 1–6.

Received: 7 March 2025 Accepted: 10 April 2025

Authors' contribution

The authors contributed equally to the writing of the article.

Information about the authors

Gasimov V. A., Dr. Sci. (Tech.), Professor, Dean of the Faculty of Informationa and Computer Technologies, Baku Engineering University (BEU)

Zeynalli-Huseynzade L. R., Lecturer at the Department of Cybersecurity and Computer Engineering, BEU

Zahidova G. A., Lecturer at the Department of Cybersecurity and Computer Engineering, BEU

Address for correspondence

AZ0101, Republic of Azerbaijan,
Khirdalan City, Hasan Aliyev St., 120
Baku Engineering University
Tel.: +994 70 213-99-98
E-mail: vaqasimov@beu.edu.az
Gasimov Vagif Alidjavad ogly

CC BY

http://dx.doi.org/10.35596/1729-7648-2025-23-3-70-76

УДК 004.93

ОЦЕНКА СХОДСТВА МЕЖДУ НАБОРАМИ ДАННЫХ С ПОМОЩЬЮ ВЕКТОРНЫХ ПРЕДСТАВЛЕНИЙ

А. А. УСАТОВ, А. М. НЕДЗЬВЕДЬ, ГО ЦЗИЖАНЬ

Белорусский государственный университет (Минск, Республика Беларусь)

Аннотация. Рассмотрен подход к определению сходства наборов данных (датасетов) для обучения алгоритмов на примере датасетов с лицами людей. Такой подход позволяет находить похожие наборы данных из разных источников, расширяя детектирование признаков и классов и не нанося серьезного вреда балансировке. Для каждого объекта датасета получено векторное представление (эмбеддинг), затем выполнено сравнение эмбеддингов в обоих датасетах. Эксперименты проводились на примере датасетов с изображениями лиц людей. Для получения эмбеддингов использовалась предобученная сеть ResNet. В процессе исследований один датасет делился на две части, представляющие собой схожие датасеты, затем каждая из частей сравнивалась с отличающимся набором данных. Предлагается новая метрика сходства, которая обладает рядом преимуществ и позволяет находить наиболее похожие датасеты.

Ключевые слова: набор данных, векторное представление, ResNet, сходство датасетов, глубокое обучение.

Конфликт интересов. Авторы заявляют об отсутствии конфликта интересов.

Благодарность. Авторы выражают благодарность Полине Игоревне Тишковской за помощь в оформлении статьи.

Для цитирования. Усатов, А. А. Оценка сходства между наборами данных с помощью векторных представлений / А. А. Усатов, А. М. Недзьведь, Го Цзижань // Доклады БГУИР. 2025. Т. 23, № 3. С. 70–76. http://dx.doi.org/10.35596/1729-7648-2025-23-3-70-76.

ASSESSING SIMILARITY BETWEEN DATASETS USING VECTOR REPRESENTATIONS

ALEXANDER A. USATOFF, ALEXADER M. NEDZVED, GUO JIRAN

Belarusian State University (Minsk, Republic of Belarus)

Abstract. The article considers an approach to determining the similarity of datasets for training algorithms using datasets with human faces as an example. This approach allows finding similar datasets from different sources, expanding the detection of features and classes and significantly affecting dataset balance. For each dataset object, a vector representation (embedding) was obtained, then the embeddings in both datasets were compared. The experiments were conducted using datasets with images of human faces as an example. To obtain embeddings, a pre-trained ResNet network was used. During the research, one dataset was divided into two parts, which were similar datasets, then each of the parts was compared with a different dataset. The new similarity metric is proposed, which has several advantages and allows to find the most similar datasets.

Keywords: dataset, vector representation, ResNet, dataset similarity, deep learning.

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

Gratitude. The authors express their gratitude to Polina Igorevna Tishkovskaya for her help in the design of the article.

For citation. Usatoff A. A., Nedzved A. M., Guo Jiran (2025) Assessing Similarity Between Datasets Using Vector Representations. *Doklady BGUIR*. 23 (3), 70–76. http://dx.doi.org/10.35596/1729-7648-2025-23-3-70-76 (in Russian).

Введение

Для решения задач в глубоком обучении [1, 2] необходим большой объем данных. Обучающая выборка должна быть внушительной, чтобы модель определила зависимости и признаки детектируемых объектов и при этом не переобучилась. Как правило, в случае не очень больших наборов данных (датасетов) используется лишь дообучение уже готовой модели, для этого может быть достаточно нескольких тысяч объектов. Тем не менее в случае изображений собрать и разметить даже такой объем данных может быть проблематично. Поэтому сеть часто обучают на публичном датасете, а свой небольшой набор данных используют лишь для валидации модели. Однако, чем меньше публичный набор данных похож на то, что модель будет обрабатывать на практике, тем хуже будет качество модели в реальных условиях.

Совмещение различных датасетов в машинном обучении представляет собой сложную задачу, так как возникает множество факторов, влияющих на качество модели и точность прогнозов. Основная проблема заключается в том, что данные из разных источников часто имеют различия в структуре, формате или распределении признаков. Это может привести к тому, что модель будет обучаться на непоследовательных или даже противоречивых примерах. Различия в масштабах и единицах измерения между данными из разных источников также создают проблемы. Если одни признаки представлены в больших численных значениях, а другие – в меньших, это может исказить важность таких признаков для модели. Даже после нормализации или стандартизации различия в характере данных могут сохраняться, особенно, если они были собраны с использованием разных методологий или инструментов. Это может привести к появлению шума в данных или усилению некоторых биасов, которые снижают обобщающую способность модели.

Еще одна трудность связана с тем, что различные датасеты могут содержать разные уровни полноты или иметь разные стратегии обработки пропущенных значений. В одном наборе данных могут быть удалены записи с пропусками, тогда как в другом такие записи могли быть заполнены средними значениями или другими методами. Это может создать искусственные различия между данными, которые не отражают реальной картины. Таким образом, совмещение данных требует тщательной предварительной обработки, чтобы минимизировать подобные искажения и обеспечить согласованность между источниками информации. Одно из решений перечисленных проблем – нахождение похожих датасетов. Но для этого необходимо определять некую меру сходства наборов данных, чтобы использовать для обучения тот набор, который ближе всего к реальным данным, с которыми модель будет работать на практике.

Цель исследований авторов – повышение эффективности определения сходства датасетов с использованием векторных представлений (эмбеддингов) – объектов фиксированной и, как правило, относительно невысокой размерности, которые содержат наиболее важную информацию об исходных данных. Эмбеддинги несут в себе определенный «смысл» объекта, а близкие эмбеддинги означают, что объекты похожи между собой [3]. Эксперименты проводились на примере датасетов с изображениями лиц людей.

Определение сходства изображений

Математическое определение сходства изображений основывается на вычислении расстояния или сходства между их представлениями в численном пространстве. Один из классических подходов – использование метрик, таких как среднеквадратичная разность (MSE) или коэффициент корреляции Пирсона, которые сравнивают пиксельные значения двух изображений напрямую. Например, MSE вычисляется как среднее значение квадратов разностей между соответствующими пикселями двух изображений, что позволяет количественно оценить различия [4]. Однако такие методы имеют существенные ограничения: они чувствительны к небольшим изменениям, не учитывают более сложные семантические характеристики изображений и игнорируют структурные особенности объектов [5]. Другим подходом является использование гистограмм цветов или текстур для сравнения изображений. В этом случае каждое изображение представляется в виде распределения значений, характеризующих его цветовые или текстурные свойства. Затем применяются метрики, такие как расстояние Хэмминга или Кульбака – Лейблера между двумя распределениями [6]. В данном случае оценка более устойчива к некоторым преобразованиям, но ограничена в восприятии высокоуровневых признаков, включая форму объектов. Кроме того, она чувствительна к шуму или изменению освещенности [7].

Использование эмбеддингов для сравнения изображений – современный и эффективный подход. Векторные представления отражают высокие уровни абстракции, связанные с содержанием изображений, и позволяют сравнивать их с использованием таких метрик, как евклидово расстояние или косинусная близость [8].

Особенности применения векторных представлений

При нейросетевом методе получения эмбеддингов для изображений обычно используются сверточные нейронные сети, принцип устройства которых продемонстрирован на рис. 1.

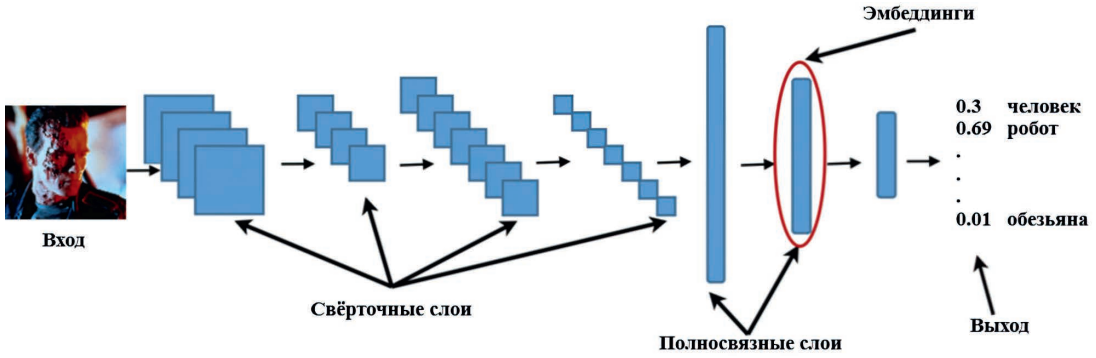


Рис. 1. Получение векторных представлений с помощью нейронной сети
Fig. 1. Obtaining vector representations using a neural network

Сверточные слои являются ключевыми для сверточной нейронной сети. Они выполняют операцию свертки, применяя фильтры (ядра) к исходному изображению, чтобы выделить важные признаки. Эти признаки могут включать края, углы, текстуры и другие визуальные элементы. После применения фильтров результаты передаются через функции активации (например, ReLU) и слои субдискретизации, которые уменьшают размерность и обобщают признаки. Обычно применяется несколько сверточных слоев, где каждый последующий извлекает признаки более высокого уровня. Например, на первом сверточном слое могут извлекаться такие признаки, как границы, углы, текстуры, на втором – более сложные формы – изгибы бровей или колеса автомобилей, на третьем сверточном слое могут выявляться уже конкретные объекты или их крупные части. За сверточными слоями следуют полносвязные слои, которые выполняют классификацию на основе извлеченных сверточными слоями признаков. Последний полносвязный слой обычно содержит столько нейронов, сколько классов в задаче классификации, а функция активации softmax преобразует выходные данные в вероятности каждого класса, что и является выходом нейронной сети. Эмбеддингами обычно являются выходы одного из полносвязных слоев сети. Эти признаки затем сравниваются с помощью таких метрик, как косинусное сходство или манхэттенское расстояние, чтобы определить степень их схожести [9].

Воспользуемся вышеупомянутым свойством, что похожие изображения имеют близкие эмбеддинги, чтобы считать сходство изображений. Будем считать сходство изображений как косинусное сходство между их эмбеддингами согласно формуле

$$cosine(A,B) = \frac{\sum_{i=1}^{n} A_i B_i}{\sqrt{\sum_{i=1}^{n} A_i^2}\sqrt{\sum_{i=1}^{n} B_i^2}}, \quad (1)$$

где A, B – векторы, между которыми считается косинусное сходство; n – размер векторов A и B.

Формула (1) – один из самых распространенных способов вычисления близости эмбеддингов, который используется, например, в модели word2vec [10]. Для определения сходства датасетов проводилось сравнение содержащихся в них элементов и их распределения:

– для каждого объекта из первого датасета находили самый близкий объект во втором датасете, затем наоборот;

– для каждого объекта первого датасета считали близость со всеми объектами второго датасета, затем усредняли.

Первый вариант не учитывает плотность распределения объектов в признаковом пространстве, поэтому для оценки сходства использовалось второе определение, где оценка выполнялась по формуле

$$datasets_similarity(D1, D2) = \frac{\sum_{i=1}^{n}\sum_{j=1}^{m} cosine(D1_i, D2_j)}{nm}, \quad (2)$$

где $D1$, $D2$ – датасеты, сходство которых вычислялось; *cosine* – косинусное сходство, описанное в (1); n, m – размер датасетов $D1$ и $D2$ соответственно.

Такой подход имеет важное преимущество: эмбеддинги учитывают не только низкоуровневые признаки, но и высокоуровневые семантические характеристики, что делает их гораздо более релевантными для оценки человеческой перцепции сходства. Исследования показывают, что эмбеддинги, полученные с помощью предобученных CNN, демонстрируют высокую корреляцию с результатами тестов, основанных на субъективной оценке людей [11].

Использование объектов при сравнении наборов данных

Для проверки работоспособности предлагаемого подхода оценки сходства между датасетами проводились эксперименты на датасетах с фотографиями лиц людей. Чтобы получить эмбеддинги таких изображений, использовалась предобученная сеть ResNet18 [12] без слоя классификации. В экспериментах применялось три датасета: два – очень близкие, а третий несколько от них отличался. Такая организация данных позволила проверить адекватность оценки на основе эмбеддингов, поскольку сходство между первыми двумя датасетами было больше, чем их сходство с третьим. Для получения двух максимально схожих между собой наборов данных один датасет разбивался на два тестовых набора данных.

Поскольку в датасетах для глубокого обучения обычно минимум десятки тысяч изображений, а для статистической значимости результатов необходимо провести эксперименты много раз, применялись подвыборки размером 100 изображений. Кроме того, поиск для каждого изображения наиболее близкого выполнялся квадратичное время, что затрудняло использование больших подвыборок. Для эксперимента наборы данных с идентичным содержанием, включающим объекты одного типа в одинаковой проекции и при одинаковых условиях съемки, определялись как очень похожие датасеты, а наборы данных объектов одного типа, но в разных условиях съемки и в разном положении, – как похожие датасеты. Эксперименты проводились следующим образом:

1) использовалось два датасета – $D1$ и $D2$. Датасет $D1$ делился на две части $D1_1$ и $D1_2$, которые представляли собой очень похожие наборы данных. Датасет $D2$ имел небольшие отличия от $D1$ (ниже приведены примеры и описание отличий для конкретных датасетов, но в общем случае это не принципиально);

2) далее 100 раз повторялись следующие действия:

a) из датасетов $D1_1$, $D1_2$ и $D2$ брались выборки размером 100 элементов – $d1_1$, $d1_2$ и $d2$ соответственно;

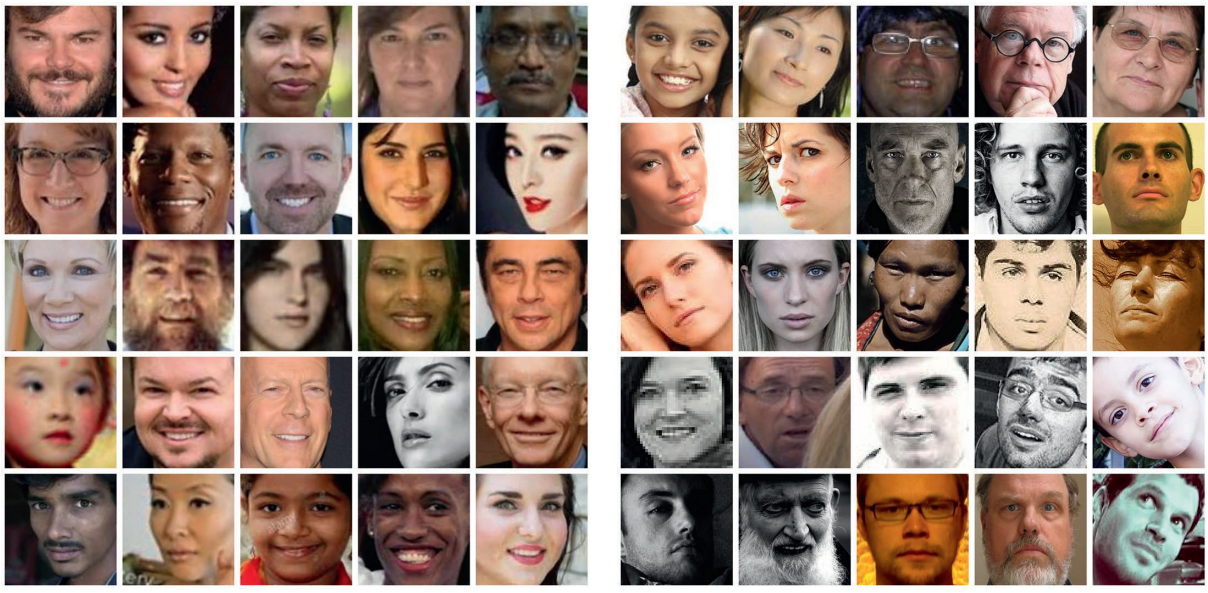
b) по формуле (2) считалось сходство между выборками $\{d1_1, d1_2\}$, $\{d1_1, d2\}$ и $\{d1_2, d2\}$;

3) по результатам вычислений по пункту 2b строились гистограммы сходства между всеми датасетами (точнее, их оценки по выборкам) и вычислялась точность определения более похожего датасета.

На рис. 2, *a* показан фрагмент набора изображений для классификации лиц. Этот набор делился на два, представлявшие собой очень похожие датасеты. В качестве не очень похожего датасета были использованы картинки из набора данных для определения ключевых точек лица (рис. 2, *b*), где присутствуют различные искажения – контрастные тени, засветки, повернутые изображения и т. д. На рис. 3 приведено распределение косинусной меры сходства для различных датасетов.

В ходе экспериментов рассчитанное в соответствии с предложенным подходом сходство между датасетами $D1_1$ и $D1_2$ всегда было больше, чем сходство между $D1_1$ и $D2$ и датасетами $D1_2$ и $D2$, то есть:

$$\forall d1_1 \in D1_1, d1_2 \in D1_2, d2 \in D2$$
$$datasets_similarity(d1_1, d1_2) > datasets_similarity(d1_1, d2)$$
$$datasets_similarity(d1_1, d1_2) > datasets_similarity(d1_2, d2).$$



a *b*

Рис. 2. Примеры изображений: *a* – датасетов, которые делились на два; *b* – несколько отличающегося датасета

Fig. 2. Examples of images: *a* – datasets that were divided into two; *b* – several different datasets

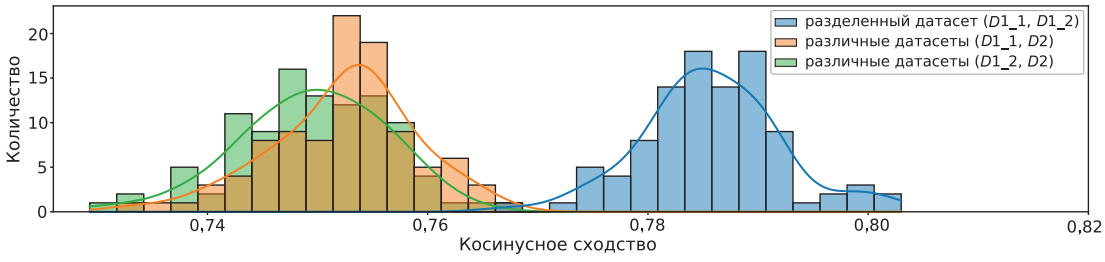


Рис. 3. Распределение косинусной меры сходства для различных датасетов

Fig. 3. Distribution of cosine similarity measure for different datasets

Одной из популярных метрик определения сходства изображений является MSE. Предложенный в статье подход превосходит метод оценки расстояния между датасетами на основе MSE. Для уменьшения влияния случайностей при сравнении методов эксперименты с применением MSE выполнялись на тех же подвыборках изображений, что и с использованием эмбеддингов. На рис. 4 приведено распределение среднеквадратичной разности для различных датасетов.

Из рис. 4 видно, что в части случаев оценка с использованием MSE дает неверный результат. Кроме того, оценка получается неустойчивой, поскольку датасеты *D*1_1, *D*1_2 являются частями одного и того же датасета, но их расстояния от *D*2 сильно отличаются, чего не наблюдается при оценке сходства с использованием эмбеддингов. При использовании MSE точность определения более схожего датасета составила 0,975, в то время как применение формулы (2) позволило полностью исключить ошибки, обеспечив точность, равную единице. Это подтверждает эффективность предложенной метрики в сравнении наборов данных

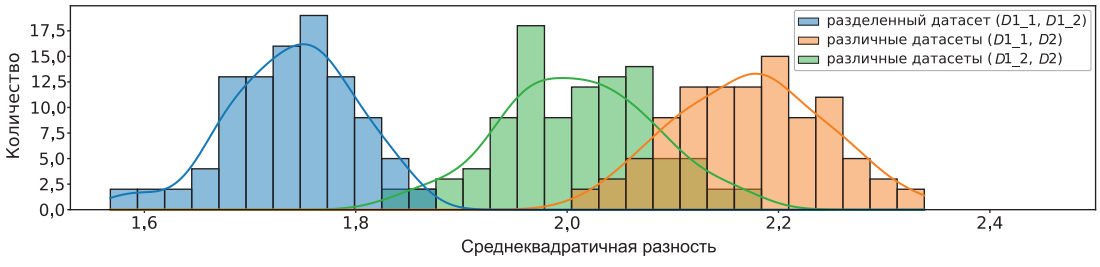


Рис. 4. Распределение MSE для различных датасетов

Fig. 4. Distribution of MSE for different datasets

Заключение

1. Рассмотрен подход к определению сходства набора данных (датасетов) на примере датасетов с изображениями лиц людей. Для получения векторных представлений использовалась предобученная сеть ResNet18 без слоя классификации. Предложена эффективная метрика сходства датасетов, которая позволяет определять меру сходства наборов данных и выбирать датасет, наиболее близкий к образцу.

2. Качество оценки, основанной на векторных представлениях, зависит от архитектуры модели и данных, на которых она была обучена, особенно если модель обучалась на ограниченном наборе категорий [13]. Однако использование векторных представлений остается одним из самых перспективных направлений для решения задачи определения сходства изображений благодаря способности выявлять сложные паттерны и отношения между объектами.

3. Для специфических данных, таких как снимки с дрона или медицинские изображения, имеет смысл использовать дообученную сеть из той же области, поскольку датасет ImageNet [14], на котором была обучена сеть ResNet, во многом отличается от узкоспециализированных наборов данных [15]. Предложенную метрику можно использовать не только для определения сходства датасетов, но и, например, для расширения набора данных путем добавления в него похожих объектов из других датасетов.

Список литературы

1. Ивахненко, А. Г. Кибернетические предсказывающие устройства / А. Г. Ивахненко, В. Г. Лапа. Киев: Акад. наук Укр. ССР, 1965.
2. Gradient-Based Learning Applied to Document Recognition / Y. Lecun [et al.] // Proceedings of the IEEE. 1998. Vol. 86, Iss. 11. P. 2278–2324.
3. Label-Embedding for Image Classification / Z. Akata [et al.] // IEEE Transactions on Pattern Analysis and Machine Intelligence. 2015. Vol. 38, No 7. P. 1425–1438. DOI: 10.1109/TPAMI.2015.2487986.
4. Image Quality Assessment: From Error Visibility to Structural Similarity / Z. Wang [et al.] // IEEE Transactions on Image Processing. 2024. Vol. 13, No 4. P. 600–612. DOI: 10.1109/TIP.2003.819861.
5. Rubner, Y. The Earth Mover's Distance as a Metric for Image Retrieval / Y. Rubner, C. Tomasi, L. J. Guibas // International Journal of Computer Vision. 2000. Vol. 40, No 2. P. 99–121. DOI: 10.1023/A:1026543900054.
6. Lin, J. Divergence Measures Based on the Shannon Entropy / J. Lin // IEEE Transactions on Information Theory. 1991. Vol. 37, Iss. 1. P. 145–151. DOI: 10.1109/18.61115.
7. Swain, M. J. Color Indexing / M. J. Swain, D. H. Ballard // International Journal of Computer Vision. 1991. Vol. 7, No 1. P. 11–32.
8. Simonyan, K. Very Deep Convolutional Networks for Large-Scale Image Recognition / K. Simonyan, A. Zisserman // arXiv.1409.1556. 2014. Vol. 1.
9. Self-Similarity Guided Probabilistic Embedding Matching Based on Transformer for Occluded Person Re-Identification / Y. Pang [et al.] // Expert Systems with Applications. 2024. Vol. 237. https://doi.org/10.1016/j.eswa.2023.121504.
10. Efficient Estimation of Word Representations in Vector Space / T. Mikolov [et al.] // arXiv:1301.3781. 2013. http://arxiv.org/abs/1301.3781.
11. The Unreasonable Effectiveness of Deep Features as a Perceptual Metric / R. Zhang [et al.] // arXiv:1801.03924. 2023. https://doi.org/10.48550/arXiv.1801.03924.
12. Deep Residual Learning Forimage Recognition / K. He [et al.] // arXiv:1512.03385. 2015. https://doi.org/10.48550/arXiv.1512.03385.
13. Learning Transferable Visual Models from Natural Language Supervision / A. Radford [et al.] // arXiv:2103.00020. 2021. https://doi.org/10.48550/arXiv.2103.00020.
14. Imagenet: A Large-Scale Hierarchicalimage Database / Jia Deng [et al.] // 2009 IEEE Conference on Computer Vision and Pattern Recognition. P. 248–255.
15. Недзьведь, А. М. Анализ изображений для решения задач медицинской диагностики / А. М. Недзьведь, С. В. Абламейко. Минск: Объедин. ин-т проблем информ. Нац. акад. наук Беларуси, 2012.

Поступила 10.12.2024 Принята в печать 25.04.2025

References

1. Ivakhnenko A. G., Lapa V. G. (1965) *Cybernetic Predictive Devices*. Kyiv, Academy of Sciences of the Ukrainian SSR (in Russian).
2. Lecun Y., Bottou L., Bengio Y., Haffner P. (1998) Gradient-Based Learning Applied to Document Recognition. *Proceedings of the IEEE*. 86 (11), 2278–2324.
3. Akata Z., Perronnin F., Harchaoui Z., Schmid C. (2015) Label-Embedding for Image Classification. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 38 (7), 1425–1438. DOI: 10.1109/TPAMI.2015.2487986.

4. Wang Z., Bovik A. C., Sheikh H. R., Simoncelli E. P. (2024) Image Quality Assessment: From Error Visibility to Structural Similarity. *IEEE Transactions on Image Processing*. 13 (4), 600–612. DOI: 10.1109/TIP.2003.819861.
5. Rubner Y., Tomasi C., Guibas L. J. (2000) The Earth Mover's Distance as a Metric for Image Retrieval. *International Journal of Computer Vision.* 40 (2), 99–121. DOI: 10.1023/A:1026543900054.
6. Lin J. (1991) Divergence Measures Based on the Shannon Entropy. *IEEE Transactions on Information Theory*. 37 (1), 145–151. DOI: 10.1109/18.61115.
7. Swain M. J., Ballard D. H. (1991) Color Indexing. *International Journal of Computer Vision*. 7 (1), 11–32.
8. Simonyan K., Zisserman A. (2014) Very Deep Convolutional Networks for Large-Scale Image Recognition. *arXiv.1409.1556*. 1.
9. Pang Y., Zhang H., Zhu L., Liu D., Liu L. (2024) Self-Similarity Guided Probabilistic Embedding Matching Based on Transformer for Occluded Person Re-Identification. *Expert Systems with Applications*. 237. https://doi.org/10.1016/j.eswa.2023.121504.
10. Mikolov T., Chen K., Corrado G., Dean J. (2013) Efficient Estimation of Word Representations in Vector Space. *arXiv:1301.3781*. http://arxiv.org/abs/1301.3781.
11. Zhang R., Isola P., Efros A. A., Shechtman E., Wang O. (2023) The Unreasonable Effectiveness of Deep Features as a Perceptual Metric. *arXiv:1801.03924*. https://doi.org/10.48550/arXiv.1801.03924.
12. He K., Zhang X., Ren S., Sun J. (2015) Deep Residual Learning for Image Recognition. *arXiv:1512.03385*. https://doi.org/10.48550/arXiv.1512.03385.
13. Radford A., Kim J. W., Hallacy C., Ramesh A., Goh G., Agarwal S., et al. (2021) Learning Transferable Visual Models from Natural Language Supervision. *arXiv:2103.00020*. https://doi.org/10.48550/arXiv.2103.00020.
14. Jia Deng, Wei Dong, Richard Socher, Li-Jia Li, Kai Li, Li Fei-Fei, et al. (2009) ImageNet: A Large-Scale Hierarchical Image Database. *2009 IEEE Conference on Computer Vision and Pattern Recognition*. 248–255.
15. Nedzved A., Ablameyko S. (2012) *Image Analysis for Tasks of Medical Diagnostic*. Minsk, United Institute of Informatics Problems of the National Academy of Sciences of Belarus (in Russian).

Received: 10 December 2024 Accepted: 25 April 2025

Вклад авторов

Усатов А. А. провел эксперименты и подготовил рукопись статьи.

Недзьведь А. М. осуществил постановку задачи для проведения исследования, дал рекомендации по выполнению экспериментов и написанию статьи.

Го Цзижань предложил возможные применения предложенного подхода, принимал участие в подготовке рукописи статьи.

Authors' contribution

Usatoff A. A. conducted experiments and prepared the manuscript of the article.

Nedzved A. M. set the task for the research, gave recommendations on conducting the experiments, and writing the article.

Guo Jiran suggested possible applications of the proposed approach, participated in the preparation of the manuscript of the article.

Сведения об авторах

Усатов А. А., магистр, асп. каф. информационных систем управления, Белорусский государственный университет (БГУ)

Недзьведь А. М., д-р техн. наук, доц., зав. каф. информационных систем управления, БГУ

Го Цзижань, асп. каф. информационных систем управления, БГУ

Адрес для корреспонденции

220030, Республика Беларусь,
Минск, просп. Независимости, 4, к. 501
Белорусский государственный университет
Тел.: +375 17 209-52-45
E-mail: alexander.usatoff@gmail.com
Усатов Александр Андреевич

Information about the authors

Usatoff A. A., Master's, Postgraduate at the, Department of Information Management Systems, Belarusian State University (BSU)

Nedzved A. M., Dr. Sci. (Tech.), Associate Professor, Head of the Department of Information Management Systems, BSU

Guo Jiran, Postgraduate at the Department of Information Management Systems, BSU

Address for correspondence

220030, Republic of Belarus,
Minsk, Nezavisimosti Ave., 4, Off. 501
Belarusian State University
Tel.: +375 17 209-52-45
E-mail: alexander.usatoff@gmail.com
Usatoff Alexander Andreevich

CC BY

http://dx.doi.org/10.35596/1729-7648-2025-23-3-77-85

УДК [621.384.64:539.1.089.6]:615.849.1

МАТЕМАТИЧЕСКАЯ МОДЕЛЬ РЕСПИРАТОРНЫХ ИСКАЖЕНИЙ ПАТОЛОГИЧЕСКОГО ОЧАГА ПРИ ВИЗУАЛИЗАЦИИ НА ПЭТ/КТ-ИЗОБРАЖЕНИЯХ

М. Н. ПЕТКЕВИЧ[1], Д. В. КОРОВКО[2], Е. В. ЕМЕЛЬЯНЕНКО[1]

[1]Республиканский научно-практический центр онкологии и медицинской радиологии имени Н. Н. Александрова (Минский район, Республика Беларусь)
[2]Международный государственный экологический институт имени А. Д. Сахарова Белорусского государственного университета (Минск, Республика Беларусь)

Аннотация. Проведены оценка и анализ неопределенностей визуализации патологического очага на ПЭТ/КТ-изображениях, возникающих вследствие респираторных движений биологического объекта, оказывающих непосредственное влияние на определение геометрических характеристик патологического очага, его локализацию и корректность моделирования трехмерного распределения дозы излучения в теле пациента. На основании экспериментально установленных зависимостей влияния величины смещения и диаметра исследуемого объекта (сферы) на величину несоответствия визуализированного объема исследуемого объекта разработана математическая модель респираторных искажений патологического очага при визуализации на ПЭТ/КТ-изображениях, позволяющая осуществить количественную и качественную оценку влияния респираторного движения пациента на геометрическую точность визуализации патологического очага на ПЭТ/КТ-изображениях с точностью до 98 %. Верификация модели показала высокую степень согласованности между модельными расчетами и экспериментальными данными. Средние значения относительной неопределенности для КТ и ПЭТ, соответственно, составили (2,123 ± 1,051) % и (2,661 ± 0,870) %, а для абсолютной – (2,096 ± 0,941) % и (1,992 ± 0,782) %, что подтверждает корректность и практическую применимость разработанной методики.

Ключевые слова: математическая модель, искажения, ПЭТ/КТ, фантом, величина несоответствия, респираторные движения, лучевая терапия.

Конфликт интересов. Авторы заявляют об отсутствии конфликта интересов.

Для цитирования. Петкевич, М. Н. Математическая модель респираторных искажений патологического очага при визуализации на ПЭТ/КТ-изображениях / М. Н. Петкевич, Д. В. Коровко, Е. В. Емельяненко // Доклады БГУИР. 2025. Т. 23, № 3. С. 77–85. http://dx.doi.org/10.35596/1729-7648-2025-23-3-77-85.

MATHEMATICAL MODEL OF RESPIRATORY DISTORTIONS OF A PATHOLOGICAL FOCUS DURING VISUALIZATION ON PET/CT IMAGES

MAKSIM N. PIATKEVICH[1], DMITRIY V. KOROVKO[2], EVGENIY V. EMELYANENKO[1]

[1]N. N. Aleksandrov National Cancer Center of Belarus (Minsk Region, Republic of Belarus)
[2]International Sakharov Environmental Institute of Belarusian State University (Minsk, Republic of Belarus)

Abstract. The article presents an assessment and analysis of uncertainties in pathological lesion visualization on PET/CT images that arise due to respiratory movements of a biological object, which have a direct impact on determining the geometric characteristics of the pathological lesion, its localization, and the correctness of modeling the three-dimensional distribution of the radiation dose in the patient's body. Based on the experimentally established dependencies of the influence of the displacement value and the diameter of the object (sphere) under study on the value of the discrepancy between the visualized volume of the object under study, a mathematical model of respiratory distortions of the pathological lesion during visualization on PET/CT images has been de-

veloped, which allows for a quantitative and qualitative assessment of the influence of the patient's respiratory movement on the geometric accuracy of pathological lesion visualization on PET/CT images with up to 98 % accuracy. Verification of the model has shown a high degree of consistency between the model calculations and experimental data. The average values of relative uncertainty for CT and PET, respectively, were (2.123 ± 1.051) % and (2.661 ± 0.870) %, and for absolute uncertainty – (2.096 ± 0.941) % and (1.992 ± 0.782) %, which confirms the correctness and practical applicability of the developed method.

Keywords mathematical model, distortion, PET/CT, phantom, uncertainties value, respiratory motion, radiation therapy.

Conflict of interests. The author declares no conflict of interests.

For citation. Piatkevich M. N., Korovko D. V., Emelyanenko E. V. (2025) Mathematical Model of Respiratory Distortions of a Pathological Focus During Visualization on PET/CT Images. *Doklady BGUIR*. 23 (3), 77–85. http://dx.doi.org/10.35596/1729-7648-2025-23-3-77-85 (in Russian).

Введение

Компьютерное и математическое моделирование представляет собой метод исследования сложных систем посредством построения и анализа ее компьютерной модели. Основная цель моделирования заключается в получении неизвестных ранее свойств сложной системы, а также в прогнозировании и объяснении различных ее состояний. В частности, компьютерное и математическое моделирование в лучевой терапии направлено на изучение физиологических процессов в организме человека в норме и патологии, а также на оценку воздействия ионизирующего излучения на паталогический очаг и здоровые ткани с помощью компьютерных моделей.

Благодаря эволюции методов моделирования возможно создать модель распределения дозы в теле пациента с высокой степенью соответствия геометрическим характеристикам целевого объема облучения, однако определение самого целевого объема является сложной и нетривиальной задачей [1]. Каждое поле облучения должно быть сформировано таким образом, чтобы оно охватывало патологический очаг с учетом всех его перемещений, и при этом объем здоровой ткани, попадающий в геометрические границы поля облучения, должен быть сведен к минимуму [2]. Особое внимание уделяется локализациям грудной и брюшной полостей в связи с физиологической подвижностью органов в процессе дыхания человека. Движения внутренних органов приводят к возникновению артефактов и неточностям в получаемой компьютерной анатомической информации о пациенте.

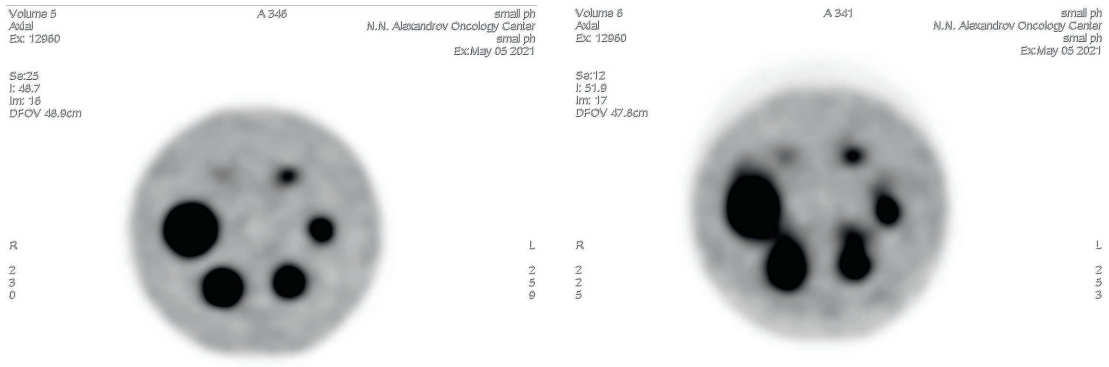
В качестве стандарта при моделировании дозового распределения принято применение КТ-изображений как исходной диагностической информации [3]. Без измерения движения в реальном времени моделирование распределения дозы ионизирующего излучения на основе компьютерной томографии предполагает, что изображение, полученное при свободном дыхании, является статичным и репрезентативным для среднего положения целевого объема в течение временного интервала, в котором проводится облучение [4]. Поскольку КТ-изображение – это только снимок в определенный момент времени, каждый 2D-участок паталогического очага может быть зафиксирован в любой точке диапазона перемещения во время дыхательного цикла, а результирующие изображения могут отображать некорректную информацию о положении или геометрической форме паталогического очага (эффект частичного объема).

Внедрение в клиническую практику РНПЦ ОМР имени Н. Н. Александрова ПЭТ/КТ позволило значительно повысить качество получаемой диагностической информации. Преимущество данного метода, по сравнению с классической компьютерной томографией, заключается в возможности получить совмещенные ПЭТ- и КТ-изображения, позволяющие оценить топологию патологического очага накопления радиофармацевтического препарата и его метаболическую активность. Вместе с тем, согласно результатам исследования [5], применение ПЭТ/КТ-изображений для предлучевой подготовки пациента при динамической лучевой терапии требует учета респираторных искажений, возникающих вследствие респираторных движений паталогического очага. Под респираторными искажениями понимается возникновение артефактов на ПЭТ/КТ-изображениях, вызванных движением органов и тканей грудной и брюшной полостей во время респираторного цикла биологического объекта, приводящих к ухудшению качества визуализации и точности определения геометрических параметров исследуемых объектов.

В связи с вышесказанным, определение индивидуальных для каждого клинического случая параметров движения паталогического очага при получении компьютерной объемно-анатомической информации на ПЭТ/КТ является одним из приоритетных направлений к снижению уровня облучения нормальных тканей и корректному облучению мишени. В процессе исследований разработана математическая модель респираторных искажений паталогического очага при визуализации на ПЭТ/КТ-изображениях. Выполнена количественная и качественная оценка влияния респираторного движения на геометрическую точность визуализации паталогического очага на ПЭТ/КТ-изображениях.

Методика проведения эксперимента

В [5] представлена установка для оценки количественных характеристик ПЭТ-изображений при исследовании динамических объектов, с помощью которой обнаружено несоответствие в визуализации геометрических параметров паталогического очага с учетом его перемещений во время респираторного цикла пациента. Установка имитирует тело пациента (фантом) с расположенными в нем патологическими очагами (сферами) различного диаметра (10, 13, 17, 22, 28 и 37 мм) и приводится в движение мотором-редуктором с валом с кулачком. Фантом и сферы наполняются радиофармацевтическим препаратом, обеспечивая отношение фоновой удельной активности в фантоме к удельной активности в каждой сфере, равное 1:3. Таким образом, установка имитирует тело пациента с патологическими очагами, находящимися в движении, аналогичными при дыхании человека. Установка сканируется на ПЭТ/КТ-томографе с различными начальными данными (амплитуда движения, диаметр исследуемой сферы) для дальнейшего анализа несоответствий в геометрических параметрах патологического очага, визуализированного на ПЭТ/КТ-изображениях, истинным значениям. На рис. 1 приведены примеры визуализации фантома на ПЭТ/КТ-изображениях в статическом и динамическом состояниях.



a *b*

Рис. 1. Визуализация фантома по ПЭТ-изображениям в динамическом (*a*) и статическом (*b*) состояниях
Fig. 1. Visualization of the phantom using PET images in dynamic (*a*) and static (*b*) states

Оценка геометрических характеристик исследуемых сфер в статическом режиме при ПЭТ-визуализации показала несоответствие истинным геометрическим характеристикам до 4 % в зависимости от размеров моделируемых патологических очагов. Наибольшее несоответствие определено при визуализации сферы диаметром 10 мм. Для выполнения сравнительного анализа исследование проводилось для сфер, визуализированных с применением метода компьютерной томографии. При КТ-исследовании наибольшее несоответствие объема – 6,5 %.

Аналогичные операции проводились для ПЭТ- и КТ-изображений, полученных в динамическом режиме. Оценка геометрических характеристик исследуемых сфер в динамическом режиме по КТ-сериям показала несоответствие истинным характеристикам до 52 % в зависимости от размеров моделируемых патологических очагов. Наибольшее несоответствие определено при визуализации сферы диаметром 10 мм. Оценка геометрических характеристик исследуемых сфер в динамическом режиме по ПЭТ-сериям показала несоответствие истинным характеристикам до 27 % в зависимости от размеров моделируемых патологических очагов. Наибольшее несоответствие определено при визуализации сферы диаметром 10 мм.

По полученным в результате экспериментального исследования данным установлены зависимости влияния величины смещения (рис. 2, *a*) и диаметра (рис. 2, *b*) исследуемого объекта (сферы) на величину несоответствия визуализированного объема исследуемого объекта.

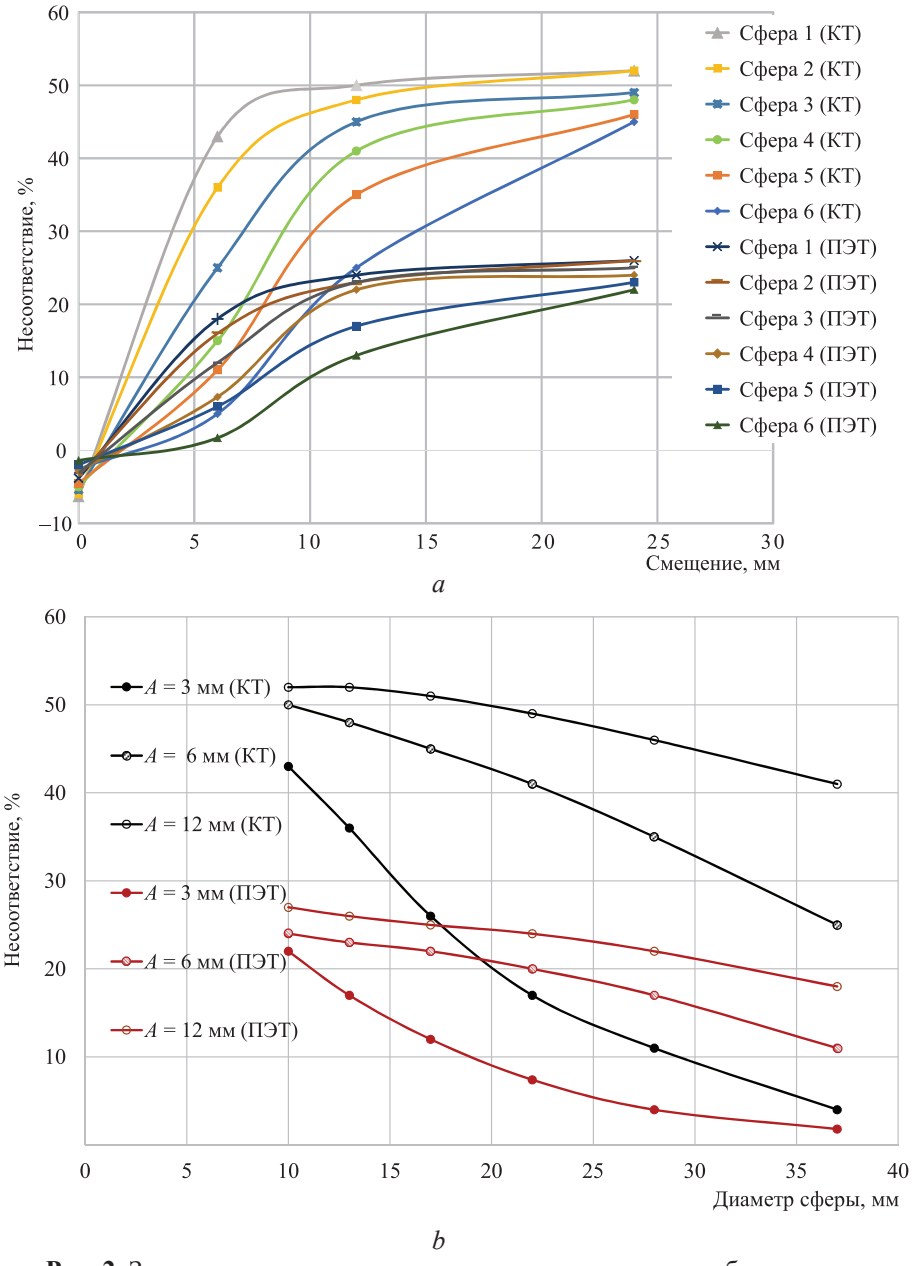


Рис. 2. Зависимость величины несоответствия истинному объему от:
a – смещения движущейся модели; *b* – диаметра сферы; *A* – амплитуда смещения сферы
Fig. 2. Dependence of the magnitude of the discrepancy with the true volume on:
a – displacement of the moving model; *b* – diameter of the sphere; *A* – amplitude of the sphere displacement

Недостатком установки является то, что она оценивает ограниченное количество вариаций патологического очага (имеет ограниченный набор сфер и их перемещений), что не позволяет определить несоответствия визуализированных объектов, находящихся в движении, для всех возможных клинических случаев. В связи с чем и с учетом полученных экспериментальных данных разработана математическая модель респираторных искажений патологического очага при визуализации на ПЭТ/КТ-изображениях.

Разработка математической модели

Модель разработана в среде программирования MATLAB R2023b Update 4 (23.2.0.2428915) и реализована в виде единого м-файла. При запуске модели активируется блок симуляции с гра-

фическим интерфейсом, где пользователь задает исходные параметры. Принцип работы блока симуляции представлен на схеме (рис. 3).

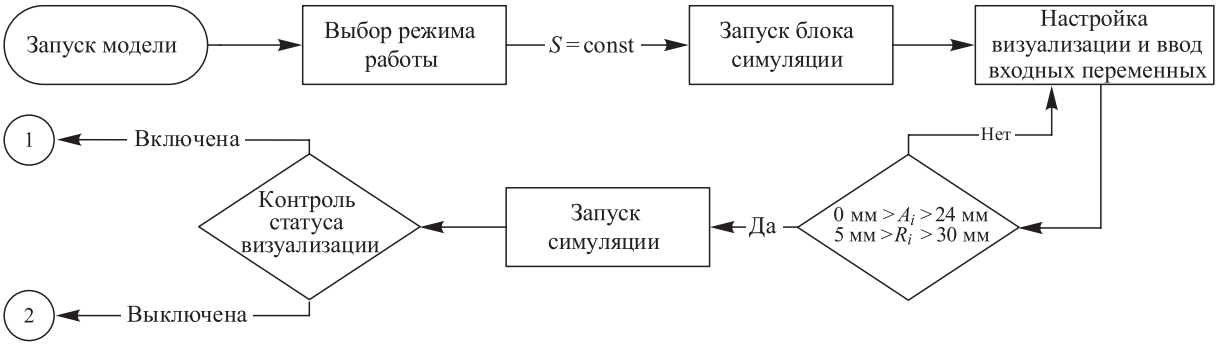


Рис. 3. Схема работы блока симуляции
Fig. 3. Simulation unit operation diagram

Интерфейс поддерживает два режима: фиксированная амплитуда (A_i) с изменяющимся радиусом сферы и фиксированный радиус (R_i) с варьированием амплитуды. Ввод параметров осуществляется с возможностью динамического добавления/удаления полей ввода, а также посредством переключателя «Включить визуализацию». При нажатии кнопки «Старт» модель проверяет корректность введенных данных, уведомляя пользователя об ошибках для их оперативного исправления. При отсутствии ошибок симуляция продолжается с учетом флага визуализации.

Симуляция представляет собой комплексную анимацию, состоящую из нескольких этапов, отображаемых на отдельных графиках в одном окне. Визуализационный блок имитирует динамику движения сферы в условиях дыхания и накапливает сигналы для моделирования ПЭТ- и КТ-изображений. Схема работы визуализационного блока показана на рис. 4 [6].

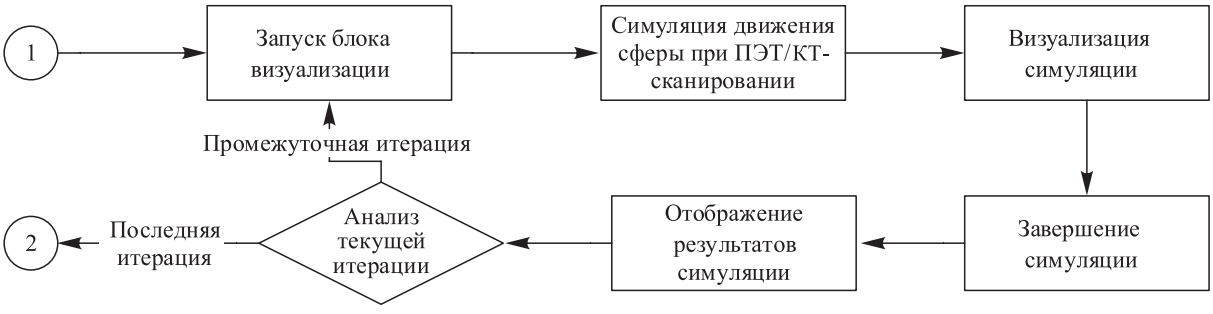


Рис. 4. Схема работы блока визуализации
Fig. 4. Visualization unit operation diagram

В начале функции вычисляется площадь поперечного сечения сферы, задается размер квадратной сетки для накопленных импульсов и инициализируются переменные для подсчета посещенных пикселей. Создается временной вектор t с шагом dt = 0,05 и длительностью total_time, указанной в интерфейсе. Далее рассчитывается положение сферы из траектории ее движения по эллиптической кривой. Значения по осям y и x вычисляются с помощью уравнений:

$$l_y = \frac{A}{2}\sin\left(\frac{2\pi t}{t_\text{р}}\right); \tag{1}$$

$$l_x = 0,225A\sin\left(\frac{2\pi t}{t_\text{р}}\right), \tag{2}$$

где A – амплитуда; t – момент времени; $t_\text{р}$ – период респираторного цикла.

Для накопления сигналов и сохранения положения всех точек поперечного сечения сферы формируются матрицы. После инициализации параметров открывается окно симуляции в полном экране с текстовыми полями для отображения промежуточных данных. Основной цикл по временной сетке обновляет состояние модели: вычисляется положение сферы, определяют-

ся соответствующие индексы в матрицах и в диапазоне пикселей, охватывающих сферу, накапливаются случайные импульсы. Одновременно обновляются счетчики посещенных пикселей, вычисляется накопленная площадь и отображается в графическом окне визуализационного блока модели. Графическое окно обновляется, создавая четыре внутренних окна для различных этапов визуализации, на которых вычисляется синусоида дыхания по формуле [7]:

$$A(t) = \frac{A}{2}\sin\left(\frac{2\pi t}{t_{\text{р}}}\right). \tag{3}$$

Моделируется положение сферы в фантоме: сначала отрисовывается окружность, представляющая цилиндр, затем сфера по текущим координатам, что позволяет наблюдать перемещение поперечного сечения по эллиптической траектории и контролировать площадь. Также выполняется моделирование ПЭТ- и КТ-сканирований [8].

В конце каждой итерации делается краткая пауза для плавного воспроизведения анимации. После завершения всех внутренних итераций модель анализирует результаты текущего цикла. Если итерация последняя, запускается расчетный блок, выводящий итоговые данные о влиянии изменяющихся параметров; иначе модель переходит к следующей итерации.

После завершения визуализации автоматически запускается расчетный блок, схема работы которого для случая опухоли диаметром 19,6 мм при амплитуде колебаний 13,2 мм представлена на рис. 5.

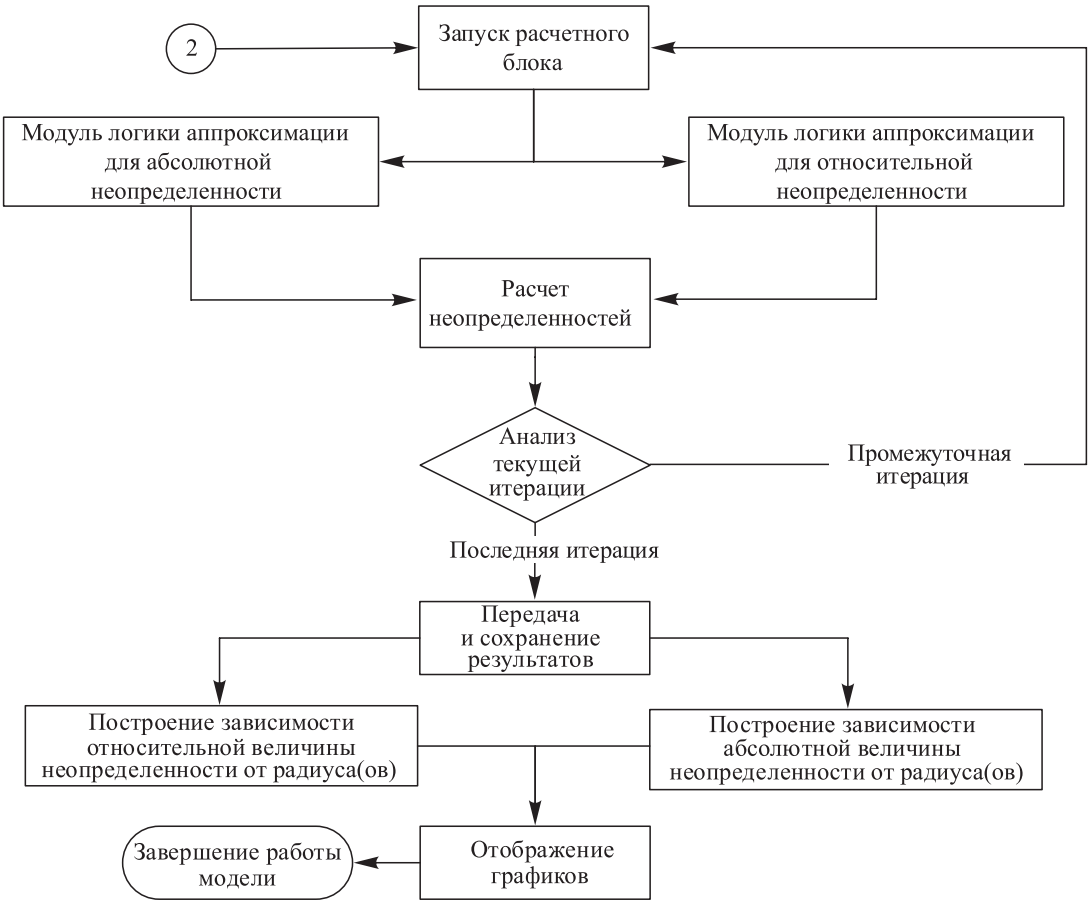


Рис. 5. Схема работы расчетного блока
Fig. 5. Calculation unit operation diagram

Основные расчетные процедуры основаны на экспериментальных данных, полученных с физического фантома. Функции с использованием условных конструкций (if-elseif-else) вычисляют относительные и абсолютные неопределенности. При попадании значений в рамки измеренного диапазона применяется линейная интерполяция, а при выходе за пределы – экстраполяция с экспоненциальными или логарифмическими зависимостями.

После завершения расчетов в рамках симуляционной итерации модель анализирует результаты и, если итерация промежуточная, запускает расчеты для нового набора входных данных. При завершении последней итерации результаты передаются в блок симуляции, в котором сохраняются входные параметры и результаты. Создаются два графика: зависимость относительной неопределенности от амплитуды и зависимость абсолютной неопределенности от амплитуды [9]. После отображения результатов пользователь может закрыть все окна для завершения работы модели или изменить входные данные и запустить симуляцию повторно.

Верификация модели проведена на основе анализа ПЭТ/КТ-изображений, обработанных с использованием программного обеспечения Ellipse v.16.1. В рамках эксперимента контурированы 12 патологических очагов на КТ-изображениях, из них семь – совмещенных с ПЭТ-изображениями, что позволило точно определить их геометрические параметры, включая амплитуду дыхания и размеры патологических очагов. На основании полученных данных вычислены относительные и абсолютные неопределенности, характеризующие погрешности измерений, возникающие вследствие респираторной подвижности [10].

Для каждого из измеренных образований рассчитана погрешность в процентах как для относительной, так и для абсолютной неопределенности. Затем определены средние значения этих погрешностей – для относительной неопределенности средние значения для КТ и ПЭТ, соответственно, составили (2,123 ± 1,051) % и (2,661 ± 0,870) %, для абсолютной – (2,096 ± 0,941) % и (1,992 ± 0,782) %. Анализ результатов выявил, что наибольшие значения абсолютной и относительной неопределенностей для ПЭТ- и КТ-симуляций наблюдаются у больших опухолей, а именно – диаметрами 30,8 и 42,3 мм при средней амплитуде дыхания, а наименьшие погрешности получены у малых опухолей диаметром 13,3 и 11,2 мм при близком значении средней амплитуды дыхания.

Сравнение экспериментальных значений с результатами, предсказанными моделью, позволило оценить точность и воспроизводимость симуляционных алгоритмов. Итоговый анализ демонстрирует высокую степень согласованности между модельными расчетами и экспериментальными данными, что подтверждает корректность и практическую применимость разработанной методики.

Результаты исследований и их обсуждение

Оценка геометрических характеристик исследуемых объектов (сфер) в динамическом режиме показала их несоответствие характеристикам истинных объемов по КТ-сериям до 52 %, по ПЭТ-сериям – до 27 %, в зависимости от размеров моделируемых патологических очагов. Наибольшее несоответствие определено при визуализации сферы диаметром 10 мм. Установлено, что значение несоответствия исследуемых динамических объектов при ПЭТ/КТ-визуализации увеличивается с ростом амплитуды движения (до 23 %). С увеличением объема патологического очага значение несоответствия уменьшается до 4 % (на 40 %) для КТ-визуализации и до 3 % (на 20 %) – для ПЭТ-визуализации в зависимости от амплитуды его движения. Также проведенное исследование показало, что визуализация динамических объектов по ПЭТ-изображениям приблизительно в два раза точнее визуализации по КТ-изображениям.

Заключение

1. Разработана математическая модель респираторных искажений, позволяющая оценить неопределенности визуализации патологического очага на ПЭТ/КТ-изображениях, возникающих вследствие респираторных движений биологического объекта.

2. Применение результатов моделирования в клинической практике радиологических отделений позволит повысить точность определения границ и локализации патологического очага, что, в свою очередь, минимизирует дозовую нагрузку ионизирующего излучения на здоровые органы и риск возникновения рецидива при прохождении пациентом курса лучевой терапии.

Список литературы

1. Murshed, H. Fundamentals of Radiation Oncology: Physical, Biological and Clinical Aspects / H. Murshed, 3rd ed. US: Academic Press, 2019.

2. Gibbons, J. P. Khan's the Physics of Radiation Therapy / J. P. Gibbons, 6th ed. US: Wolters Kluwer Health, 2019.
3. Image Guidance in Radiation Therapy for Better Cure of Cancer / V. Gregoire [et al.] // Molecular Oncology. 2020. Vol. 14, No 7. P. 1470–1491. https://dx.doi.org/10.1002/1878-0261.12751.
4. Хансен, Э. К. Лучевая терапия в онкологии / Э. К. Хансен, М. Роач III. М.: ГЭОТАР-Медиа, 2014.
5. Емельяненко, Е. В. Методика и аппаратные средства оценки количественных характеристик ПЭТ-изображений при исследовании динамических объектов / Е. В. Емельяненко, М. Н. Петкевич, И. Г. Тарутин // Весці Нацыянальнай акадэміі навук Беларусі. Серыя фізіка-тэхнічных навук. 2021. Т. 66, № 4. С. 496–504. https://dx.doi.org/10.29235/1561-8358-2021-66-4-496-504.
6. Development of the Open-Source Dose Calculation and Optimization Toolkit matRad / H. P. Wieser [et al.] // Medical Physics. 2017. Vol. 44, Iss. 6. P. 2556–2568. https://dx.doi.org/10.1002/mp.12251.
7. Simulation on Human Respiratory Motion Dynamics and Platform Construction / Y. Bao [et al.] // Biocybernetics and Biomedical Engineering. 2023. Vol. 43, Iss. 4. P. 736–750. https://dx.doi.org/10.1016/j.bbe.2023.09.002.
8. Optimal Design for Thermodynamic System with OpenModelica and MATLAB/Simulink / Y. Kazuki [et al.] // Asian Modelica Conference, Tokyo, 24–25 Nov. 2022. Tokyo, 2022. P. 93–100. https://dx.doi.org/10.3384/ecp19385.
9. Simulation Tools for Two-Dimensional Experiments in X-Ray Computed Tomography Using the FORBILD Head Phantom / Y. Zhicong [et al.] // Physics in Medicine & Biology. 2012. Vol. 57, Iss. 13. P. 237–252. https://dx.doi.org/10.1088/0031-9155/57/13/N237.
10. PETSTEP: Generation of Synthetic PET Lesions for Fast Evaluation of Segmentation Methods / B. Berthon [et al.] // Physica Medica. 2015. Vol. 31, Iss. 8. P. 969–980. https://dx.doi.org/10.1016/j.ejmp.2015.07.139.

Поступила 01.04.2025 Принята в печать 06.05.2025

References

1. Murshed H. (ed.) (2019) *Fundamentals of Radiation Oncology: Physical, Biological and Clinical Aspects*. US, Academic Press.
2. Gibbons J. P. (2019) *Khan's the Physics of Radiation Therapy*. US, Wolters Kluwer Health.
3. Gregoire V., Guckenberger M., Haustermans K., Lagendijk J., Menard C., Potter R., et al. (2020) Image Guidance in Radiation Therapy for Better Cure of Cancer. *Molecular Oncology.* 14 (7), 1470–1491. https://dx.doi.org/10.1002/1878-0261.12751.
4. Hansen E. K., Roach III M. (2014) *Radiation Therapy in Oncology*. Moscow, GEOTAR-Media Publ.
5. Emelyanenko E. V., Piatkevich M. N., Tarutin I. G. (2021) Methodology and Hardware for Assessing the Quantitative Characteristics of Pet Images in the Study of Dynamic Objects. *Vesti National Academy of Sciences of Belarus. Gray Physics and Technical Sciences*. 66 (4), 496–504. https://dx.doi.org/10.29235/1561-8358-2021-66-4-496-504 (in Russian).
6. Wieser H. P., Cisternas E., Wahl N., Ulrich S., Stadler A., Mescher H., et al. (2017) Development of the Open-Source Dose Calculation and Optimization Toolkit matRad. *Medical Physics*. 44 (6), 2556–2568. https://dx.doi.org/10.1002/mp.12251.
7. Bao Y., Li X., Wei W., Qu S., Zhan Y. (2023) Simulation on Human Respiratory Motion Dynamics and Platform Construction. *Biocybernetics and Biomedical Engineering*. 43 (4), 736–750. https://dx.doi.org/10.1016/j.bbe.2023.09.002.
8. Kazuki Y., Hitoi O., Hirotaka O. (2022) Optimal Design for Thermodynamic System with OpenModelica and MATLAB/Simulink. *Asian Modelica Conference, Tokyo, 24–25 Nov.* 93–100. https://dx.doi.org/10.3384/ecp19385.
9. Yu Z., Noo F., Dennerlein F., Wunderlich A., Lauritsch G., Hornegger J. (2013) Simulation Tools for Two-Dimensional Experiments in X-Ray Computed Tomography Using the FORBILD Head Phantom. *Physics in Medicine & Biology*. 57 (13), 237–252. https://dx.doi.org/10.1088/0031-9155/57/13/N237.
10. Berthon B., Häggström I., Apte A., Beattie B. J., Kirov A. S., Humm J. L., et al. (2015) PETSTEP: Generation of Synthetic PET Lesions for Fast Evaluation of Segmentation Methods. *Physica Medica*. 31 (8), 969 – 980. https://dx.doi.org/10.1016/j.ejmp.2015.07.139.

Received: 1 April 2025 Accepted: 6 May 2025

Вклад авторов / Authors' contribution

Авторы внесли равный вклад в написание статьи. / The authors contributed equally to the writing of the article.

Сведения об авторах

Петкевич М. Н., нач. отдела по инженерному обеспечению лучевой терапии, Республиканский научно-практический центр онкологии и медицинской радиологии имени Н. Н. Александрова (РНПЦ ОМР)

Коровко Д. В., студент, МГЭИ имени А. Д. Сахарова Белорусского государственного университета

Емельяненко Е. В., канд. техн. наук, инж., РНПЦ ОМР

Адрес для корреспонденции

223040, Республика Беларусь,
Минский р-н, а/г Лесной
Республиканский научно-практический центр
онкологии и медицинской радиологии
имени Н. Н. Александрова
Тел.: +375 29 577-09-75
E-mail: MaxPetkevichN@gmail.com
Петкевич Максим Николаевич

Information about the authors

Piatkevich M. N., Head of the Radiation Therapy Engineering Department, N. N. Alexandrov National Cancer Center of Belarus (NCCB)

Korovko D. V., Student, International Sakharov Environmental Institute of Belarusian State University

Emelyanenko E. V., Cand. Sci. (Tech.), Engineer, NCCB

Address for correspondence

223040, Republic of Belarus,
Minsk Region, Agro-Town Lesnoy
N. N. Alexandrov National
Cancer Center
of Belarus
Tel.: +375 29 577-09-75
E-mail: MaxPetkevichN@gmail.com
Piatkevich Maksim Nikolaevich

CC BY

http://dx.doi.org/10.35596/1729-7648-2025-23-3-86-94

УДК 123; 616.894-053.8-09

УМНЫЙ КОНТРАКТ БЛОКЧЕЙНА ДЛЯ ПОДДЕРЖКИ ТЕРАПИИ НЕВРОЛОГИЧЕСКИХ БОЛЬНЫХ

В. А. ВИШНЯКОВ, ЧУЮЭ ЮЙ

Белорусский государственный университет информатики и радиоэлектроники (Минск, Республика Беларусь)

Аннотация. Научной новизной статьи является технология использования медицинского блокчейна на базе умного контракта для поддержки терапии пациента после IT-диагностики болезни Альцгеймера. В комплексе с подсистемой IT-диагностики пациента реализуется обмен данными распознавания степени болезни Альцгеймера. Умным контрактом выполняется автоматизированное управление поддержкой терапии пациентов с диагностированной болезнью Альцгеймера. Разработана функциональная структура подсистемы поддержки терапии (умный контракт в блокчейне Ethereum), включающая инициализацию данных о пациенте, планы лечения, регистрацию медперсонала, управление записями, динамическое переключение планов терапии на основе диагностических данных, контроль предупреждений медперсонала, а также управление выплатами и их распределением. Предлагаемый умный контракт позволяет выработать новый подход к взаимодействию и сотрудничеству в области управления терапией и может использоваться для неврологических больных с другими заболеваниями.

Ключевые слова: умный контракт, терапия пациента, болезнь Альцгеймера, IT-диагностическая подсистема, блокчейн.

Конфликт интересов. Авторы заявляют об отсутствии конфликта интересов.

Для цитирования. Вишняков, В. А. Умный контракт блокчейна для поддержки терапии неврологических больных / В. А. Вишняков, Чуюэ Юй // Доклады БГУИР. 2025. Т. 23, № 3. С. 86–94. http://dx.doi.org/10.35596/1729-7648-2025-23-3-86-94.

SMART CONTRACT BLOCKCHAIN FOR SUPPORTING THERAPY OF NEUROLOGICAL PATIENTS

U. A. VISHNIAKOU, CHUYUE YU

Belarusian State University of Informatics and Radioelectronics (Minsk, Republic of Belarus)

Abstract. The scientific novelty of the article is the technology of using a medical blockchain based on a smart contract to support patient therapy after IT diagnosis of Alzheimer's disease. In combination with the subsystem of IT diagnostics of a patient, the exchange of recognition data for the degree of Alzheimer's disease is implemented. The smart contract performs automated management of therapy support for patients with diagnosed Alzheimer's disease . The functional structure of the therapy support subsystem (smart contract in the Etherium blockchain) has been developed, including initialization of patient data, treatment plans, registration of medical staff, record management, dynamic switching of therapy plans based on diagnostic data, monitoring of medical staff warnings, as well as management of payments and their distribution. The proposed smart contract enables a new approach to interaction and collaboration in the field of therapy management and can be used for neurological patients with other diseases.

Keywords: smart contract, patient therapy, Alzheimer's disease, IT diagnostic subsystem, blockchain.

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

For citation. Vishniakou U. A., Chuyue Yu (2025) Smart Contract Blockchain for Supporting Therapy of Neurological Patients. *Doklady BGUIR*. 23 (3), 86–94. http://dx.doi.org/10.35596/1729-7648-2025-23-3-86-94 (in Russian).

Введение

Технологии блокчейн и умные контракты (УК) благодаря распределенной архитектуре и автоматизированным функциям предоставляют возможность решения проблем обмена информацией между медицинскими учреждениями, а также безопасного хранения данных пациентов и авторизованного доступа к ним [1]. Умный контракт, являясь протоколом, основанным на компьютерном коде, функционирует в системах блокчейна и позволяет автоматизировать проверку, выполнение и запись заранее определенных условий поддержки терапии [2]. Интеграция УК с электронными медицинскими записями [3] улучшает процессы сбора, хранения и авторизованного доступа к данным пациентов.

Использование УК позволяет ввести IT-контроль лечения и изменить качество терапии пациентов с болезнью Альцгеймера (БА) посредством механизмов многостороннего взаимодействия. Предлагаемая в статье подсистема поддержки терапии является продолжением ранней IT-диагностики БА, изложенной в [4], и рассматривает как применение технологий блокчейн и УК для надежного обмена данными, так и возможность автоматизированного ухода за пациентом и многостороннее сотрудничество (между врачами, членами семей пациентов и медицинским персоналом).

Поддержка терапии пациентов с болезнью Альцгеймера

С учетом обычной терапии пациентов с БА [5–9] предложены следующие методы поддержки лечения с записью процедур для мониторинга и контроля на смартфон для обратной связи:

• регулярный уход: мониторинг применения лекарств, поддержка здорового рациона и плана применения медикаментов, восстановление ориентировки посредством напоминаний о времени, месте и людях;

• когнитивные тренировки: упражнения включают сортировку чисел, распознавание фигур, языковые упражнения и др.;

• музыкальная терапия: использование определенной музыки для снижения тревожности, улучшения настроения и когнитивных функций;

• терапия воспоминаний: использование предметов или окружения, связанных с жизнью пациента (фотографии, старые вещи), для стимуляции памяти, улучшения когнитивных и эмоциональных функций;

• осознанная терапия: фокусирование внимания пациента на настоящем моменте с помощью техник осознанности (осознанное дыхание, ходьба), направленных на улучшение психологической регуляции;

• нарративный подход: посредством воспоминаний и организации жизненных событий усиливаются психологическая мотивация и качество жизни пациента.

Структура подсистемы поддержки терапии пациентов с БА приведена на рис. 1. Согласно рисунку, подсистема IT-диагностики анализирует состояние пациента и формирует показатель тяжести заболевания (severityValue). На основе этого значения в модуле «Динамическое планирование» автоматически выбирается подходящий план ухода (обычный или интенсивный). Модуль «Управление медперсоналом» контролирует исполнение выбранного плана и фиксирует выполненные процедуры через «Вспомогательные модули», формируя историю событий. При корректном соблюдении плана производится оплата услуг через «Вспомогательные модули». Таким образом, выход подсистемы IT-диагностики напрямую влияет на выбор терапевтического сценария. Подсистема поддержки терапии позволяет динамически адаптировать ход лечения для пациентов с БА.

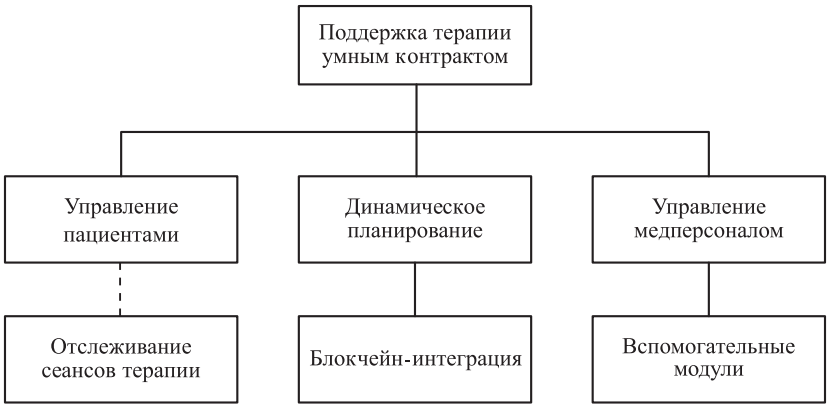


Рис. 1. Структурная схема подсистемы поддержки терапии пациентов с болезнью Альцгеймера
Fig. 1. Structural diagram of the subsystem for supporting therapy of patients with Alzheimer's disease

Структура умного контракта для поддержки терапии пациента с болезнью Альцгеймера

Функциональная структура УК является основой для разработки эффективной поддержки терапии. Приведем структуру проекта, основанного на функциях УК и его взаимодействии с IT-диагностической системой БА.

Взаимодействие УК с IT-диагностической системой. IT-диагностическая система [4] выполняет количественный анализ текущего состояния пациента с БА и выдает значение степени тяжести в диапазоне от 0 до 100 («вероятность заболевания» или «оценка когнитивного снижения»). Это значение вводится в блокчейн УК через интерфейс медицинского персонала, на основании этого УК определяет необходимость переключения на интенсивный или стандартный план ухода. После получения новых диагностических данных УК вызывает функцию updateCarePlanByDiagnosis (uint severityValue), чтобы определить, превышает ли состояние пациента установленный порог (severityThreshold). Если порог превышен, контракт автоматически переключает план ухода на интенсивный. В противном случае остается активным стандартный план ухода. Расписание смен медперсонала, частота осмотров и продолжительность активности автоматически управляются текущим активным планом ухода.

Функциональные модули УК:

• управление медперсоналом и оценка эффективности: медперсонал регистрируется через функцию registerCaregiver(), после чего вся регистрационная информация записывается в блокчейне. Для завершения процесса необходимо выполнение вторичного подтверждения со стороны системного администратора через функцию approveCaregiver (address_caregiver). Медперсонал вводит записи о предоставленных услугах через функцию submitCareRecord(), после чего контракт автоматически рассчитывает количество завершенных услуг, их продолжительность и другие показатели. При достаточном балансе через функцию payCaregiver() осуществляется начисление заработной платы. В случае выявления нарушений или ненадлежащего ухода функция issueWarning() фиксирует количество предупреждений. Если количество предупреждений превышает лимит, контракт автоматически вызывает функцию dismissCaregiver(), аннулируя доступ медперсонала к уходу;

• настройка планов поддержки терапии: УК предусматривает два плана терапии – интенсивный и стандартный – и позволяет владельцу добавлять дополнительные планы через функцию addCarePlan() на разных стадиях болезни. В каждом плане определяются «количество выполнений в неделю» (frequencyPerWeek) и «продолжительность одного выполнения» (durationMinutes). Через интерфейс IT-системы можно наглядно отобразить расписание смен и прогресс выполнения. События, встроенные в контракт (DiagnosisUpdated, PaymentMade), автоматически инициируют логические действия и делают результаты публично доступными. Это обеспечивает надежные и отслеживаемые данные для разных участников (семей пациентов, медицинских учреждений, страховых компаний), минимизируя вероятность ошибок из-за человеческого фактора. Блок-схема смарт-контракта AlzheimerCare представлена на рис. 2.

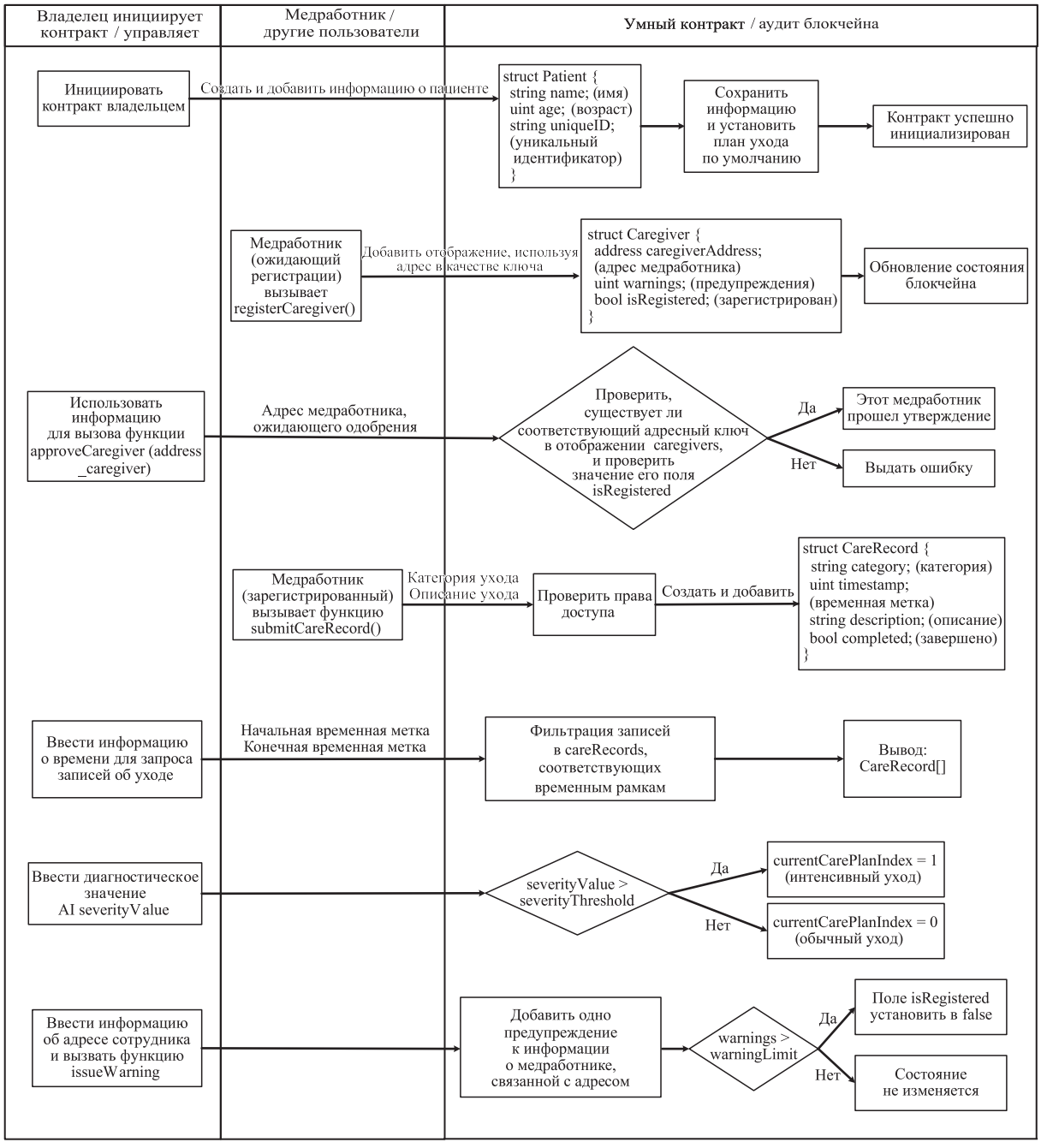


Рис. 2. Блок-схема смарт-контракта AlzheimerCare
Fig. 2. AlzheimerCare smart contract flowchart

Структура умного контракта для пациентов с болезнью Альцгеймера

1. Инициализация контракта.

Вход: имя пациента (name) – строка; возраст пациента (age) – целое число; уникальный идентификатор пациента (uniqueID) – строка.

Процесс: устанавливается владелец контракта (owner) – адрес, который разворачивает контракт. Сохраняется информация о пациенте в структуру Patient. Определяются два плана ухода: Normal Care Plan – стандартный план ухода, High-Intensity Care Plan – интенсивный план ухода. Устанавливается текущий план ухода (currentCarePlanIndex) как стандартный (0).

Выход: контракт развернут. Пациент зарегистрирован. Планы ухода добавлены.

1.1. Регистрация и подтверждение входа. Регистрация входящих.

Вход: адрес водящего (msg.sender).

Процесс: проверяется, зарегистрирован ли данный адрес в системе isRegistered. Если адрес не зарегистрирован, он добавляется в caregivers как новый уходящий:

isRegistered = true.

warnings = 0.

Выход: входящий зарегистрирован.

1.2. Подтверждение ухода.

Вход: адрес входящего (address_caregiver).

Процесс: владелец контракта проверяет, существует ли указанный адрес в caregivers. Если адрес найден, его статус обновляется как подтвержденный (isRegistered = true).

Выход: входящий подтвержден.

2. Добавление записей ухода.

Вход: категория ухода (category) – строка; описание ухода (description) – строка.

Процесс: входящий проверяется на статус регистрации (isRegistered). Если статус зарегистрирован, создается новая запись ухода:

category – категория ухода;

timestamp – текущая временная метка;

description – описание ухода;

completed = true.

Запись добавляется в массив careRecords.

Выход: запись ухода добавлена в блокчейн.

3. Запрос записей ухода.

Вход: начальная дата (_monthStart) – временная метка; конечная дата (_monthEnd) – временная метка.

Процесс: производится фильтрация записей в массиве careRecords на основе временных меток. Возвращаются записи, соответствующие указанному диапазону.

Выход: список записей ухода.

4. Обновление плана ухода на основе диагностики.

Вход: значение диагноза (severityValue) – целое число (0–100).

Процесс: сравнивается severityValue с пороговым значением severityThreshold: если severityValue > severityThreshold, переключается на интенсивный план ухода (currentCarePlanIndex = 1), в противном случае используется стандартный план ухода (currentCarePlanIndex = 0). Вызывается событие DiagnosisUpdated для фиксации обновления.

Выход: план ухода обновлен.

5. Выдача предупреждений и увольнение.

5.1. Выдача предупреждений.

Вход: адрес входящего (address_caregiver).

Процесс: проверяется регистрация указанного адреса в caregivers. Если уходящий зарегистрирован, к его счетчику предупреждений (warnings) прибавляется 1. Если количество предупреждений превышает порог (warningLimit), вызывается процесс увольнения.

Выход: предупреждение выдано.

5.2. Увольнение входящих.

Вход: адрес входящего (address_caregiver).

Процесс: входящий исключается из списка зарегистрированных: isRegistered = false. Вызывается событие CaregiverDismissed.

Выход: входящий уволен.

6. Управление заработной платой.

6.1. Депозит средств.

Вход: сумма депозита (msg.value) – число.

Процесс: владелец контракта переводит средства в контракт. Баланс контракта (balance) увеличивается на сумму депозита.

Выход: средства зачислены.

6.2. Выплата заработной платы.

Вход: адрес входящего (address_caregiver); сумма выплаты (_amount) – число.

Процесс: проверяется, достаточно ли средств на балансе контракта. Выплата осуществляется на указанный адрес. Баланс контракта уменьшается на сумму выплаты.

Выход: выплата произведена.

Реализация умного контракта

Реализация умного контракта для поддержки терапии пациентов с болезнью Альцгеймера осуществляется в приведенной последовательности.

1. Конструктор контракта: `constructor (string memory_name, uint_age, string memory_uniqueID)`. При инициализации контракта записывается информация о пациенте (имя, возраст и уникальный идентификатор), а владелец контракта устанавливается как пользователь.

2. Добавление плана ухода (рис. 3): `function addCarePlan (string memory_category, uint_frequencyPerWeek, uint_durationMinutes, string memory_goals, string memory_requirements) external onlyOwner`. Позволяет владельцу контракта добавлять планы ухода с указанием категории, частоты выполнения, продолжительности, целей и конкретных требований.

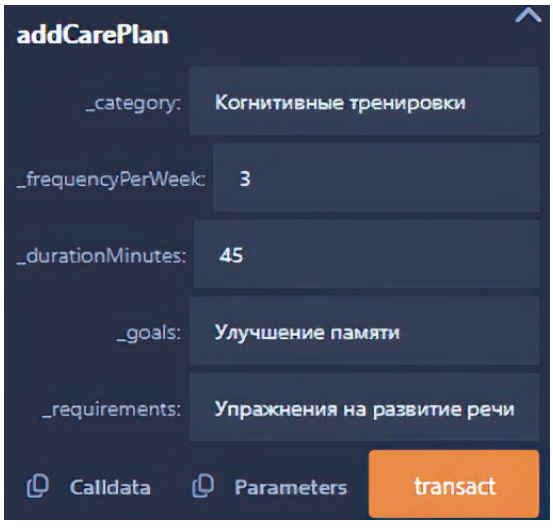


Рис. 3. Запрос добавленного плана ухода
Fig. 3. Request for an added care plan

3. Получение всех планов ухода: `function getCarePlans() external view returns (string[] memory)`. Возвращает подробную информацию обо всех планах ухода. Эта функция отличается от массива `carePlans(index)`, который позволяет получать отдельный план по индексу, тогда как `getCarePlans` обычно возвращает все планы ухода.

4. Регистрация и подтверждение специалистов по уходу: `registerCaregiver`. Специалисты по уходу могут вызвать данную функцию для регистрации. Для этого необходимо открыть MetaMask на адресе кошелька специалиста по уходу и подключиться к Remix. `approveCaregiver`: подтверждение личности специалиста по уходу осуществляется владельцем контракта, что позволяет специалисту выполнять задачи по уходу (рис. 4).

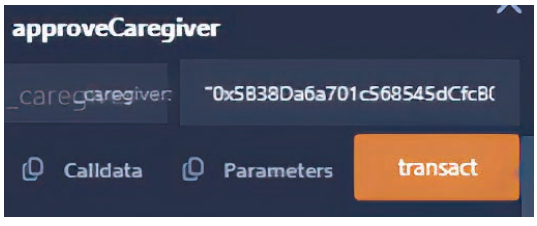


Рис. 4. Подтверждение специалистов по уходу
Fig. 4. Confirmation of care professionals

5. Внесение записи о предоставленном уходе: `function submitCareRecord (string calldata_category, string calldata_description) external onlyRegisteredCaregiver`. Позволяет зарегистрированным

и подтвержденным специалистам по уходу вносить записи о предоставленном уходе, указывая категорию и краткое описание (рис. 5). Введите категорию ухода (например, «Когнитивные тренировки») и краткое описание (например, «Завершена языковая тренировка, пациент отреагировал хорошо»).

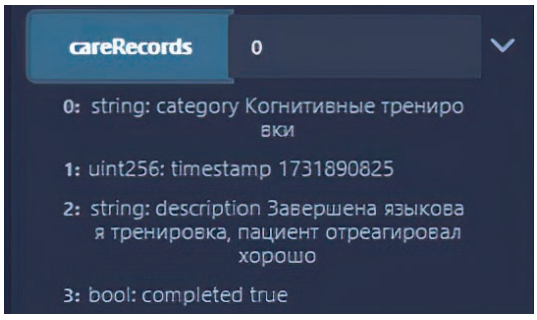


Рис. 5. Проверка выполнения работы по уходу
Fig. 5. Checking the completion of the care work

6. Запрос записей о предоставленном уходе (рис. 6): `function getCareRecordsByMonth (uint_monthStart, uint_monthEnd) external view returns (CareRecord[] memory)`. `monthStart` и `monthEnd` используются в качестве параметров для запроса записей о предоставленном уходе, чтобы указать начальную и конечную даты. В процессе выполнения контракта специалистами по уходу записи автоматически получают временную метку.

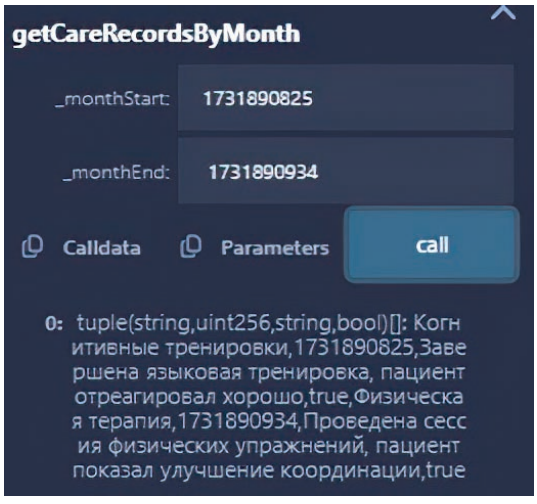


Рис. 6. Запрос записей о предоставленном уходе за период времени
Fig. 6. Requesting records of care provided over a period of time

7. Внесение и выплата зарплаты: `deposit`: владелец контракта может внести средства на контракт для выплаты зарплаты специалистам по уходу, необходимо указать сумму для внесения. `payCaregiver`: выплата зарплаты зарегистрированным специалистам по уходу.

8. Выдача предупреждения специалисту по уходу: `issueWarning`. Если количество предупреждений превышает установленное ограничение, автоматически вызывается функция `dismissCaregiver` для увольнения специалиста.

Рассмотренный смарт-контракт реализован и проверен в частной сети блокчейна Etherium.

Заключение

1. Представлена структура поддержки терапии ухода за пациентами с болезнью Альцгеймера на основе блокчейна и умного контракта, интегрированная с подсистемой IT-диагностики болезни Альцгеймера для реализации обмена данными и автоматизированного контроля. В рамках проектирования умного контракта определены категории планов ухода (зависят от выхода подсистемы IT-диагностики болезни Альцгеймера), их частота, цели и механизмы оценки результатов.

2. Предложены взаимодействия между управляющим контрактом и обслуживающими, что обеспечивает прозрачность и отслеживаемость выполнения задач IT-терапии. Умный контракт реализован и проверен в частной сети блокчейна Etherium.

Список литературы

1. Вишняков, В. А. Технология блокчейн в образовании и ИТ-медицине: модели, алгоритмы, программные средства / В. А. Вишняков, Д. А. Качан. Минск: Республ. ин-т высш. шк., 2023.
2. Blockchain-Enabled Smart Contracts: Architecture, Applications, and Future Trends / Shuai Wang [et al.] // IEEE Transactions on Systems Man and Cybernetics Systems. 2019. Vol. 49, No 11. P. 2266–2277.
3. Long, Jiujun. Design and Implementation of a Blockchain-Based Electronic Medical Records System / Long Jiujun, Li Hongjin, Pan Shiqing // Network Security Technologies and Applications. 2021. No 9. P. 58–60.
4. Вишняков, В. А. Разработка и моделирование сети интернета вещей для IT-диагностики пациентов / В. А. Вишняков, Юй Чу Юэ // Доклады БГУИР. 2024. Т. 22, № 5. С. 104–112. http://dx.doi.org/10.35596/1729-7648-2024-22-5-104-112.
5. A Survey of Multidisciplinary Team-Based Comprehensive Nursing on ADL And Cognitive Function of Alzheimer Disease Patients / Zhao Xuli [et al.] // Chinese Journal of Alzheimer's Disease and Related Disorders. 2023. Vol. 6, Iss. 1. P. 73–76. https://doi.org/10.3969/j.issn.2096-5516.2023.01.013.
6. Tu, Hanlu. A Brief Overview of Product Design to Meet the Needs of Elderly People with Asthma / Tu Hanlu // Science and Innovation. 2016. No 12.
7. Xu, Jing. The Progress of Research on the Care of Patients with Asthma / Xu Jing, Cai Huazhuan // Nursing. 2021. Vol. 10, No 5. P. 465–470.
8. The Progress of Research on the Care of Patients with Asthma / Wang Jinxiao [et al.] // Modern Medicine in China. 2024. Vol. 31, No 12. P. 190–194.
9. Zhang, Huanhuan. The Effectiveness of Using a Model of Joint Management of Patients and Medical Staff in the Care of Elderly Patients with Early-Stage Asthma to Prevent Falls / Zhang Huanhuan // Chinese and Foreign Medicine Research. 2024. Vol. 3, No 1. P. 129–131.

Поступила 08.01.2025 Принята в печать 15.04.2025

References

1. Vishniakou U. A., Kachan D. A. (2023) *Blockchain Technology in Education and IT Medicine: Models, Algorithms, Software Tools*. Minsk, Republican Institute of Higher Education (in Russian).
2. Shuai Wang, Liwei Ouyang, Yong Yuan, Xiaochun Ni (2019) Blockchain-Enabled Smart Contracts: Architecture, Applications, and Future Trends. *IEEE Transactions on Systems Man and Cybernetics Systems*. 49 (11), 2266–2277.
3. Long Jiujun, Li Hongjin, Pan Shiqing (2021) Design and Implementation of a Blockchain-Based Electronic Medical Records System. *Network Security Technologies and Applications*. (9), 58–60.
4. Vishniakou U. A., Yu Chu Yeu (2024) Development and Modeling of the Internet of Things Network for Patients IT Diagnostics. *Doklady BGUIR*. 22 (5), 104–112. http://dx.doi.org/10.35596/1729-7648-2024-22-5-104-112 (in Russian).
5. Zhao Xuli, Wang Huazhi, Sun Yan, Xu Hongjie, Zhang Qian, Chen Xi, et al. (2023) A Survey of Multidisciplinary Team-Based Comprehensive Nursing on ADL And Cognitive Function of Alzheimer Disease Patients. *Chinese Journal of Alzheimer's Disease and Related Disorders*. 6 (1), 73–76. https://doi.org/10.3969/j.issn.2096-5516.2023.01.013.
6. Tu Hanlu (2016) A Brief Overview of Product Design to Meet the Needs of Elderly People with Asthma. *Science and Innovation*. (12).
7. Xu Jing, Cai Huazhuan (2021) The Progress of Research on the Care of Patients with Asthma. *Nursing*. 10 (5), 465–470.
8. Wang Jinxiao, Zhao Qianqian, Zhao Qingqing, Wang Fan (2024) The Progress of Research on the Care of Patients with Asthma. *Modern Medicine in China*. 31 (12), 190–194.
9. Zhang Huanhuan (2024) The Effectiveness of Using a Model of Joint Management of Patients and Medical Staff in the Care of Elderly Patients with Early-Stage Asthma to Prevent Falls. *Chinese and Foreign Medicine Research*. 3 (1), 129–131.

Received: 8 January 2025 Accepted: 15 April 2025

Вклад авторов

Вишняков В. А. выполнил постановку задачи, предложил концепцию использования умного контракта для поддержки терапии пациентов с болезнью Альцгеймера.

Чуюэ Юй провела детализацию разработки умного контракта и его программирование.

Authors' contribution

Vishniakou U. A. completed the task statement, proposed the concept of using a smart contract to support the therapy of patients with Alzheimer's disease.

Chuyue Yu has detailed the development of the smart contract and its programming.

Сведения об авторах

Вишняков В. А., д-р техн. наук, проф. каф. инфокоммуникационных технологий, Белорусский государственный университет информатики и радиоэлектроники

Чуюэ Юй, асп. каф. инфокоммуникационных технологий, Белорусский государственный университет информатики и радиоэлектроники

Адрес для корреспонденции

220013, Республика Беларусь,
Минск, ул. П. Бровки, 6
Белорусский государственный университет
информатики и радиоэлектроники
Тел.: +375 44 486-71-82
E-mail: vish@bsuir.by
Вишняков Владимир Анатольевич

Information about the authors

Vishniakou U. A., Dr. Sci. (Tech.), Professor at the Department of Infocommunication Technologies, Belarusian State University of Informatics and Radioelectronics

Chuyue Yu, Postgraduate at the Department of Infocommunication Technologies, Belarusian State University of Informatics and Radioelectronics

Address for correspondence

220013, Republic of Belarus,
Minsk, P. Brovki St., 6
Belarusian State University
of Informatics and Radioelectronics
Tel.: +375 44 486-71-82
E-mail: vish@bsuir.by
Vishniakou Uladzimir Anatol'evich

CC BY

http://dx.doi.org/10.35596/1729-7648-2025-23-3-95-101

УДК 616-006

ВЛИЯНИЕ ПОЛОЖЕНИЯ БОЛЮСА НА ВЕЛИЧИНУ ДОЗЫ ИОНИЗИРУЮЩЕГО ИЗЛУЧЕНИЯ ПРИ МОДЕЛИРОВАНИИ ОБЛУЧЕНИЯ НА ЛИНЕЙНОМ УСКОРИТЕЛЕ

И. Н. ЧИРКОВА[1,2], М. В. ТУМИЛОВИЧ[2]

[1]Республиканский научно-практический центр онкологии и медицинской радиологии имени Н. Н. Александрова (Минский район, Республика Беларусь)
[2]Белорусский государственный университет информатики и радиоэлектроники (Минск, Республика Беларусь)

Аннотация. Основная цель лучевой терапии – подведение дозы ионизирующего излучения к опухоли. При облучении очагов, расположенных на поверхности кожи либо на небольшой глубине, возникает проблема переоблучения здоровых тканей. С целью оптимизации распределения дозы ионизирующего излучения в практике лечения используются вспомогательные устройства в виде болюсов. При этом важное значение имеет корректная фиксация болюса, обеспечивающая его плотное прилегание. Экспериментально доказано повышение качества дозового распределения при плотной фиксации болюса. Получены численные значения дозы в контрольной точке при различной плотности прилегания болюса с применением разных методик расчета планов облучения.

Ключевые слова: лучевая терапия, мастэктомия, болюс, дозовое распределение, ионизирующее излучение, неопределенность, фантом, компьютерный томограф, радиационный пучок.

Конфликт интересов. Авторы заявляют об отсутствии конфликта интересов.

Для цитирования. Чиркова, И. Н. Влияние положения болюса на величину дозы ионизирующего излучения при моделировании облучения на линейном ускорителе / И. Н. Чиркова, М. В. Тумилович // Доклады БГУИР. 2025. Т. 23, № 3. С. 95–101. http://dx.doi.org/10.35596/1729-7648-2025-23-3-95-101.

EFFECT OF BOLUS POSITION ON THE DOSE OF IONIZING RADIATION IN A LINEAR ACCELERATOR SIMULATION

IRINA N. CHIRKOVA[1,2], MIROSLAV V. TUMILOVICH[2]

[1]N. N. Aleksandrov National Cancer Center of Belarus (Minsk Region, Republic of Belarus)
[2]Belarusian State University of Informatics and Radioelectronics (Minsk, Republic of Belarus)

Abstract. The main goal of radiation therapy is to deliver a dose of ionizing radiation to a tumor. When irradiating foci located on the skin surface or at a shallow depth, the problem of overirradiation of healthy tissues arises. In order to optimize the distribution of the dose of ionizing radiation, auxiliary devices in the form of boluses are used in treatment practice. In this case, correct fixation of the bolus, ensuring its tight fit, is of great importance. An increase in the quality of dose distribution with tight fixation of the bolus has been experimentally proven. Numerical values of the dose at the control point with different bolus fit density have been obtained using different methods for calculating irradiation plans.

Keywords: radiation therapy, mastectomy, bolus, dose distribution, ionizing radiation, uncertainty, phantom, computed tomograph, radiation beam.

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

For citation. Chirkova I. N., Tumilovich M. V. (2025) Effect of Bolus Position on the Dose of Ionizing Radiation in a Linear Accelerator Simulation. *Doklady BGUIR*. 23 (3), 95–101. http://dx.doi.org/10.35596/1729-7648-2025-23-3-95-101 (in Russian).

Введение

Лучевая терапия опухолей, расположенных на поверхности кожи либо близко к поверхности, имеет ряд особенностей, обусловленных спецификой взаимодействия ионизирующего излучения с веществом. Эффективность радиационного луча с медицинской точки зрения коррелирует с распределением дозы по глубине и характеристиками рассеянного излучения. Распределение дозы в радиационных пучках различных видов излучения в зависимости от глубины изображено на рис. 1. Анализ данных показывает, что максимум ионизации для фотонов располагается на некоторой глубине от поверхности, что в реальных условиях может привести к переоблучению здоровых органов и тканей, а также спровоцировать некорректный подвод дозы ионизирующего излучения к очагу опухоли [1].

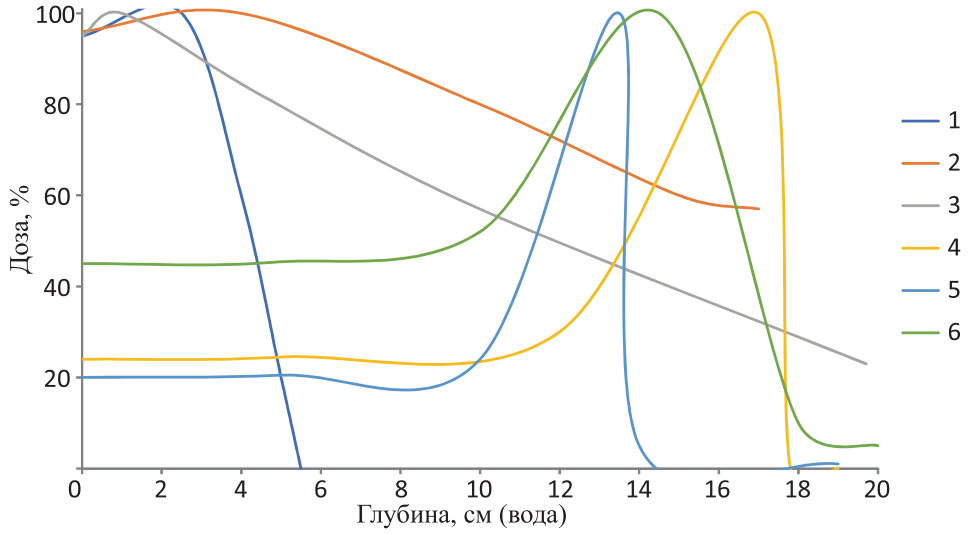


Рис. 1. Зависимость распределения дозы ионизирующего излучения в пучках различных видов излучения от глубины в воде: 1, 3 – электроны и нейтроны с энергией 2,5 и 30 МэВ соответственно; 2 – фотоны с максимальной энергией 25 МэВ; 4 – протоны с энергией 160 МэВ; 5 – ионы ^{12}C с энергией 250 МэВ/нуклон; 6 – пучок π^- мезонов

Fig. 1. Dependence of the distribution of the dose of ionizing radiation in beams of different types of radiation on the depth in water: 1, 3 – electrons and neutrons with energies of 2.5 and 30 MeV, respectively; 2 – photons with a maximum energy of 25 MeV; 4 – protons with an energy of 160 MeV; 5 – ^{12}C ions with an energy of 250 MeV/nucleon; 6 – beam of π^- mesons

На практике в лучевой терапии для минимизации последствий некорректной доставки дозы в опухолевую мишень принято применять вспомогательное устройство в виде болюса. Болюс – вспомогательное устройство, используемое в лучевой терапии для изменения дозы ионизирующего излучения, доставляемой на кожу или поверхностные ткани. Болюс действует как дополнительный слой ткани, позволяя более эффективно доставлять дозу ионизирующего излучения в объем опухоли, которая расположена вблизи либо на поверхности кожи, а также выравнивать дозу по неровным поверхностям тела пациента. Без использования болюса ионизирующее излучение проникает на большую глубину, не затронув предписанный объем облучения. Болюс переносит максимальную дозу ионизирующего излучения ближе к поверхности, что делает его эффективным для лечения поверхностных новообразований [2].

Наиболее частыми локализациями, при лечении которых применяются болюсы, являются:

– поверхностные опухоли (рак кожи, где для лечения требуется высокая поверхностная доза ионизирующего излучения);

– лечение опухолей, расположенных на неровных поверхностях (нос, уши, где болюс применяется для выравнивания объема облучения для равномерного распределения дозы ионизирующего излучения);

– послеоперационное лечение (в случаях, когда опухоль была удалена хирургическим путем и целью лучевого лечения является воздействие ионизирующим излучением на место операции, где могут остаться раковые клетки).

Сегодня в отделениях лучевой терапии активно применяются болюсы, как виртуальные, так и физические, с целью снижения вероятности местных рецидивов после воздействия на ткани ионизирующим излучением. Под физическим болюсом подразумевают тканеэквивалентный материал, который располагается на теле пациента во время сеансов предлучевой подготовки и лучевой терапии. Такой болюс представляет собой пластичный плотный нестерильный материал (например, плотный гель, парафин, силикон), который при облучении обладает свойствами кожи по своим плотностным характеристикам и используется для моделирования тканей человека (кожи) с целью изменения попадающей дозы ионизирующего излучения или глубины ее проникновения.

Виртуальный болюс – это математический инструмент, позволяющий моделировать на поверхности тела пациента либо фантома тканеэквивалентный материал, который вносится в область контурирования [3]. Под контурированием подразумевается очерчивание границ и объемов тканей для дальнейшей оценки на них дозовой нагрузки ионизирующего излучения. Суть моделирования виртуального болюса состоит в построении контуров тела и объема целевого планирования (Planning Target Volume, PTV) вокруг области облучения с заданным расширением, эквивалентным по плотности, в зоне прохождения радиационного пучка [4]. В [5] выдвинута гипотеза о том, что использование виртуального болюса применимо только для методик лучевого лечения 3D (трехмерная конформная лучевая терапия) и IMRT (лучевая терапия с модуляцией интенсивности) и не подходит для методики расчета VMAT (модулированной по интенсивности арк-терапии), так как не представляется возможным сделать отступы по телу.

На этапе предлучевой подготовки физический болюс располагается на теле пациента, после чего плотно фиксируется. Происходит процесс сканирования пациента на компьютерном томографе, затем данные передаются в систему планирования облучения со встроенным программным обеспечением. При помощи встроенных функций отсканированный физический болюс связывается с виртуальным болюсом – математическим инструментом системы.

В последнее время в литературе обсуждается применение физических болюсов для снижения вероятности местных рецидивов, описываются случаи успешного применения вспомогательного материала, когда риск местных рецидивов достаточно велик [6]. В онкологических диспансерах сейчас для лечения используются методики с применением различных материалов болюса, в частности, болюсы, напечатанные на 3D-принтере, а также латунный сетчатый материал, позволяющий уменьшить зазор между кожей и поверхностью болюса, снизить затраты времени на укладку, а также время на облучение [7, 8]. В [9] рассматривалось влияние толщины болюса на величину погрешности в доставке дозы.

Малоизученным остается вопрос влияния воздушного зазора между вспомогательным материалом в виде физического болюса и облучаемой поверхностью. В [10] установлено, что неопределенность линейно возрастает с увеличением зазора, а также зависит от методики расчета плана облучения, толщины болюса и материала, из которого он изготовлен. При лучевой терапии онкологических пациентов в отделениях лучевой терапии используются протоколы на основе фотонов с адаптированным болюсом, и данный метод остается наиболее распространенным в лечении большого числа пациентов [11].

Практически применимым способом оценки воздействия ионизирующего излучения на тело человека является моделирование. С этой целью используется антропоморфный фантом. Антропоморфный фантом – прибор, по своим физическим (плотность, элементарный состав) и техническим свойствам максимально соответствующий телу человека. В литературе описаны случаи использования антропоморфного фантома Catphan 600 (PTV, Freiburg). Прибор представляет собой устройство в виде цилиндра из плотного тканеэквивалентного материала. Тело фантома набрано из отдельных слоев толщиной 1 см со вставками материалов различных плотностей [11]. Также описаны случаи применения антропоморфного фантома Андерсона (RCD, Gardena). Данный фантом – макет тела человека от головы до паховой области из плотного тканеэквивалентного материала. Тело фантома состоит из отдельных вставок толщиной 1 см с различными плотностями.

На выбор фантома влияют его физические характеристики (антропоморфность, доступность, простота укладки и фиксации на нем дополнительных изделий медицинского назначения), а также техническое оснащение отделения лучевой подготовки и лучевой терапии.

Для построения модели воздействия на человека ионизирующего облучения использовалась компьютерная система со встроенным программным обеспечением (КСПО) Eclipse версии 16.3. Программное обеспечение выбиралось для простоты и удобства моделирования процесса облучения и было совместимо со всеми устройствами и техническими модулями оборудования отделения лучевой терапии. При помощи Eclipse реализовываются прецизионные, клинически эффективные и щадящие для пациентов методики облучения, такие как IMRT, 3D, VMAT.

При проведении исследований экспериментально установлена зависимость величины дозы ионизирующего излучения в точке на поверхности модели тела пациента при нарушении заданных условий облучения в части корректности фиксации болюса от различной толщины вспомогательного материала.

Методика проведения эксперимента

В основе оценки методик облучения при лучевой терапии с болюсом применялось оборудование и программное обеспечение отдела по инженерному обеспечению лучевой терапии РНПЦ ОМР им. Н. Н. Александрова. Блок-схема эксперимента по оценке дозы ионизирующего излучения в контрольной точке представлена на рис. 2.

Рис. 2. Блок-схема эксперимента по оценке дозы ионизирующего излучения в контрольной точке
Fig. 2. Block diagram of an experiment to reduce the dose of an ionizing device at a control point

На компьютерном томографе Aquilion Lightning (Canon) получена томограмма антропоморфного фантома Catphan с шагом 2,5 мм, реконструированная до толщины среза в 1,25 мм (рис. 3). Томограмму импортировали в компьютерную систему планирования облучения Eclipse. Антропоморфность фантома позволяет при помощи встроенных инструментов изображать на нем целевые объемы (в данном случае – PTV) как в соответствии с международными рекомендациями по оконтурированию, так и согласно национальным локальным протоколам и требованиям (ICRU).

За объем PTV в фантоме выбран верхний левый квадрат, ширина и длина PTV оконтурирована размером 5 см, что соответствует размеру болюса, который используется на аппарате лучевой терапии при лечении пациентов после мастэктомии, глубина PTV – 5 см. Следует отметить, что планируемый объем мишени на рис. 3 изображен близко к краю фантома, имитируя ситуацию расположения очага опухли непосредственно на коже. В виртуальную модель фантома также включена такая структурная модель, как болюс, оконтурированная с использованием соответствующего программного модуля КСПО Eclipse. Полученный набор структур можно использовать для создания дозиметрических планов облучения любой методикой из применяемых в практике в отделениях лучевой терапии. Далее с помощью встроенного программного обеспечения создавались планы облучения по различным методикам расчета. При этом направление радиационного пучка от линейного ускорителя при каждой методике расчета оставалось неизменным. Также на поверхности фантома была выбрана контрольная точка для оценки дозы, расположенная в центре PTV на поверхности.

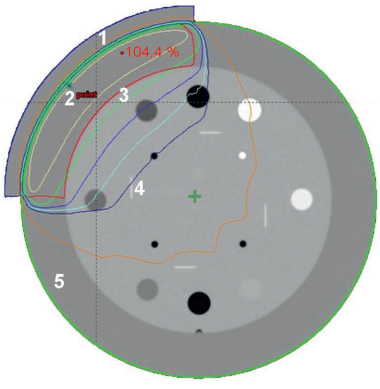


Рис. 3. Скан антропоморфного фантома Catphan с указанием структур: 1 – болюс; 2 – контрольная точка; 3 – PTV; 4 – изодозное распределение; 5 – тканеэквивалентный материал
Fig. 3. Scan of the anthropomorphic phantom Catphan indicating the structures: 1 – bolus; 2 – control point; 3 – PTV; 4 – isodose distribution; 5 – tissue equivalent material

При помощи встроенных инструментов программного модуля Eclipse менялось положение болюса, создавая погрешность его укладки путем искусственно созданного зазора между болюсом и поверхностью фантома с шагом 1 мм от 0 до 1 см. Оценка доставленной дозы и дозового распределения проводилась в выбранной контрольной точке для каждой методики расчета для болюса толщиной 0,5 и 1,0 см [4].

Результаты исследований и их обсуждение

В результате моделирования процесса облучения была установлена зависимость дозы ионизирующего излучения в контрольной точке от величины ошибки в воспроизведении заданных условий облучения в части корректности фиксации болюса при двух различных толщинах вспомогательного материала (рис. 4).

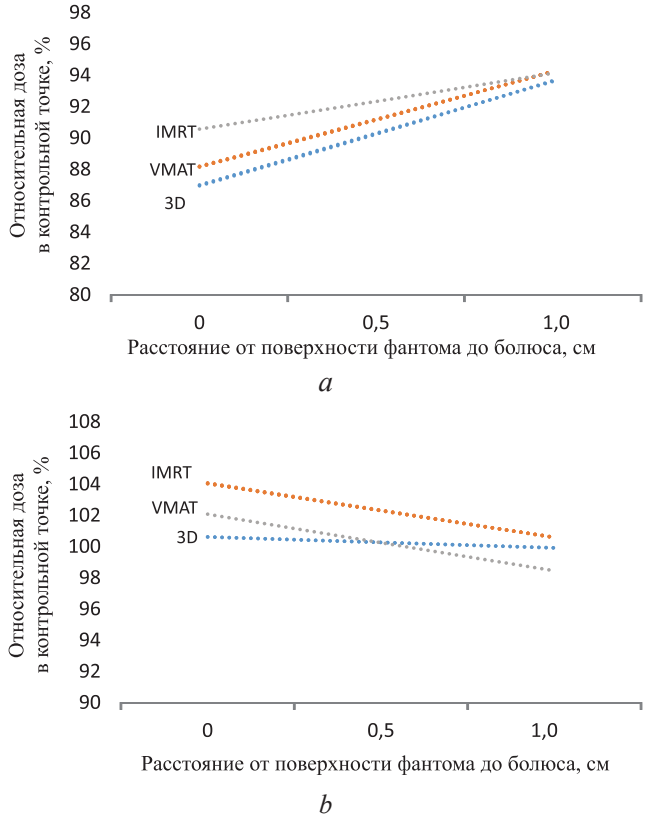


Рис. 4. Зависимость дозы ионизирующего излучения в контрольной точке на поверхности тела фантома от положения болюса толщиной, см: *a* – 0,5; *b* – 1,0
Fig. 4. Dependence of the ionizing radiation dose at the control point on the surface of the phantom body on the position of the bolus thickness, cm: *a* – 0.5; *b* – 1.0

Анализ дозиметрических параметров модели распределения дозы ионизирующего излучения в контрольной точке в объеме мишени, полученных при моделировании планирующей системой Eclipse, показал возможность применения всех методик расчета для облучения поверхностных опухолей как в случае использования болюса толщиной 0,5 см, так и при его толщине 1,0 см. Для каждой из методик с увеличением расстояния между болюсом и поверхностью фантома возникает ошибка, связанная с некорректной фиксацией болюса. При наличии зазора между болюсом и поверхностью в пределах 0–0,5 см значение дозы в контрольной точке линейно увеличивается, но при этом ухудшается и дозовое распределение в объеме PTV, а именно – на 15 % увеличивается максимум ионизации на поверхности, что в реальных условиях ведет к нарушению заданных условий облучения и может привести к незапланированному переоблучению кожных покровов и близлежащих органов и тканей. При наличии зазора в пределах 0,5–1,0 см величина дозы излучения в контрольной точке существенно уменьшается (23 %). Также наблюдаются значительные отклонения в распределении дозы излучения по объему облучаемой мишени (16 %), что приводит к нарушению заданных условий облучения, и как следствие, к отсутствию запла-

нированного эффекта от лечения. Установлено, что доза в контрольной точке имеет большие значения (10–15 %) в случае применения болюса меньшей толщины, что обуславливает использование более толстого болюса для очагов облучения, расположенных на поверхности либо близко к поверхности. Для очагов, расположенных на небольшой глубине от поверхности, целесообразно использовать более тонкий болюс. Для тех случаев, когда болюс не использовался, доза ионизирующего излучения в контрольной точке оказалась меньше запланированной на 73–76 %, что обуславливает использование вспомогательного материала в целом. Плотное прилегание болюса к поверхности облучаемого объема обеспечивает получение значений дозы ионизирующего излучения в контрольной точке на 75–80 % соответствующее запланированному, и охват мишени составляет 95 % от предписанной дозы.

Заключение

1. Результаты исследования подтверждают возможность применения на практике в отделении лучевой терапии методик расчета дозиметрических планов с использованием болюса. При этом основной вклад в ошибки при отпуске дозы вносит некорректное воспроизведение заданных условий облучения при фиксации болюса. Правильная и плотная его фиксация позволяет получить дозу в контрольной точке на поверхности на уровне 75–80 % от запланированной.

2. Определено дальнейшее направление исследования – разработка методики воспроизведения заданных условий облучения в части корректности фиксации болюса.

Список литературы

1. Ратнер, Т. Г. Клиническая дозиметрия. Физико-технические основы / Т. Г. Ратнер, И. М. Лебеденко, 2-е изд. М.: Нац. исслед. ядер. ун-т «Москов. инж.-физ. ин-т», 2022. С. 47–49.
2. Abel, S. Local Failure and Acute Radiodermatological Toxicity in Patients Undergoing Radiation Therapy with and Without Postmastectomy Chest Wall Bolus: Is Bolus Ever Necessary / S. Abel, P. Renz, M. Trombetto // Practical Radiation Oncology. 2017. Vol. 7, Iss. 3. P. 167–172.
3. Kadar-Person, O. A Delfi Study and International Consensus Recommendation: The Use of Bolus in the Setting of Postmastectomy Radiation Therapy for Early Breast Cancer / O. Kadar-Person, H. M. Dahn, N. Bese // Critical Reviews in Oncology // Hematology. 2021. Vol. 164. P. 115–121.
4. Чиркова, И. Н. Анализ дозиметрических параметров модели распределения дозы ионизирующего излучения при моделировании облучения поверхности грудной клетки / И. Н. Чиркова, М. Н. Петкевич, М. В. Тумилович // Медэлектроника–2024: сб. науч. ст. XVI Междунар. науч.-техн. конф., Минск, 5–6 дек. 2024 г. Минск: Белор. гос. ун-т информ. и радиоэлек., 2024. С. 260–265.
5. Ugurlu, T. The Effect of Using Virtual Bolus on VMAT Plan Quality for Left-Sided Breast Cancer Patients / T. Ugurlu, C. Akbas // Applied Radiation and Isotopes. 2022. Vol. 189. P. 179–184.
6. Sakai, Y. Double Enhancement Effect of a Surface Dose with Tungsten Rubber Bolus in Photon Radiotherapy for Keloids and Superficial Tumors / Y. Sakai, H. Monzen, M. Tamuro // Physical and Engineering Sciences in Medicine. 2023. Vol. 46, No 1. P. 179–184.
7. Стуров, А. В. О возможности изменения изодозного распределения в лучевой терапии / А. В. Струков, А. О. Нештак, С. Ф. Четвериков // Медицина и здравоохранение: матер. III Междунар. науч. конф., Казань, май 2015 г. Казань, 2015. С. 25–28.
8. Mieville, F., Post-Mastectomy Radiotherapy: Impact of Bolus Thickness and Irradiation Technique on Skin Dose / F. Mieville, N. Pitteloud, V. Achard // Journal of Medical Physics. 2023. Vol. 46. P. 1–13.
9. Quantify the Effect of Air Gap Errors on Skin Dose for Breast Cancer Radiotherapy / C. Tang [et al.] // Technology in Cancer Research & Treatment. 2024. Vol. 23. DOI: 10.1177/15330338241258566.
10. Гольдман, Е. И. Использование антропоморфного фантома тела человека для осуществления комплексного тестирования технологического процесса лучевой терапии / Е. И. Гольдман, Е. В. Титович // Доклады БГУИР. 2019. № 7–8. С. 133–140.
11. Возможности использования анатомических пластинатов в качестве антропоморфных фантомов // Радиационная гигиена: итоги и перспективы: сб. тр. Всерос. науч.-практ. конф., 18 мая 2022 г. СПб.: Сев.-Зап. гос. мед. ун-т им. И. И.Мечникова, 2022. С. 6–7.

Поступила 08.04.2025 Принята в печать 27.05.2025

References

1. Ratner T. G., Lebedenko I. M. (2022) *Clinical Dosimetry. Physical and Technical Foundations*. Moscow, National Research Nuclear University “Moscow Engineering Physics Institute”. 47–49.

2. Abel S., Renz P., Trombetto M. Local (2017) Failure and Acute Radiodermatological Toxicity in Patients Undergoing Radiation Therapy with and Without Postmastectomy Chest Wall Bolus: Is Bolus Ever Necessary. *Practical Radiation Oncology*. 7 (3), 167–172.
3. Kadar-Person O., Dahn H. M., Bese N. (2021) A Delfi Study and International Consensus Recommendation: The Use of Bolus in the Setting of Postmastectomy Radiation Therapy for Early Breast Cancer. *Critical Reviews in Oncology/Hematology*. 164, 115–121.
4. Chirkova I. N. (2024) Dependence of Dose Distribution of Ionizing Radiation on the Selected Method for Calculating the Irradiation Plan. *Collection of Scientific Articles of the XVI International Scientific and Technical Conference, Minsk, Dec. 5–6.* Minsk, Belarussian State University of Informatics and Radioelectronics. 260–265 (in Russian).
5. Ugurlu T., Akbas C. (2022) The Effect of Using Virtual Bolus on VMAT Plan Quality for Left-Sided Breast Cancer Patients. *Applied Radiation and Isotopes*. 189, 179–184.
6. Sakai Y., Monzen H., Tamuro M. (2023) Double Enhancement Effect of a Surface Dose with Tungsten Rubber Bolus in Photon Radiotherapy for Keloids and Superficial Tumors. *Physical and Engineering Sciences in Medicine*. 46 (1), 179–184.
7. Sturov A. V., Neshtjak S. F., Chetverikov S. F. (2015) On the Possibility of Changing the Isodose Distribution in Radiation Therapy. *Medicine and Healthcare*, *Proceedings of the III International Scientific Conference*. Kazan. 25–28 (in Russian).
8. Mieville F., Pitteloud N., Achard V. (2023) Post-Mastectomy Radiotherapy: Impact of Bolus Thickness and Irradiation Technique on Skin Dose. *Journal of Medical Physics*. 46, 1–13.
9. Tang C., Yuan J., Guo H., Dai Z., Liu B., Xi H., et al. (2024) Quantify the Effect of Air Gap Errors on Skin Dose for Breast Cancer Radiotherapy. *Technology in Cancer Research & Treatment*. 23. DOI: 10.1177/15330338241258566.
10. Goldman E. I., Titovich E. V. (2019) Human Body Anthropomorphic Phantom Utilisation for the Complex Testing of Radiation Therapy Technological Process. *Doklady BGUIR*. 7–8, 133–140 (in Russian).
11. Possibilities of Using Anatomical Plastinates as Anthropomorphic Phantoms. *Radiation Hygiene: Results and Prospects: Collection of Works of the All-Russian Scientific and Practical Conference, 18 May 2022*. Saint Petersburg, North-Western State Medical University named after I. I. Mechnikov. 6–7.

Received: 8 April 2025 Accepted: 27 May 2025

Вклад авторов / Authors' contribution

Авторы внесли равный вклад в написание статьи / The authors contributed equally to the writing of the article.

Сведения об авторах

Чиркова И. Н., вед. мед. физик отдела по инженерному обеспечению лучевой терапии, Республиканский научно-практический центр онкологии и медицинской радиологии имени Н. Н. Александрова; асп., Белорусский государственный университет информатики и радиоэлектроники

Тумилович М. В., д-р техн. наук, проф., Белорусский государственный университет информатики и радиоэлектроники

Information about the authors

Chirkova I. N., Leading Medical Physicist, Department of Radiation Therapy Engineering, N. N. Aleksandrov National Cancer Center of Belarus; Postgraduate, Belarusian State University of Informatics and Radioelectronics

Tumilovich M. V., Dr. Sci. (Tech), Professor, Belarusian State University of Informatics and Radioelectronics

Адрес для корреспонденции

220013, Республика Беларусь,
Минск, ул. П. Бровки, 6
Белорусский государственный университет
информатики и радиоэлектроники
Тел.: +375 17 293-89-11
E-mail: irinagutkovskaya@gmail.com
Чиркова Ирина Николаевна

Address for correspondence

220013, Republic of Belarus,
Minsk, P. Brovki St., 6
Belarusian State University
of Informatics and Radioelectronics
Tel.: +375 17 293-89-11
E-mail: irinagutkovskaya@gmail.com
Chirkova Irina Nikolaevna

CC BY

http://dx.doi.org/10.35596/1729-7648-2025-23-3-102-107

УДК 004.832.519.6

АЛГОРИТМ ШИФРОВАНИЯ ИНФОРМАЦИИ С ПРИМЕНЕНИЕМ ДИСКРЕТНОЙ КВАНТОВОЙ КАРТЫ

А. В. СИДОРЕНКО, Е. А. ВЫСОЦКАЯ

Белорусский государственный университет (Минск, Республика Беларусь)

Аннотация. Разработан комплексный алгоритм шифрования на основе традиционных алгоритмов RSA (криптографический алгоритм с открытым ключом), SHA-3 (алгоритм хеширования переменной разрядности), а также дискретной квантовой карты. Описываются особенности предлагаемого алгоритма и разработанной компьютерной программы на языке C++. Для оценки работоспособности алгоритма было проведено тестирование на трех зашифрованных изображениях, различающихся объемом передаваемой информации. Сравнительный анализ полученных изображений с исходными показал эффективность алгоритма шифрования информации с применением дискретной квантовой карты. Также с помощью этого алгоритма можно выполнять дополнительные операции, связанные с включением информации внутрь зашифрованного изображения, что обеспечит дополнительный уровень защиты при передаче информации.

Ключевые слова: шифрование, алгоритм, дискретная квантовая карта.

Конфликт интересов. Авторы заявляют об отсутствии конфликта интересов.

Для цитирования. Сидоренко, А. В. Алгоритм шифрования информации с применением дискретной квантовой карты / А. В. Сидоренко, Е. А. Высоцкая // Доклады БГУИР. 2025. Т. 23, № 3. С. 102–107. http://dx.doi.org/10.35596/1729-7648-2025-23-3-102-107.

INFORMATION ENCRYPTION ALGORITHM USING A DISCRETE QUANTUM MAP

ALEVTINA V. SIDORENKO, ELIZAVETA A. VYSOTSKAYA

Belarusian State University (Minsk, Republic of Belarus)

Abstract. A complex encryption algorithm has been developed based on traditional RSA (public key encryption algorithm), SHA-3 (hashing algorithm), and a discrete quantum map. The features of the proposed algorithm and the developed computer program in C++ are described. To evaluate the performance of the algorithm, testing was carried out on three encrypted images that differed in the volume of transmitted information. A comparative analysis of the obtained images with the original ones showed the efficiency of the information encryption algorithm using a discrete quantum map. This algorithm can also be used to perform additional operations related to including information inside the encrypted image, which will provide an additional level of protection when transmitting information.

Keywords: encryption, algorithm, discrete quantum map.

Conflict of interests. Authors declare no conflict of interests.

For citation. Sidorenko A. V., Vysotskaya E. A. (2025) Information Encryption Algorithm Using a Discrete Quantum Map. *Doklady BGUIR*. 23 (3), 102–107. http://dx.doi.org/10.35596/1729-7648-2025-23-3-102-107 (in Russian).

Введение

Интенсивное развитие информационных технологий, машинного обучения и искусственного интеллекта выдвигает на первый план вопросы, связанные с квантовой безопасностью. По мере расширения возможностей искусственного интеллекта растет число потенциальных уязвимостей, которыми могут воспользоваться злоумышленники [1, 2].

При эволюции квантовых компьютеров традиционные криптосистемы с публичным ключом и хеш-функции могут быть взломаны, и со временем их устойчивость снизится. Возникает необходимость обеспечения криптографической гибкости с помощью алгоритмов шифрования. Среди них – алгоритмы шифрования RSA (криптографический алгоритм с открытым ключом), ECDA (алгоритм с открытым ключом, использующийся для построения и проверки электронной цифровой подписи), SHA-3 (алгоритм хеширования переменной разрядности), SHA-256 (алгоритм криптографического хеширования), ECDH (протокол Диффи – Хеллмана на эллиптических кривых), DSA (алгоритм цифровой подписи). Даже на достаточно мощном квантовом компьютере конфиденциальная передача данных в перечисленных алгоритмах может быть нарушена алгоритмом Шора (квантовый алгоритм факторизации – разложение числа на простые множители) [3–7]. Следует отметить, что для ускорения генерации хеш-значений в системах с блокчейном квантовые алгоритмы могут использовать алгоритм Гровера (GSA, Grover Search Algorithm – квантовый алгоритм решения задачи перебора) [8]. GSA может быть адаптирован для детектирования хеш-коллизий и в то же время заменить блоки, формируемые в блокчейне, сохраняя интегрируемость передаваемой информации.

В статье представлен разработанный алгоритм шифрования и компьютерной программы на основе RSA, SHA-3 и дискретной квантовой карты. Выполнено тестирование программы с определением ее эффективности. Отмечены основные аспекты применения дискретной квантовой карты совместно с алгоритмами шифрования RSA и SHA-3 для обработки тестовых изображений.

Методы шифрования

Рассмотрим традиционные методы шифрования и шифрование с помощью дискретной квантовой карты. Так, RSA основан на задаче факторизации больших полупростых чисел. В алгоритме используются два ключа: открытый и закрытый. Открытый ключ известен всем, а закрытый хранится в секрете. Только авторизованные пользователи знают, как открыть сообщение. Размер блока, как правило, составляет 446 бит, а размер ключа – более 1024 бит. Для шифрования применяется только один цикл. В цикле шифрования используются три операции: генерация ключа, шифрование и расшифровка. Процесс генерации ключа состоит из следующих этапов.

1. Выбираются два случайных ключа заданного размера: p и q.
2. Вычисляется их произведение: $n = pq$.
3. Вычисляется значение функции Эйлера: $\varphi(n) = (p - 1)(q - 1)$.
4. Выбирается целое число e (простое), которое меньше значения функции Эйлера и является взаимно простым со значениями функции Эйлера, не имеющими общих делителей друг с другом, кроме 1: $1 < e < \varphi(n)$.
5. Вычисляется число d, обратное числу e по модулю $\varphi(n)$: $de \equiv 1(\mathrm{mod}\varphi(n))$.
6. После проведения вычислений получаем: (e, n) – открытый ключ; (d, n) – закрытый ключ [9].

Структурная схема алгоритма шифрования приведена на рис. 1. Для создания электронной подписи применяется алгоритм с использованием закрытого ключа DSA [10].

Дискретная квантовая карта представляет собой структуру элементов и способна обрабатывать данные, характеризущие квантовые процессы. На рис. 2 изображены стандартная и упрощенная двухкубитная схемы дискретной квантовой карты.

Квантовая карта состоит из квантового регистра памяти и квантового регистра данных [4]. В регистре памяти (рис. 2, *а*) находятся n_m кубит. Для ввода переменных в регистре памяти используется m_t кубит за время t. Регистр данных содержит в себе n_x кубит и используется для ввода памяти (скрытых переменных) x_t за время t.

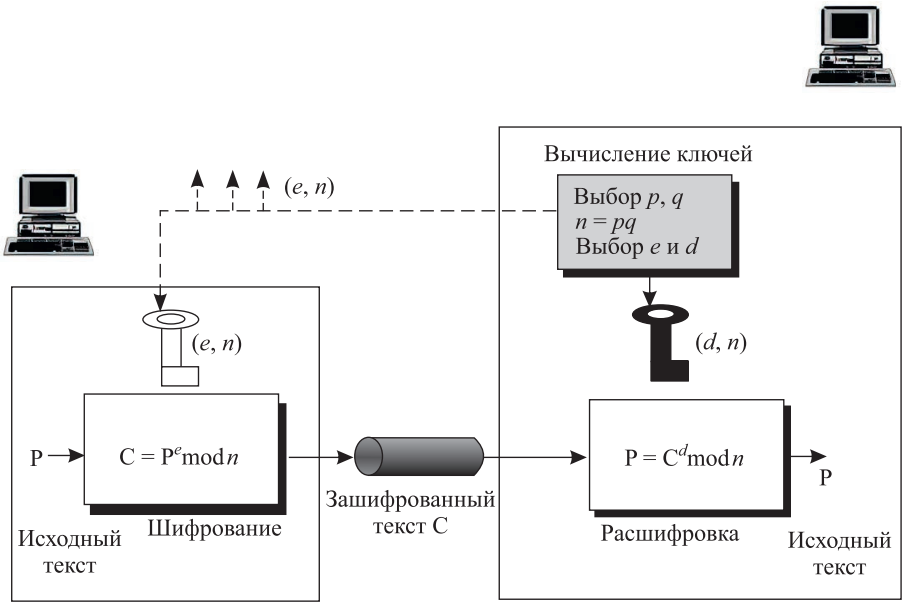


Рис. 1. Стандартная структура алгоритма RSA
Fig. 1. Standard structure of the RSA algorithm

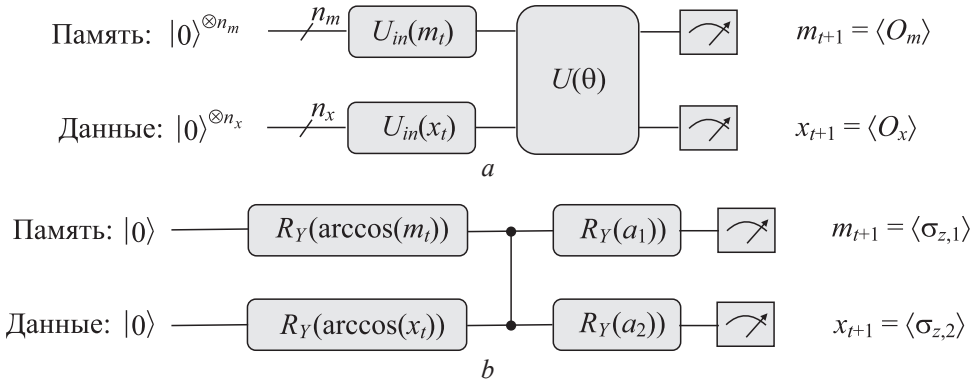


Рис. 2. Схема дискретной квантовой карты: *a* – стандартная; *b* – упрощенная двухкубитная
Fig. 2. Discrete quantum map diagram: *a* – standard; *b* – simplified two-qubit

Унитарный оператор $U(\theta)$ следует после слоя кодирования. Упрощенная двухкубитная схема (рис. 2, *b*) в регистре памяти использует функции $R_Y(\arccos(m_t))$ и $R_Y(\arccos(x_t))$ для кодирования, а также параметрические операторы $R_Y(a_1)$ и $R_Y(a_2)$. Следует отметить, что дискретная квантовая карта чувствительна к начальным условиям, то есть любое, даже небольшое, изменение в них приводит к образованию абсолютно различных текстовых последовательностей или изображений. Применение дискретной квантовой карты для шифрования позволяет обеспечить более высокий уровень безопасности благодаря квантовым свойствам карты. В частности, квантовые ключи, полученные с помощью карт, будут обладать такими параметрами, которые не позволят их копировать.

Процесс шифрования

Создание компьютерной программы включает в себя, прежде всего, разработку структурной схемы алгоритма для процесса шифрования (рис. 3). В схему включены блоки для выполнения дополнительных операций (например, перестановка значений используемой для шифрования последовательности, замена символов, перебор значений и т. д.) при подготовке и реализации процесса шифрования. Кроме того, в структурной схеме имеются блоки, реализующие алгоритмы SHA-3, RSA и дискретную квантовую карту. С их помощью получают зашифрованные данные. Поскольку шифрование в данном случае является обратимым процессом, расшифровка

происходит по той же схеме, но в обратном порядке. Зашифрованная последовательность снова подвергается процессу дополнительных операций, и на выходе получается расшифрованная исходная последовательность.

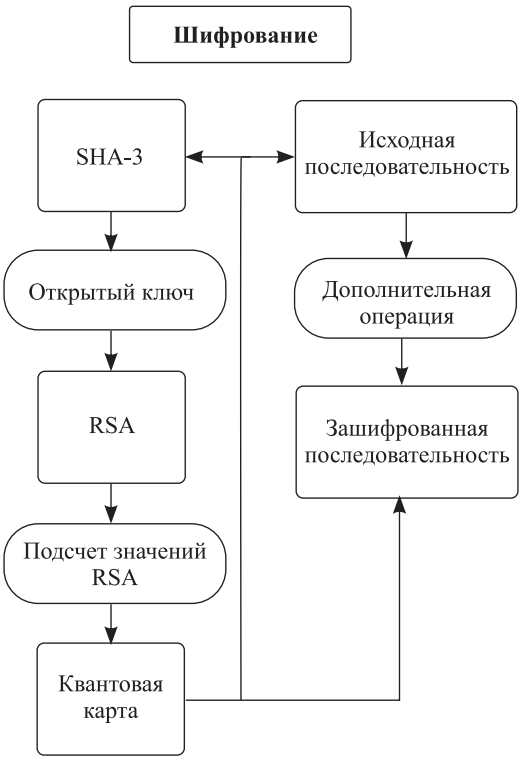


Рис. 3. Структурная схема алгоритма для процесса шифрования
Fig. 3. Block diagram of the algorithm for the encryption process

Реализация и тестирование разработанной компьютерной программы

Предлагаемая компьютерная программа разработана на языке C++. При этом дополнительные библиотеки не использовались. С помощью алгоритма SHA-3 получали хеш-значения информационной последовательности или изображения, которые в дальнейшем применялись для выполнения алгоритма RSA. После этого рассчитывались параметры для дискретной квантовой карты. При тестировании текстовой последовательности для создания ключей в разработанной программе использовали пять пар простых чисел. Результаты тестирования показали, что в случае применения алгоритма SHA-3 к анализируемой последовательности получются 32 хеш-значения. Из этого ряда находили, что наилучшими параметрами среди простых чисел являются $p = 23$ и $q = 21$. Для создания открытого ключа использовали $e = 7$. Внедрение дискретной квантовой карты в разработанный алгоритм обеспечивало дополнительную защиту передаваемой зашифрованной последовательности или изображения. Последние позволили создать максимальную запутанность зашифрованных данных. Расширение возможностей для разработанной компьютерной программы продемонстрированы при шифровании трех тестовых изображений в формате .bmp (рис. 4).

Алгоритм работы компьютерной программы можно представить в виде выполнения следующих этапов.

1. Исходный файл изображения считывается в виде матрицы.

2. Матрица вводится в алгоритм SHA-3, на выходе которого получаются хеш-значения.

3. Хеш-значения вводятся в алгоритм RSA, где рассчитываются открытый и закрытый ключи, производятся шифрование и расшифровка переданной информации.

4. Рассчитанные ключи вводятся в дискретную квантовую карту, в которой осуществляется циклическая запутанность строк и столбцов матрицы. При этом нечетные строки смещаются влево, четные – вправо. Нечетные столбцы смещаются вверх, четные – вниз. Получается измененная матрица изображения.

a *b* *c*

Рис. 4. Тестовые изображения размерами, пиксель: *a* – 512×512; *b* – 256×256; *c* – 467×467
Fig. 4. Test images with dimensions, pixel: *a* – 512×512; *b* – 256×256; *c* – 467×467

5. Применение операции XOR (исключающее ИЛИ) к полученной матрице изменяет значения в ее строках и столбцах, что способствует большей запутанности и позволяет получить зашифрованное изображение.

6. Расшифровка повторяет этапы проведения зашифровки изображения в обратном порядке.

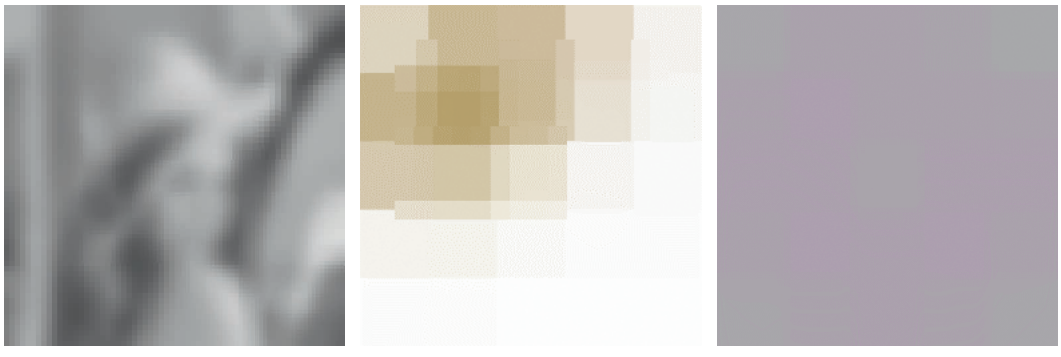
Результаты тестирования разработанной компьютерной программы с использованием тестовых изображений, приведенных на рис. 4, отражены в табл. 1.

Таблица 1. Результаты тестирования компьютерной программы на основе изображений на рис. 4
Table 1. Results of testing the computer program based on images in Fig. 4

Изображение на рис. 4	Размер, пиксель	Время, мс			Хеш-значение	Размер зашифрованных данных, Кб
		обработки матрицы	шифрования RSA	расшифровки RSA		
a	512×512	110	180	150	3a9f–7d8b	200
b	256×256	40	70	60	5f4a–9c2e	80
c	467×467	90	150	120	8c3d–b2f1	130

Как видно из табл. 1, время обработки матрицы передаваемого изображения увеличивается с ростом размера изображения, что связано с необходимостью обработки большего объема данных. Дискретная квантовая карта реализует процесс запутанности, необходимый при шифровании, а с помощью образованных алгоритмом ключей RSA изменяет порядок следования столбцов и строк изображения. Посредством алгоритма SHA-3 успешно получены хеш-значения для всех изображений с обеспечением целостности данных. Время зашифровки и расшифровки данных зависит от размеров передаваемого изображения и с их увеличением возрастает.

Зашифрованные изображения, полученные с использованием тестовых изображений на рис. 4, представлены на рис. 5. Сравнительный анализ тестовых и полученных изображений показал эффективнсть предложенного алгоритма.



a *b* *c*

Рис. 5. Зашифрованные изображения, соответствующие тестовым на рис. 4
Fig. 5. Encrypted images corresponding to the test images in Fig. 4

Заключение

1. Разработан комплексный алгоритм шифрования изображений на основе алгоритмов RSA (алгоритм шифрования с открытым ключом), SHA-3 (алгоритм хеширования) и дискретной квантовой карты. Предложенный алгоритм позволяет эффективно шифровать изображения, что подтверждает возможность сочетания традиционных методов шифрования и шифрования с использованием дискретной квантовой карты.

2. Алгоритм шифрования информации с применением дискретной квантовой карты позволит осуществлять дополнительные операции с изображениями (например, встраивать секретные изображения внутрь изначально зашифрованного). Это обеспечит дополнительный уровень защиты при передаче информации в виде изображений.

3. Работа выполнена в рамках Государственной программы научных исследований «Цифровые и космические технологии, безопасность общества и государства» (подпрограмма «Цифровые технологии и космическая информатика», задание 1.9.3.4).

Список литературы / References

1 Kiktenko E. O. (2018) Quantun Secured Blockchain. *ArXiv*. 3.
2 Bechzadan V. (2017) Vulnerability of Deep Reinforcement Learning to Policy. Induction Attacks. *ArXiv*.
3 Nai Sun Ping (2023) Quantum Discrete -Map. Based Recurrent Neural Networks. *University of Science and Technology of China*. (3), 1–11.
4 Guodong E., Kaivin J., Xiaoling H. (2021) Quantum Logistic Image Encryption Algorithm Based on SHA-3 and RSA. *Journal of Nonlinear Dynamics*. 104, 2807–2827.
5 Sidorenko A. V., Shishko M. S. (2020) Hashing Technical Algorithm Based on SHA-3 and Using Chaotic Maps. *Informatics*. 17 (1), 109–118 (in Russian).
6 Standard FIPS. 186-2. *Digital Signature Standard (DSS)*. 2000.
7 Grover L. A. (1996) First Quantum Mechanical Algorithm for Database Search. *Proc. of 28 Annual ACM Simposium on the Theory of Computing. Philadelphia. USA*. 212–219.
8 Shor P. W. (1992) Polynomial Time Algorithms for Prime Factorization and Discrete Logarithms on a Quantum Computer. *SIAM J. Comput*. 26, 1484–1509.
9 Calderbank H. (2007) *RCA Cryptosystem. History Algorithm. Primes*. Chicago, University of Chicago, Department of Mathematics. 1–7.
10 Jamak H. (2006) Digital Signature Algorithms. *Expert Conference Trends in the Development of Machinery and Associated Technology: 10th International Research*. 1–5.

Поступила 29.01.2025
Received: 29 January 2025

Принята в печать 11.04.2025
Accepted: 11 April 2025

Сведения об авторах

Сидоренко А. В., д-р техн. наук, проф., проф. каф. физики и аэрокосмических технологий, Белорусский государственный университет

Высоцкая Е. А., студент, Белорусский государственный университет

Адрес для корреспонденции

220030, Республика Беларусь,
Минск, просп. Независимости, 4
Белорусский государственный университет
Тел.: +375 29 339-14-30
E-mail: sidorenkoa@yandex.by
Сидоренко Алевтина Васильевна

Information about the authors

Sidorenko A. V., Dr. Sci. (Tech.), Professor, Professor at the Department of Physics and Aerospace Technology, Belarusian State University

Vysotskaya E. A., Student, Belarusian State University

Address for correspondence

220030, Republic of Belarus,
Minsk, Nezavisimosti Ave., 4
Belarusian State University
Tel.: +375 29 339-14-30
E-mail: sidorenkoa@yandex.by
Sidorenko Alevtina Vasilevna

ДЛЯ ЗАМЕТОК